Introduction to Experimental Economics

by

Glenn W. Harrison†

November 2002

1. Two Canonical Experiments ............................................................ -1-
   A. Double Auctions and Smith ......................................................... -2-
   B. Preference Reversals and Grether and Plott ................................... -8-

2. Literature ................................................................................... -10-
   A. Davis and Holt ........................................................................ -10-
   B. Friedman and Sunder .................................................................. -13-
   C. Hey ..................................................................................... -15-
   D. Kagel and Roth ......................................................................... -17-
   E. Omissions ............................................................................. -20-
   F. Comparative Evaluation ............................................................ -23-

3. Selected Issues ............................................................................ -24-
   A. Individual Characteristics .......................................................... -24-
   B. Hypothetical Responses ............................................................. -25-
   C. Incentives .............................................................................. -26-
   D. Recruiting .................................................................................. -28-
   E. Repetition .............................................................................. -29-
   F. Deception and Clarity ................................................................. -29-
   G. Field Experiments .................................................................... -31-
   H. Computerization ..................................................................... -34-
   I. Journals and Conferences ........................................................... -35-
   J. Pedagogy .................................................................................. -36-
   K. Smart Markets ......................................................................... -36-

4. Classics .................................................................................... -38-

5. Conclusions .................................................................................. -39-

References ...................................................................................... -40-

† Dewey H. Johnson Professor of Economics, Department of Economics, Moore School of Business, University of South Carolina, USA. E-mail contact: HARRISON@MOORE.SC.EDU. Prepared for the Nineteenth Biannual Workshop of the Economy and Environment Program for Southeast Asia (EEPSEA), Manila, November 2002.
Experimental economics is now a well developed field with economics. It has its own methodological practices, its own association that runs annual conferences, its own journal, and now its own Nobel Laureates. A review of two of the earliest experiments, which laid the foundations for the two main areas of experimental economics, provides a simple introduction to the field (section 1). There have been several general textbooks in the field, and a review of each provides a useful way to introduce the field as a whole (section 2). Selected comments of an informal nature on several areas in the field might help new researchers (section 3). Finally, I offer a personal list of “classics” that every student of the field should read.

This introduction does not focus on issues in environmental economics per se. A review of experimental work on the use and abuse of the contingent valuation method is contained in Harrison [2002]. Additional materials, and many of the papers cited here, are available at the web page for my presentation: http://dmsweb.moore.sc.edu/glenn/Manila/.

1. Two Canonical Experiments

The two main areas of experimental economics employ different approaches to controlling valuations of subjects. The first is “induced values,” where the experimenter directly controls valuations by imposing them on the subject as explained below. The second is to “elicit homegrown values” using rules which give the subject an incentive to reveal them truthfully.

---

1 Smith [1982].
2 The Economic Science Association: http://www.economicscience.org
5 There have also been several additional books on specialized issues: for example, Palfrey [1991] on experiments in political science, and Kagel and Levin [2002] on bidding behavior in common values settings. The annual compilation Research in Experimental Economics (JAI Press) has started to have “theme issues” since the journal Experimental Economics has become the main outlet for specialized papers on method.
**A. Double Auctions and Smith**

Smith [1962] is still regarded as the first modern paper in the field of experimental economics. It introduced the notion of induced values. Subjects taking on the role of buyers are told that any units that they buy in the market will be redeemed at a specified rate into money, and that they may keep the difference between the price and that redemption value as profit (or “consumer surplus”). Subjects taking on the role of sellers are told that any units that they sell in the market will cost them the specified amount to produce, and that they may keep the difference as profit. The attached schedule illustrates how this might be presented to subjects. This series of redemption values and costs generates a demand and supply schedule as shown on the following diagram, with each unit contributing one “step” to each schedule. In this manner the experimenter can be said to “know” the demand and supply schedule before seeing what prices are set by the specific trading institution. Thus, one can test whether a market equilibrium is achieved without having to use the observed data to jointly (a) estimate the demand and supply schedules, and (b) test if prices are at the equilibrium of demand and supply. The key insight here is that experimental control gives us (a) under certain assumptions, later spelt out clearly by Smith [1982].

Once values and costs are induced in this manner, one can define the trading institution over these fictitious commodities. In Smith [1962] the institution was a canonical representation of “a competitive market,” the double oral auction (DA). This auction is said to be “double” because both sides of the market can offer to trade at the same time: buyers call out prices that they are willing to purchase one unit at, and sellers call out prices that they are willing to sell one unit at, or one side or the other calls out that they will accept an offer. Trading lasts for a fixed period of time, and then a new round begins where everyone is “re-initialized.” The results of Smith’s first DA experiment are shown: notice the tendency to convergence towards the competitive equilibrium.

---

6 There are precursors, just as there were people in North America before Christopher Columbus “discovered” it. But, as the saying goes, after Columbus North America was never “lost” again!
### Redemption Values

**Individual Buyer Number:** __________

**Period:** __________

<table>
<thead>
<tr>
<th>Unit</th>
<th>Row</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1st unit redemption value</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>purchase price</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>profit</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>profit + 15¢ commission</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>2nd unit redemption value</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>purchase price</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>profit</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>profit + 15¢ commission</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>3rd unit redemption value</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>purchase price</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>profit</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>profit + 15¢ commission</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>4th unit redemption value</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>purchase price</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>profit</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>profit + 15¢ commission</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>5th unit redemption value</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>purchase price</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>profit</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>profit + 15¢ commission</td>
</tr>
</tbody>
</table>

**Total period earnings:**

### Costs

**Individual Seller Number:** __________

**Period:** __________

<table>
<thead>
<tr>
<th>Unit</th>
<th>Row</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>selling price</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>cost of 1st unit</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>profit</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>profit + 15¢ commission</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>selling price</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>cost of 2nd unit</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>profit</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>profit + 15¢ commission</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>selling price</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>cost of 3rd unit</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>profit</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>profit + 15¢ commission</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>selling price</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>cost of 4th unit</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>profit</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>profit + 15¢ commission</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>selling price</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>cost of 5th unit</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>profit</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>profit + 15¢ commission</td>
</tr>
</tbody>
</table>

**Total period earnings:**

*Figure 1. Redemption and Cost Incentive Forms*
Figure 2. Aggregated Limit Values
FIGURE 14.9. Monopoly experiment M5x.
There are many variants on this design. The institution became computerized, and many details of implementation turned out to be important (e.g., how does one handle the “open book” of back orders?). Multiple units could be offered, rather than one unit at a time. And “call market” features could be added, to generate starting prices. Perhaps the most interesting variants dealt with issues in industrial organization: how many buyers and sellers were needed to generate competitive outcomes. Early research on DA established that only 2 or 3 symmetric traders on each side were needed to generate competitive outcomes, a far cry from the “large numbers” so often taught to students. In fact, with only one seller, DA can generate outcomes at or below the competitive equilibrium, as shown on the last picture, confirming what is referred to as the “Coase Conjecture.”

Different trading institutions have been studied by experimental economists. The most important are:

• Double Auction (DA) – such as share markets;
• Posted Offer (PO) – such as retail markets, where there is a fixed “take it or leave it” price offered to sellers;
• Negotiated prices -- telephone markets, such as found in wholesale settings;
• Sealed Bid auctions – such as government procurement, a deceptively large fraction of GDP, and more recently in privatization auctions;
• Real Time auctions – such as used by art auction houses, some real estate markets, or emerging on-line institutions (eBay);
• Call markets – such as the auction procedures used to establish opening bids in stock markets; and
• “Smart markets” – where pricing, packaging and network logistics are aided by computer (e.g., electricity markets).

One of the strengths of experimental economics is the relative ease with which such a wide variety

---

7 When “posted offer” trading institutions are employed, much more familiar monopoly pricing outcomes obtain when there is only one potential seller. However, the fact that only one seller is observed to supply the market does not mean that the market is not contested, and in fact one observed competitive outcomes in such settings. See Harrison and McKee [1985] for examples of monopoly and contested markets in experiments.
of institutions can be studied, and their performance evaluated in different environments.

B. Preference Reversals and Grether and Plott

The other side of experimental economics has to do with what are now known as “homegrown values,” or the valuations that individuals put on real goods and services. This is to differentiate them from experimenter-induced values. The idea is that they are “grown at home” in the sense that they derive from the preferences of the individual, rather than the valuations imposed by the experimenter. When people are asked to choose between two lotteries, or asked to state a willingness to pay for some real commodity, the experimenter is eliciting the homegrown values of the subject.

One of the most important experiments in this vein was by Grether and Plott [1979]. They studied a “preference reversal” phenomenon long examined by psychologists. The underlying objects being valued are simple lotteries: the experimenter offers to pay you $X if a coin comes up heads, and $Y if it comes up tails. The phenomenon in question refers to a situation in which the same individual exhibits a preference for lottery A over lottery B when asked directly which one he would prefer, and then exhibits a preference for B over A when asked to place a value on each lottery. In the first case, the subject is simply asked to state which lottery he would prefer, and then exhibits a preference for B over A when asked to place a value on each lottery. In the first case, the subject is simply asked to state which lottery he would prefer to play out, knowing that there is some probability that he will have to live by this choice. In the second case the subject is asked to state a valuation for the lottery, using a procedure we explain below. The valuation elicited for lottery A can be compared to the valuation elicited for lottery B, and a preference implied from the valuation. It is the inconsistency between the direct preference and the implied preference that is said to be a reversal.

One of the most important features of this study was how it identified “sneaking suspicions” that economists had with the way that psychologists had demonstrated this phenomenon experimentally, and proceeded to address each one. For example, the lack of financial incentives for truth-telling: psychologists simply asked hypothetical questions, and provided subjects with no
incentive to respond truthfully. This can be addressed easily, and did not make the phenomenon go away (this issue is a recurring, and depressing, theme of some relevance for students of contingent valuation). Similarly, the lack of procedures to guard against strategic mis-representation of valuations: psychologists did not use what we now call incentive-compatible procedures for eliciting valuations. This was addressed by using the “BDM procedure” developed by Becker, DeGroot and Marschak [1964] for eliciting the certainty-equivalent of a lottery: endow the subject with the lottery, ask them to state a minimum selling price for it, draw a random buying price from a pre-determined distribution, give the subject the buying price if it exceeds the stated selling price, and let the subject play out the lottery if the buying price equals or falls below the stated selling price. If the subject believes that the buying price is indeed drawn without any influence from the stated selling price, it is in the subject’s best interests to reveal that price honestly.

Grether and Plott [1979] designed an experiment which used procedures that met all of the criteria that economists would want, and still observed preference reversals. This generated a flood of interest in psychology experiments, and modifications of theory, and effectively paved the way for what is now known as “behavioral economics.” Although the prospect theory of Kahneman and Tversky [1979] is justly acknowledged as the forefather of this work, in the sense of offering an alternative hypothesis for such phenomena, Grether and Plott [1979] is the study that really legitimized this line of research for economists.

---

8 In my experience many subjects do not believe this unless one uses a physical randomizing device, such as a bingo cage or many-sided die, to generate the buying price. Subjects often harbor beliefs that their selling price somehow influences what they perceive as a “counter offer” in a bargaining game. Thus one might be wary of experiments that implement the BDM procedure in a computerized environment, since the subject does not see how the buying price is generated. This point also illustrates how properties of an institution that are clear to the experimenter might not be so clear to the subject (and that it is not entirely the subject’s fault, given the extent of deception in many experiments they encounter in “Psych 101 classes”).

9 There is no question that Plott was deserving of the 2002 Nobel Prize, along with Smith and Kahneman.
2. Literature

A. Davis and Holt

Davis and Holt [1993] is clearly positioned as the best introductory text of the group. The scope of topics covered is wide, with good discussions of markets, auctions, bargaining games, normal-form and extensive-form games, free riding, individual decision-making, and statistical methods. There are many examples of instructions, typically presented in a more standard form than found in the original sources. These examples help the reader follow the actual experiment, as distinct from the more abstract theory being tested. At just less than 600 pages this is a fat text. However, the chapters are complete as stand-alone readings. For classroom use they could be rearranged or some could be skipped.

For non-specialists the first chapter provides an introduction to the methodological arguments for using experimental methods. It also contains a useful typology of different types of experiments, discusses procedural considerations, introduces terminology, and presents several trading institutions. In this chapter Davis and Holt discuss the feature that distinguishes experimental economics from other types of experiments: the inclusion of salient rewards. They point out that it is critical for participants to receive rewards that correspond to the incentives assumed in the relevant theory and that changes in decisions should have a noticeable effect on the rewards.

The second chapter on decisions and games introduces a variety of theoretical topics. It then provides an introduction to the way in which these fundamental economic theories can be operationalized in the lab. This is a useful place to begin pedagogically, since jumping straight into a market experiment can require the explanation of several methodological innovations at once. A key feature of modern work in experimental economics has been the attempt, generally successful, to keep experiments tied to theory. Thus the experiments are not offered as “black behavioral boxes,” but as systematic and controlled tests of operationally meaningful theories.

The third and fourth chapters focus on competitive price theories and provide balanced
introductions to market experiments, presenting the double-auction and posted-offer institutions. The double-auction, where buyers compete by raising price bids and sellers compete by reducing price offers, is similar to the institution used in major securities markets. It has proven itself a fascinating institution in terms of efficiency and robustness. It is also one of the hardest to formally model using the tools of non-cooperative game theory, which may explain some of the longevity of the mystique it holds for experimenters. The posted-offer institution, where sellers offer “take it or leave it” prices, is common in retail markets. It has been the most popular vehicle of experimental explorations of industrial organization questions.

The fifth chapter covers bargaining, where two parties negotiate (often with no formal institutional structure), and auctions, where one (or more) trader(s) on one side of a market bargains with a number of traders on the other side of a market. It would be fair to say that this chapter could easily have been two, since it is here that the links between formal theory and experiments have been the tightest. It is also here that many of the methodological battles have taken place (e.g, with respect to “fairness” or “altruism” in bargaining, and with respect to “payoff dominance” and “risk attitudes” in auctions). This chapter covers many topics and discusses a diverse set of institutional structures. Nevertheless, the coverage is excellent for the scale attempted, with all of the major developments well surveyed.

The sixth chapter deals with public goods, externalities and voting. In contrast to the private goods discussed in previous chapters, public goods are characterized by nonrivalrous consumption (multiple users can simultaneously consume the same unit of the good) and nonexcludability (it is not feasible to prevent consumption by those who fail to pay). The exposition is admirable given the size of the literature. Topics covered include the voluntary contribution mechanism; the common pool resource mechanism; effects of repetition, experience, group size, and communication; and provision points. Although the use of these particular experimental methods has exploded in political science, it is appropriate to provide a minimal introduction to this topic in a book oriented towards economics applications.
Chapter seven covers asymmetric information, such as occurs in a market where product quality is known to sellers but unknown to buyers, or in an asset market where different traders poses different information about asset value. The questions considered are interesting and applicable to many situations. However, large jumps are required over different institutions, and this is a difficult chapter to cover in an integrated manner. For classroom use, parts of this chapter could be better used in conjunction with the chapters discussing asset markets and (repeated) games.

A chapter on individual decision-making under uncertainty closes the substantive coverage. It includes an excellent discussion of a variety of topics on violations of expected utility maximization and individual decision making in general. The exposition of probability-triangle representations and the discussion of problems in preference elicitation are particularly valuable pedagogically.

A final chapter provides a summary of experimental results and discusses the relationship among theory, experiments, and natural economic environments. It also discusses experimental design and statistical procedures. This is an excellent concluding chapter. It provides the reader with often neglected tools needed to critically evaluate the experimental literature and design new experiments. Of particular interest is the discussion of one of the major goals of experimentation: to induce the structural assumptions of the theory being examined in order to observe the behavioral outcome. The ability to control the structure in line with theory, and manipulate key variables independently, is the real advantage of the experimental method.

The book has a good index and relatively complete listings of references at the end of each chapter. Most chapters also include appendices containing subject instructions for experiments or more technical detail related to chapter material (such as calculation of equilibria or derivation of optimal strategies).
B. Friedman and Sunder

Friedman and Sunder [1994] positions itself as a “primer” for students and researchers and is designed to provide a “readily accessible, self-contained summary of experimental method and technique.” In the Preface the authors point out that most experimentalists have learned their craft through apprenticeship. As such, the book attempts to emphasize concrete procedures. At just over 200 pages, the book does not attempt to be encyclopedic so much as broad and readable.

The first chapter discusses the role of experiments in economics. It introduces, in simple non-technical language, the interaction between theory and empirics, the differences between experimental and non-experimental data, and some of the purposes of experiments. Friedman and Sunder point out that data obtained through controlled experiments allow more reliable inferences to be made than happenstance data (generated by uncontrolled processes).

The second chapter covers the principles of economics experiments. It provides a concise discussion of induced-value theory and the necessary conditions to achieve control over subjects’ preferences. These include monotonicity (subjects prefer more reward to less and do not become satiated), salience (a subject’s reward depends on his actions), and dominance (changes in subject utility come predominantly from the reward medium). Friedman and Sunder argue that salience differentiates surveys from controlled experiments and cite evidence indicating that the use of salient rewards tends to increase the reliability of results.

The third chapter on experimental design introduces the concepts of treatment and control. It discusses several types of basic designs and suggests situations where each might be applicable. It also provides some practical advice which may be particularly useful for those trying to learn about experiments without the benefit of an experienced mentor.

Chapter four addresses issues related to the use of human subjects. The authors point out that economic theories have traditionally made convenient assumptions about human behavior. Yet, the questions of how humans actually learn, form expectations, adapt, choose strategies, and make decisions are crucial to many theories. Through experimental studies, the researcher can observe
human behavior in the context of the theory being examined. The chapter also discusses several practical design choices such as what type of subjects to use, how many subjects, commissions and rewards to be paid, instruction, recruitment, and record-keeping.

Laboratory facilities are the subject of chapter five. The choice between manual and computer models is briefly discussed. Next, laboratory setup in terms of space, layout, and furnishings is considered. Several specific university labs are mentioned in terms of both their hardware and software. The chapter also includes a discussion of possible procedures to use in random number generation.

Chapter six provides a concise practical checklist for actually conducting an experiment. It discusses the importance of maintaining a record book and of conducting pilot experiments. Details such as lab setup, registration, instruction of subjects, handling questions, and paying subjects are all considered.

Chapters seven and eight address data analysis and reporting of results, respectively. The authors discuss a variety of statistical approaches that are commonly used to analyze experimental data, but that may not be familiar to researchers accustomed to analyzing other types of data. They also discuss the importance of anticipating the data analysis while planning the experiment. A fairly standard structure for academic paper organization in presented in chapter eight. The authors note that complete documentation of procedures is particularly important in experimental studies.

Chapter nine provides a mix of concluding historical and methodological perspectives. Of particular interest is the section on the divergence of experimental economics from experimental psychology. In addition to differences in the use of salient rewards, Friedman and Sunder point out that experimental economics studies tend to focus on behavior in specific institutions, and to develop new theories based on a strong tie to existing theories. Differences in attitudes toward deception between the two fields are also discussed: the tolerance toward deception of any kind is noticeably less tolerated in economics-based experiments.

Several appendices are included. These provide reading lists from experimental economics
courses, sample instructions, useful forms for conducting experiments, *Econometrica* guidelines for submitting experimental papers, and a partial list of researchers and labs involved in experimental research.

Each chapter concludes with an applications section presenting actual experiments related to the chapter material. These applications sections tend to rely on theories that are not well developed in the text, and thus require the reader to have a deeper level of economics knowledge than the rest of the text. This treatment is fitting with the authors’ goal of presenting the experimental method to economists who are accustomed to other types of research, but does not serve to make the material accessible to researchers in other related fields trying to learn the experimental method.

*Hey [1991]* is designed as a textbook for advanced undergraduates, graduate students, and general economists hoping to learn about the experimental field. Accordingly, the style is non-technical, assuming a general knowledge of theory and econometrics, but not of the experimental method. At 242 pages, the book does not offer a comprehensive review of work to date, but rather has a coverage intended to provide the nonspecialist with an orientation to the field.

Part I considers methodological issues, such as the use of experimental methods in economics and practical details of laboratory experimentation. An amusing discussion of empirical theory testing using field data is presented. The objective is to convince the reader that nearly all areas within economics can and should be subject to experimental investigation. The practical details chapter is particularly well written, with discussions on such topics as deciding whether to computerize the experiment, run pilot studies, how to analyze data, and deciding where to send finished papers.

Part II provides a detailed review of experiments on individual decision-making under risk. Subjective expected utility theory is introduced, and experimental evidence of violations of the theory’s implications are discussed. Several alternative theories of individual decision making are
then presented. Experiments testing these theories are surveyed. Although the exposition is centered on Hey’s own published papers in the area, the goal is to take the reader through the many steps necessary in designing an experiment. This research project has recently culminated in an important and widely-cited study, Hey and Orme [1994]. Thus this section provides rare insight into the “black box” of the years of experimental design behind a modern classic.

Part III continues the examination of experiments on individual decision-making. There is a great deal of economic theory which takes as maintained the assumptions of one theory of decision-making (typically the standard assumptions of Von Neumann and Morgenstern). Tests of those foundations are the topic of part II; tests of theories building on those foundations are the topic of part III. The specific topics are again drawn from Hey’s own published research in search and optimal consumption under uncertainty. Although the focus is again idiosyncratic, the goal of providing insight into the design process is accomplished. A sketchy survey of the work of other experimenters in this general area is also provided.

Part IV attempts a broad brush overview of work on interactive economic behavior. Simple normal form and extensive form games are examined, bargaining experiments are mentioned, and then auctions and markets are reviewed. The sketchy nature of the review is tempered by the inclusion of numerous editorial comments questioning certain issues of design. Although virtually all of the questions are posed and then left unanswered, they offer the type of self-doubts that all experimenters carry with them about some of the work reviewed here. For example, concern that the modern rush to computerize experiments may lead to a breakdown of trust between the subjects and the experimenter is discussed (p. 179). In some situations, the experimenter may need to make specific design choices in order to ensure credibility.

Part V offers brief parting thoughts on other experimental work and pointers to current trends. The existence of a variety of experiments on the interface between economics and other disciplines (including accounting, education, history, law, politics, psychology, public choice, and sociology) is also briefly mentioned.
Overall, this book is a pleasant surprise. It is highly readable, written in a light-hearted, often even amusing, style. Although the coverage may be seen initially as a bit patchy, in the end it offers a better insight into the nature of experimental economics than any of the books reviewed here. The reason is undoubtedly the attention to design detail in parts II and III, and the willingness to “editorialize” on main findings in part IV rather than catalogue details of long series of experiments.

D. Kagel and Roth

Kagel and Roth [1995] does attempt to be encyclopedic, and largely succeeds. The book consists of eight chapters: the first is an introductory chapter, and the others are surveys of areas in which there have been a concentration of experiments. The latter chapters are each written by a different experimenter and are designed to provide a critical review of the work to date. At 721 pages, this text wins the weight sweepstakes of the four, but this is offset by the ambitious nature of the project.

Chapter one by Roth provides some historical perspective, discusses the uses of experiments, and introduces the following chapters. This chapter is quite complete by itself as a review of the field. For the non-specialist, it does an excellent job of introducing each research area, providing context, and discussing general methodological issues.

Chapter two by Ledyard reviews public goods experiments. It begins by briefly mentioning some everyday public goods and asking how well existing institutions can provide these. It then asks a perhaps even more interesting question of whether we can discover other organizational arrangements that might better serve the interests of society. This up-front motivation is a refreshing addition. The remainder of the chapter focuses on experiments using the voluntary contribution mechanism, a common situation resulting when voluntary contributions to a public good are socially desirable but individually bad. Several papers are covered in depth, and general conclusions about what has been learned from these experiments are presented. Although a great deal of material is presented in the chapter, the coverage is still patchy. Many contributions to the literature are
surveyed in a superficial manner or ignored altogether.

Chapter three by Ochs looks at coordination problems. In this literature games with multiple equilibria are studied. In such situations, a particular equilibrium may arise as a focal point because players have come to expect play consistent with this equilibrium from others. Three particular types of these experiments are discussed in some detail. Although the discussion is representative of the published literature, this topic seems a bit out of place. The issues might have been better dealt with in a broader setting of “experimental games” in general, or as an addition to the asset markets chapter.

Chapter four by Roth reviews bargaining experiments. This work involves many issues and complexities. As Roth points out, the related experimental results defy easy summary. The chapter does an admirable job of stressing the tight link between formal theory and experimental design. However, much of the text is written at a fairly sophisticated level, and some narrow issues receive excessive coverage. Given that this is a handbook with one stated goal of providing the nonspecialist with a review of work to date, it may be that much of this material will be inaccessible to at least some portion of the intended audience.

Chapter five by Holt is perhaps the most balanced chapter of the handbook, reviewing the large literature related to industrial organization. This literature is important historically, since most of the earliest contributions and findings of experimenters resulted from an interest in industrial organization issues (such as competition, collusion, and market efficiency). It is also important substantively, since these findings are likely to provide a fertile basis for policy work in future years. The chapter is written at an appropriate level and offers several very useful discussions of general topics nonspecialists should consider (for example, a warning to “generalize with caution”).

In chapter six Sunder provides a useful review of experimental work in asset markets. Although there is little in the way of basic introductory material for the nonspecialist, the chapter does an excellent job of summarizing important laboratory findings. Major sections include discussions of informational efficiency, futures contracts, price “bubbles,” learning, investment and
public policy, lab modelling, and comparisons of field and lab data. This last topic is particularly
important. Solid arguments are made as to how lab environments can provide for tests that cannot
be conducted with field data. The chapter concludes that researchers can gain insights into poorly
understood asset market behavior by using lab markets in conjunction with theoretical analysis and
field data.

Chapter seven by Kagel reviews auction experiments. Because the work covered in chapters
five and six utilize the double auction institution, the focus here is predominantly on one-sided
auctions (those with many buyers and one seller or with many sellers and one buyer). Discussions of
private value auctions, where bidders know their own personal value of the item being auctioned
with certainty, and common value auctions, where bidders share a common value but are
differentially informed as to what that value is, comprise the two main sections. A third section
contains shorter discussions of other related topics. Each subsection contains a useful summary of
important findings and evaluative comments. Kagel provides a careful discussion of the issue of
“payoff dominance”. This problem results when the reward structure is relatively flat near the
optimum. In such situations, subjects likely do not have adequate incentives to adjust their behavior
toward the optimum. This issue has provided a battlefield for several groups of experimenters, and
many mines remain in wait for the unwary. Particular attention is also given to work on the
“winner’s curse,” a judgment failure in common value auctions in which bidders fail to adjust for the
fact that they only win in cases where they have the highest estimate of the underlying value. This
failure results in winning bids that produce low (or negative) profits on average. This chapter is quite
detailed in its coverage, yet it is written at a level so as to be accessible to the nonspecialist.

Chapter eight by Camerer on individual decision making is one of the most interesting. This
discussion is organized around two main topics: judgment and choice. A variety of results from
psychology research are first introduced. While some fields, such as accounting, have a well-
established tradition of considering behavioral research with hypothetical (or no) reward structures,
the value of applying these theories with real economic incentives is a topic of heated debate.
Camerer contends that it is useful for economists to study individual decision-making and decision-making errors because errors may affect economic efficiency. While many readers may not agree with these points of view, all should agree that Camerer provides a clear articulation of the issues. Whether one sees the discussion as convincing or as finally providing a well-defined target, one thing is clear: the examination of how findings from behavioral research fare in the presence of economic incentives promises to be a topic of continuing interest. The chapter does an excellent job of presenting a large amount of material, which may have been completely unfamiliar to the reader, in a clear manner. It is also a rich source of references to the related literature.

The book contains one comprehensive author index. Detailed references follow each chapter.

E. Omissions

What is missing in these books? The answer to this question should be constrained by looking only at sub-fields that are sufficiently mature or important to warrant coverage in book form. Many sub-fields are still very much on the frontier, and it would be inappropriate to expect these books to engage in speculation as to where they will eventually fit in the broader experimental field.

None of the books cover the relationship between lab experiments and field experiments. A popular criticism of lab experiments as opposed to field experiments has been the almost exclusive use of student subjects. There are some exceptions in which non-student subjects have been recruited, such as Hoffman, et al. [1993], but the common procedure has been to use the “convenience samples” that are most convenient to academics.

The vague criticism of this practice, that “students are not real people,” may be easily dismissed. A more legitimate criticism is that these people may not be representative of some other group of people we wish to generalize to. The problem with student samples is thus that they display a relatively narrow range of socio-demographic characteristics. If we ignore the possibility that
individuals self-select into the student population, or more specifically into our experiment sample, due to these characteristics, it is a simple statistical matter to correct for the non-representative nature of convenience samples (see Harrison and Lesley [1996]). However, the remaining problem of the narrowness of the characteristics of student samples should not be minimized: see Blackburn, Harrison and Rutström [1994] for an example of the problem.

In general, lab experiments are likely to be used as cheap and fast alternatives to field experiments. The pecuniary cost of substantial field surveys is apparent, but an additional hidden cost arises due to difficulties in field logistics. In particular, making simple design changes as experience is gained and information is gathered is difficult. Conversely, design modifications are particularly cheap in the lab. Because the lab does offer these advantages, lab experiments are useful as complements to field analyses. Plott [1987] offers a stimulating review of this use of lab experiments.

All of the books appear to be dominated by the “induced values” tradition in experimental economics. While this may be historically important in market experiments, it is not a defining feature of the experimental method per se. A completely different purpose of conducting experiments is to elicit values that the subject brings to the lab. These are often referred to as “homegrown values,” to indicate that they are values that the subject brings to the lab rather than has induced at the lab.

Although many experiments have tangled indirectly with the notion of homegrown values, the focus has traditionally been on these as a contaminant rather than as a legitimate object of inquiry. However, many experimenters do talk about the role of homegrown beliefs or expectations about other players’ behavior in games. For example, much discussion has been centered on the homegrown value of “fair” outcomes in bargaining games and on the homegrown value of altruism in public goods experiments. Some experiments have also focused on the differences obtained when eliciting values for lab-created objects (e.g., preference reversal experiments).

In recent years, a major sub-field of experimental economics has emerged in which the goal is to truthfully elicit homegrown values for specific goods. The motivation for this line of inquiry is
Similarly, he may reveal a value that is higher than his true value if he believes he can re-sell the item in the field and make a profit.

The need to evaluate the reliability of field survey instruments. These instruments have developed as important tools in marketing and environmental damage assessment. Controversy surrounds the extent of “hypothetical bias” that many suspect arises from the use of survey questions that ask for a hypothetical economic payment, sometimes for a hypothetical commodity. Lab experiments provide a sensible way to evaluate these surveys. Simplified and scaled down versions of the field instrument can be tested in a lab. The results can then be compared to the valuations elicited when real economic payments are required for real commodities. Reviews of the issues and major findings can be found in Cummings and Harrison [1994] and Harrison and Rutström [2003].

This sub-field is methodologically important because of several problems that can arise when eliciting rather than inducing values. Because these problems do not typically arise in induced value experiments, the issue of dealing with them is also left unaddressed in the books. However, unless these problems are properly handled, serious errors in interpreting experimental results are likely to be made.

The first problem is the possibility that lab responses may be censored by field opportunities (Harrison [1992]). In other words, a subject may not reveal his true valuation for a good in the lab if it exceeds the value he could obtain the same good (or an acceptable substitute) for outside the lab. The second problem is the possibility that certain elicitation procedures themselves might influence the subject’s perception of field substitutes. In other words, the subject may believe that other subjects have better information about field substitutes, and after observing others’ valuation decisions, he may use that information to revise his censoring value. Because these institutions provide feedback to subjects about others’ valuations, English auctions and repeated Vickrey auctions can engender uncontrolled affiliation of beliefs about the field (see Rutström [1996]). The third problem is that these institutions might also influence the subject’s perception of the commodity itself. In particular, when the subject is unfamiliar with the commodity and views the commodity as

---

10 Similarly, he may reveal a value that is higher than his true value if he believes he can re-sell the item in the field and make a profit.
a lottery with uncertain characteristics, he may use others’ valuations to revise his subjective beliefs. These uncontrolled affiliated beliefs about characteristics may lead to a loss of control over “which” commodity is being valued (see Harrison, Harstad and Rutström [1995]).

F. Comparative Evaluation

Each of the four books discussed here has its own particular strengths. The question of which is most useful is dependent on the individual reader’s background and their motivation for reading a book on experimental economics. We therefore evaluate the books from several likely perspectives.

As a comprehensive classroom text, Davis and Holt is the best choice. The wide variety of topics covered makes this book useful as a complete introduction to the field. Davis and Holt is also probably the best single resource for researchers in other disciplines seeking to learn about the experimental method and some of its applications.

For a new experimenter with a solid background in economics, Friedman and Sunder provides a concise source of highly applicable “how-to’s”. This book is likely to be the most useful for economics researchers who do not have the opportunity to learn the method first-hand from a seasoned experimentalist.

For those with some exposure to the field seeking to become well-versed in a particular area, Kagel and Roth is a valuable resource. Each chapter represents a comprehensive literature review of a particular area, even though the critical depth of the reviews is uneven.

For those seeking to learn about the field in general without necessarily becoming an expert in conducting experiments or reading detailed surveys, Hey is the best choice. It provides a readable perspective that may be of most interest to economists who hope to maintain an informed appreciation of a variety of areas, such as experimental, without getting too deeply involved in the nitty-gritty.
3. Selected Issues

A. Individual Characteristics

It is a simple matter to ask experimental subjects to answer a series of socio-demographic questions. Information on age, sex, education, race, nationality and marital status are relatively basic. Information on household and personal income levels are also valuable, and rarely skipped by subjects if the data is requested in interval form. We ask for household and personal income separately because most of our subjects are students, and this distinction is relatively important in that case.

One obvious reason for collecting these data is to see if the sample responses vary in terms of the characteristics of the subjects. For example, is there a sex effect? Many studies examine this issue without collecting information on other characteristics, raising doubts about any conclusions being confounded by other variables. What if women have lower incomes, or more education, or are older?

A less obvious reason for collecting these data is to check to see if the sample if representative of the population to which one wants to make inferences. Even if it is not, one can make simple statistical corrections to generate predictions that are conditioned on the behavioral responses measured in the lab for given characteristics, and then weighted by the prevalence of those characteristics in the field. For instance, if women are more risk averse than men, a lab experiment with a sample of 70% men and 30% women should be able to detect this with a large enough sample size. But when it comes time to drawing inferences about the average risk aversion of the population, which we will assume is a happily balanced one of 50% men and women, one cannot simply apply the sample average from the lab. Instead, a weighted average of the sex-conditioned averages is appropriate. Recognition of this simple point makes the results from lab experiments much more relevant for drawing field inferences.11

---

Finally, there is some evidence that recruitment methods can affect the sample characteristics of subjects that volunteer for experiments. For example, recruiting subjects with a huge, fixed show-up fee and low variable earnings will tend to lead to less risk averse subjects turning up compared to a situation in which subjects are recruited with a low, fixed show-up fee and high variable earnings, even when the expected payoff is the same. Apart from testing for sample-selection effects, one can only control for the effects of different recruitment methods if information on individual characteristics is collected.\textsuperscript{12}

For all of these reasons, and since it does not add significantly to the time needed for an experiment, we always have each subject complete a questionnaire with a dozen or so standard questions.

\textbf{B. Hypothetical Responses}

For some reason the debate over the validity of hypothetical bias persists long beyond it’s relevance for any serious policy debate. In the context of hypothetical surveys of environmental goods with a significant \textit{a priori} non-use value, one cannot easily deliver the good to lab subjects. Hence there is an innate need for hypothetical surveys, and legitimate questions about the ability to calibrate for hypothetical bias.

The problem is that many of the experiments testing the validity of expected utility theory use hypothetical responses. And a great deal of effort is allocated defending the validity of those responses. This is a red herring, which should simply be discarded. The cost of paying subjects is a budgetary factor, but hardly a convincing one.

Most economics journals are reluctant to accept experimental papers that do not have salient rewards, in the sense of using real monetary incentives. This is largely due to early traditions in experimental economics, when it was important to differentiate the academic activity from the non-salient opinion polls and surveys long-used by psychologists. This may be the wrong reason for the

\textsuperscript{12} See Rutström [1998] for an extended illustration of this effect.
right policy, but it is worth being aware of.

C. Incentives

One of the hallmarks of the field of experimental economics has been the use of real monetary payoffs for decisions. This is referred to, following Smith [1982], as salience: the idea that rewards should be linked to the choices. The fact that they are linked says nothing about the strength of the incentives to tell the truth.

Consider this issue in the context of the payoffs to a subject bidding in a First-Price sealed bid auction. Assume for simplicity that the subject is risk-neutral, and that all other subjects are risk-neutral. Let the number of bidders be 4, and let induced valuations be drawn from a uniform distribution defined over the interval [0, 1000], where all monetary units are pennies. Consider a subject that has been told that his valuation is 700 cents. The standard symmetric Nash Equilibrium bidding function derived by Cox, Roberson and Smith [1982] allows us to trace out the expected income to this subject from alternative bids. The result is shown below. The vertical axis shows expected income in pennies, and the horizontal axis shows possible bids. There is a horizontal line in the picture where expected income is zero, and a vertical line in the picture where the true valuation of 700 cents equals the bid. Obviously, given the rules of the First-Price auction, the subject ensures an expected payoff of zero by bidding equal to his valuation, whether or not he wins the auction. Bidding above the valuation of 700 cents ensures an expected loss. The payoff
function has an apparent optimum, and indeed it is at the predicted Nash Equilibrium bid of 525 for this subject.

However, we can focus on the expected payoff alternatives closer to the optimum to assess the nature of the problem facing this subject if we view him as applying an algorithm to this task. This is shown in the next picture, which deliberately uses the same scaling of the vertical axis to contrast the cognitive difference between the “big picture” and the “local picture.” The second picture actually displays all possible bids that generate at least 55 cents expected profit; the optimal bid generates an expected payoff of 60 cents. So the implication is that bids that deviate by as much 65 cents from the optimal bid of 525 still generate expected payoffs within 5 pennies of the optimum. This is simply a statement of the facts of the experimental task facing the subject; it is another matter how one interprets observed deviations by subjects.13

The implications for the interpretation of experimental behavior are apparent, but open-ended, since control has been lost. von Winterfeldt and Edwards [1982; p.620] posed the issue succinctly:

If a subject can gain little by emulating the ant rather than the grasshopper, what will he or she do? Random behavior is one possibility. A more likely one arises because most experiments imply nonmonetary payoffs, and subjects are likely to pay attention to them. They may minimize effort (e.g., by ignoring the stimulus or hunting for a very simple strategy for responding). They like to be right and so may not respond appropriately to asymmetric payoffs. They may look for hidden meanings in the experiment and thus develop self-instructions. Or they may just be bored.

---

13 Harrison [1989][1992] raised this issue, arguing that it was inappropriate to reject Nash equilibrium bidding theory if the subjects got within a “reasonable” tolerance of the optimum in terms of expected income.
D. Recruiting

Subjects for experiments are typically recruited from student populations, in the nature of a convenience sample for academics. There are three types of ways of recruiting subjects in general use:

- Students are “commandeered” in large classes, and asked to “volunteer” for an experiment that will be completed in class. Such students have virtually zero opportunity cost of participating in the experiment, and may be reluctant to refuse lest it reflect on them in terms of grades. The advantage of this method is that it avoids sample selection effects within the population that has selected to attend university. Since the subjects have such a low opportunity cost, such experiments often dispense with a show-up or fixed participation fee. The disadvantage is that the students may have no interest in the task, and be rationally uninformed about it.

- Students are recruited in class, or at popular venues such as student cafeterias, or by poster announcements with large dollar signs. They are then contacted to attend a specific session, usually with a fixed show-up fee.

- Students are recruited into a permanent database of potential subjects, typically online, and contacted to attend one or more experiments over time.

Each of these methods has strengths and weaknesses.

The weakness of the first method is that it includes people who could care less about the task. This may not seem a problem, until one considers the relevance of their responses for the field. If people only undertake certain tasks because they are disposed to them (e.g., people that can’t stand the sight of blood tend not to become surgeons), odd lab behavior in that task by people that have no intention to undertaken it in the field is not particularly interesting. Of course, if the task or the theory is general, as many claim economic theory to be, then any subject is “fair game.”

The weakness of the third method is that it may restrict the sample to those that have self-selected for experimental tasks. It is not obvious that such people are the same as those that would
self-select for those tasks in the field.

\textit{E. Repetition}

In the earliest experiments it was standard for the game to be repeated for many periods, typically 10 or 20. The idea was to allow subjects to learn about the “rules of the game,” and often the published report would only look at the last period or the last few periods. With the growth and popularity of game theory, and the understanding of the multiple equilibria possible in repeated games, experimenters started to be more concerned about “one shot” designs. The tension between playing games only once, and allowing subjects time to learn about the game, was resolved by developing the “strangers” design. This is where subjects are randomly grouped in each round, and then shuffled from round to round. If the sample size is large enough, and the number of players in any group small enough, this method reduces the chance of any player meeting another player again. Indeed, in some cases one can ensure that this does not happen, and tell this to subjects.

\textit{F. Deception and Clarity}

Deception is regarded as unethical by experimental economists, who are sharply distinguished from experimental psychologists on this. A distinction can be usefully drawn between outright acts of deception (the use of statements which are knowingly false) and the failure to provide certain information (which may be contrary to the beliefs of subjects). The latter is a grey area, but one that is accepted in the field.

Clarity might seem to be an obviously desirable characteristics for all scientists, not just experimental economists. However, one of the first “cultural” differences that strikes an experimental economist dipping his toes into the sea of CVM studies of the environment is how careful those studies are in their choice of language on some matters and how appallingly vague they are on other matters. The best CVM studies spend a lot time and money on “focus groups” in which they tinker with minute details of the scenario and the granular resolution of pictures used in
Actually, the subject just needs to attach some positive subjective probability to the chance of being the decisive voter. As that probability declines, so does the (hypothetical) incentive to tell the truth. So, to paraphrase Dirty Harry the interviewer, “do you feel like a specific order statistic today, punk?” Tough question, and presumably one that the subject has guessed at an answer to. I am just adding additional layers of guesswork to the main story.

The subjects were asked if they had any questions about how the program would be paid for (p. 55), and had any confusions clarified then. But this is no substitute for the control of being explicit and clear in the prepared part of the survey instrument.

Each household was given a “price” which suggested that others may pay a different “price.” This is standard in such referendum formats, and could be due to the vote being on some fixed formula that taxes the household according to assessed wealth. Although the survey does not clarify the possible basis for household-specific prices for the subjects, it would be an easy matter to do so.

---

For example, consider the scenario in the influential Exxon Valdez study by Carson et al. [1992]. Forget the simple majority-rule referendum interpretation used by the researchers, and focus on the words actually presented to the subjects. The relevant passages concerning the provision rule are quite vague on that rule, as illustrated below.

How might the subjects be interpreting specific passages? Consider one hypothetical subject. He is first told, “In order to prevent damages to the area’s natural environment from another spill, a special safety program has been proposed. We are conducting this survey to find out whether this special program is worth anything to your household.” (p.52). Are the proposers of this program going to provide it no matter what I say, and then come for a contribution afterwards? In this case I should free-ride, even if I value the good. Or are they actually going to use our responses to decide on the program? If so, am I that Mystical Measure-Zero Median voter whose response might “pivot” the whole project into implementation? In this case I should tell the truth. Later the subjects are told, “If the program was approved, here is how it would be paid for.” But who will decide if it is to be approved? Me, or is that out of my hands as a respondent? As noted above, the answer matters for my rational response.

Later in the survey the subjects are told, “Because everyone would bear part of the cost, we are using this survey to ask people how they would vote if they had the chance to vote on the program.” (p.55). OK, this suggests that the provision rule would be just like those local public school bond issues I always vote on, so the program will (hypothetically) go ahead if more than 50% of those that vote say “yes” at the price they are asking me to pay. But I am bothered by that
phrase “if they had the chance to vote”: does this mean that they are not actually going to ask me to vote and decide if the program goes ahead, but are just floating the idea to see if I would be willing to pay something for it after they go ahead with the program? Again, the basic issue of the provision rule is left unclear. The final statement of relevance does nothing to resolve this possible confusion: “If the program cost your household a total of $(amount) would you vote for the program or against it?” (p.56).

The upshot is that we cannot claim as outside observers of the survey response of an individual that we know what the subject is guessing at. We can, of course, guess at what the subject is guessing at. This is what Carson et al. [1992] do when they choose to interpret the responses in one way rather than another, but this is still just a dressed-up guess. Moreover, it is a serious one for the claim that subjects have an incentive to free ride, quite aside from the hypothetical bias problem.

The general point is that one can avoid these problems with more explicit language about the exact conditions under which the program would be implemented and payments elicited. We conjecture that CVM researchers will shy away from such language since it would likely expose to the subject the truth about the hypothetical nature of the survey instrument. But that is just the illusory attraction of staying in the frying pan. My point here is that these are issues that running experiments requires you to face, rather than dodge.

G. Field Experiments

Sometimes nature or policy provides us with a “natural experiment” in a field setting. For example, in 1992 the United States Department of Defense started offering substantial early retirement options to nearly 300,000 individuals in the military. This voluntary separation policy was instituted as part of a general policy of reducing the size of the military as part of the “Cold War dividend.” Warner and Pleeter [2001] recognize how the options offered to military personnel could be viewed as a natural experiment with which one could elicit information on individual discount rates. At the other end of the spectrum are the experiments of Coller and Williams [1999], which
imposed a similar set of choices on college students recruited for a general economics experiment. These are traditional lab experiments with student subjects. Somewhere in between lies the experiments of Harrison, Lau and Williams [2002], in which the imposed lab design of Coller and Williams [1999] was applied to subjects in the field. How is one to differentiate these experiments?

We propose identifying a set of factors that can be used to determine the field context of an experiment: the nature of the subject pool, the nature of the commodity, the nature of the task or trading rules applied, the nature of the stakes, and the nature of the information that the subjects bring to the task. The taxonomy that results will be important as comparisons between lab and field experimental results become more common.

Student subjects can be viewed as the standard subject pool used by experimenters, simply because they are a convenience sample for academics. Thus when one goes “outdoors” and uses field subjects, they should be viewed as non-standard in this sense. But we argue that the use of non-standard subjects should not automatically qualify the experiment as a field experiment. The experiments of Cummings, Harrison and Rutström [1995], for example, used individuals recruited from churches in order to obtain a wider range of demographic characteristics than one would obtain in the standard college setting. The importance of such non-standard subject pools varies from task to task: in this case it simply provided a less-concentrated set of socio-demographic characteristics with respect to age and education level, which turned out to be important when developing statistical models to adjust for hypothetical bias (Blackburn, Harrison and Rutström [1994]). Alternatively, the subject pool can be designed to represent the national population, so that one can make inferences that are representative of the general population (Harrison, Lau and Williams [2002]).

The commodity itself can be an important part of the field. Recent years has seen a growth in experiments concerned with eliciting valuations over actual goods, rather than using induced valuations over virtual goods. The distinction here is between physical goods or actual services and abstractly defined goods. The latter have been the staple of experimental economics since Smith
We would exclude experiments in which the commodity was gamble, since very few of those gambles take the form of naturally occurring lotteries. Such influences are actually of great interest, or should be. If the nature of the commodity itself affects behavior, in a way that is not accounted for by the theory being applied, then the theory at best has a limited domain of applicability that we should know about, and at worse is simply false. In either case, one can only know the limitations of the generality of theory if one tests for it, by considering physical goods and services.

Again, however, just having one field characteristic, in this case a physical good, does not constitute a field experiment in any fundamental sense. Rutström [1994] sold lots and lots of chocolate truffles in a laboratory study of different auction institutions designed to elicit values truthfully, but hers was very much a lab experiment despite the tastiness of the commodity. Similarly, Bateman et al. [1997] elicited valuations over pizza and dessert vouchers for a local restaurant. While these commodities were not physical pizza or dessert, but vouchers entitling the subject to obtain these, they are not abstract. There are many other examples in the experimental literature of designs involving physical commodities.17

The nature of the task that the subject is being asked to undertake is an important component of field experiment, since one would expect that field experience could play a major role in helping individuals develop heuristics for specific tasks. The experiments of Kagel and Levin [1999] illustrate this point, with “super-experienced” subjects behaving fundamentally differently than inexperienced subjects in terms of their propensity to fall prey to the “winners curse.”

The nature of the stakes can affect field responses. The stakes in the laboratory might be very different than those encountered in the field, and hence have an effect on behavior. If someone is experienced with taking valuations seriously when they are in the tens of dollars, or in the hundreds, but simply taking or leaving a price when it is less than $10, laboratory experiments with stakes below $10 could easily engender imprecise bids. Of course, people buy cheap things in the

---

17 We would exclude experiments in which the commodity was gamble, since very few of those gambles take the form of naturally occurring lotteries.
field as well, but the valuation process they use might be keyed to different stake levels.

In the field subjects bring certain information to their trading activities, quite apart from knowledge of the trading institution. In abstract settings the importance of this information is diminished, by design, and that can lead to changes in the way that individuals behave. Absent such information, and risk aversion can lead to subjects requiring a risk premium when bidding for objects with uncertain characteristics. List [2001] is a good example of a field experiment that utilizes a naturally occurring setting in which subjects are asked controlled tasks for commodities that they are familiar with.

**H. Computerization**

Many experiments are now computerized. There are many obvious advantages to computerizing experiments, not the least being the automatic collection of data in a machine-readable format! Moreover, with the development of popular “toolbox” programs such as *Ztree*, it is possible to develop working experiments relatively quickly. The great advantage of toolbox programs is that they offer template programs that are already working, so that one can simply modify that experiment for one’s own purposes. The other advantage of computerized experiments is the additional control they provide over the process of conducting the experiment. This is particularly evident when detailed record-keeping is required, or detailed information displays desired.

The disadvantages of computerization are that one can easily spend a lot of time on the implementation stage, that one typically needs access to computer labs to run the experiments,\(^{18}\) that one can easily become dependant on lazy programmers, and that the computer interface can be intimidating for some subjects.

On balance, and speaking as a programmer, it is best to develop pilot experiments that can

\(^{18}\) This is rapidly changing, however, with many experiments being implemented directly on the web. For example, see the WebLab of the University of South Carolina at http://weblab.moore.sc.edu/ and the web-based experiments developed by Holt [2002].
be run “by hand” if possible. Once a design has been developed to the point that one wants to replicate under more controlled conditions, it is appropriate to consider incurring the fixed costs of a computerized version.

I. Journals and Conferences

The field of experimental economics now has a specialist journal, *Experimental Economics*. However, most substantive contributions that use experiments still get published in the best general journals, such as the *American Economic Review*, *Journal of Political Economy*, and *Econometrica*. Many experiments in the area of expected utility theory have appeared in the *Journal of Risk and Uncertainty* over the years, and many experiments in game theory have appeared in *Games and Economic Behavior*. The major journal in environmental economics, the *Journal of Environmental Economics and Management*, has long been open to experimental studies. Just as there are few journals that are so specialized that they would reject a paper that uses econometric methods, there are few journals that reject the use of experimental methods.

It was not always this way. Many older experimentalists can recall battles over many years to get major papers published. Indeed, I can recall one memorable referee report on Friedman, Harrison and Salmon [1984] that was written by an experimenter who was clearly convincing the editors of the *Journal of Political Economy* to accept the paper! Those were, indeed, memorable times.

The Economic Science Association holds an annual conference, and alternates between continental Europe and the United States. Although well attended by experimenters, these conferences tend to be “too busy” in the sense of having too many parallel sessions and too little plenary time. They can be invaluable for junior researchers wanting to get to meet people, however. More productive in some cases are conferences such as the Southern Economic Association or Public Choice Society, which tend to just have one experimental session at a time, but to have multiple sessions over several days. In this manner there is much more of a “shared dialogue” over several sessions.
Experimental economics has reached the stage where specialist conferences are being held again. For example, Jeffrey Carpenter and I are organizing a conference on Field Experiments at Middlebury College in April 2003, and another conference is being organized by Morten Lau and I on Discount Rate Experiments in Copenhagen later in 2003. The Economic Science web site is perhaps the best place to keep track of forthcoming conferences.

J. Pedagogy

A major development has been the use of experiments for pedagogic purposes. This refers to classroom pedagogy as well as policy pedagogy.

The best single source for work on classroom experiments is the collection of papers that Charles Holt has co-authored, mainly in the Journal of Economic Perspectives. These have been collected in a valuable reference available on his web page: Holt [2002]. This provides tools for web experiments for classes conducted in web-enabled computer labs, as well as hand-run experiments for more traditional classroom settings. These can be an excellent way for junior researchers to “try out” experiments as part of their teaching, and to see if they find it attractive to pursue as a research tool. Students tend to love these experiments, particularly if the instructor makes it clear that these experiments are for legitimate pedagogic purposes and are not a cheap way to generate slave labor for their research (as in psychology).

Policy pedagogy occurs where experimenters use different institutions to show regulators or policy-makers the effects of different rules. Early examples of some importance for environmental economics include the various ways of implementing tradeable emissions permit markets, collected in Holt and Isaac [1999]. More recently, experimenters were heavily involved in the test-bed design of auctions for privatizing “spectrum rights.”

K. Smart Markets

One of the most exciting areas of experimental research is the design and evaluation of
computerized resource allocation mechanisms, or “smart markets”. In these markets the computer plays an active role in information processing in trading institutions, rather than simply being a passive medium for the transmission of information. The idea originated in Smith [1977][1979], which attempted to implement a new auction procedure for public goods evaluation. This institution required that the computer undertake simple calculations in real time so that the subjects could see the results of their trial messages and adapt their behavior.

Subsequent work in this area has examined much more sophisticated resource allocation mechanisms that require substantial algorithms to be solved in real time. The path-breaking paper in this area is by Rassenti, Smith and Bulfin [1983]. They examined a fascinating combinatorial problem concerning airport landing slot allocations. The problem was derived from a field policy issue, the concern of the federal aviation authorities in the United States over whether airport landing slots could be privatized. As a commodity, airport landing slots have an interesting feature: an agent can be expected to value a takeoff slot only if it has access to a landing slot somewhere else. This suggests that the commodity to be valued, pairs of takeoff and landing slots, takes on a combinatorial dimension. With a moderately large number of slots available in different cities, where slots are also differentiated by time of arrival or departure, there are a large number of commodities that could be evaluated and would need to be considered simultaneously for a systematic allocation of resources.

Rassenti, Smith and Bulfin [1983] proposed that agents be allowed to submit conditional and logical demand functions for such combinatorial commodities, and that these conditional demand functions be entered into a formal mathematical programming problem that would seek to optimize system surplus in some well-defined sense. The shadow prices to be returned to subjects would then guide their revisions of these bids, and the process would iterate in this fashion.

An essential finding from these experiments is that subjects are not able to strategically

---

19 The idea of “responsive pricing” was also a major contribution of William Vickrey. For example, see Arnott et al. [1994, ch.12].

20 The basic idea is one that has also been applied in subsequent work to electricity transmission networks and to gas pipeline networks: see McCabe, Rassenti and Smith [1989]. And many of the principles were used in the design and evaluation of the auction of “spectrum rights” in the United States and other countries.
manipulate the mechanism, although this may be a result that does not stand the test of time or the test of field incentives to do so. It would thus appear, initially at least, that a new economic institution has been developed in the lab. This is in sharp contrast to the typical approach of simply examining existing institutions that have evolved over time. In fact, the control and cost-effectiveness of the laboratory make it an excellent test bed for new institutions. This area of research promises to grow dramatically in the next few years, as game theorists discover possibilities for trying out other new institutions in this manner.

4. Classics

What papers should every experimental economist read? Every list of this kind is subjective, but the field already has several classics that few would question. The ones listed here are papers that I have all of my graduate students review, apart from Smith [1962] and Grether and Plott [1979].

Cox, Roberson and Smith [1982] did much to define the field of experimental research into bidding behavior. Not because it was the first experiment in the area, but because of the way it combined rigorous new theory and careful experimental controls.

Fiorina and Plott [1978] opened up the use of induced value to the study of voting behavior, and is a careful example of how one “builds bridges” between a field and experimental methods. In this case a whole new area of experimental research was developed in political science (e.g., the collection in Palfrey [1991]).

Forsythe, Palfrey and Plott [1982] and Plott and Sunder [1982][1988] broke new ground by examining a rich series of asset markets, and giving structure to the fundamental notion of “informational efficiency.” In doing so, it defined another new area of experimental research in finance and accounting.

Hey and Orme [1994] provide an integrated statement of the testable implications of

---

21 Friedman, Harrison and Salmon [1984] and Forsythe, Nelson, Neumann and Wright [1992] illustrate two directions that this research initiated.

-38-
expected utility theory over a simple array of lottery choice problems. My students only really understand the alternatives to expected utility theory when they work through the implications stated here, which also helps provide a bridge to earlier studies which focused on one specific alternative or another.

Rassenti, Smith and Bulfin [1983] defined the area of “smart markets,” as discussed earlier. Although Smith’s [1977][1979] work on public goods mechanism were precursors, this study really identified the manner in which computers could play a non-trivial role in operationalizing the invisible hand for network markets.

Roth and Malouf [1979] provided a rigorous statement of how one can make bargaining theory operational, introducing the notion of a “binary lottery game” to ensure that subjects could bargain over expected utility without it being interpersonally comparable. The paper went beyond that major contribution to show how one could modify the axiomatic Nash Solution to accommodate the fact that subjects often do make such comparisons, and that it will influence behavior in predictable ways.

5. Conclusions

Experimental economics has grown rapidly in recent decades. It is now viewed by many as an empirical tool that most economists should be familiar with, in the same way that any economist should know the basics of econometrics. Fortunately, the main ideas are relatively simple and accessible, no doubt because of the intrinsic need to explain the experimental task to the “ordinary folk” that make up the subject pool. The ability of economists to gain more control over the institutional and informational setting that individuals act in should enhance the importance of rigorous theory, at the very same time that experimental results are challenging important parts of existing theory.
References


