
Introduction

John A. List and Michael K. Price

Understanding the types of behaviors that generate public goods (bads) and the ability to measure the total value of non-market goods are cornerstones of environmental and resource economics. Identifying and designing socially optimal policy measures requires both an understanding of the theoretical models that best predict behavior and the ability to obtain credible estimates of the underlying impact on benefits and costs.

However, a fundamental challenge facing researchers who wish to estimate the causal effect of a given action or policy is the construction of the correct counterfactual. As the action of interest is either taken or not, the researcher is unable to observe what would have happened if different actions were taken. Yet, it is possible to observe outcomes for similar others – or a control group – who were not treated or who took a different action. In doing so, economists have turned to the experimental model of the physical sciences as a method to create valid control groups and isolate important aspects of human behavior.

Intuitively, controlled experiments rely upon exogenous variation in the variable of interest to facilitate causal identification. Randomization provides an instrument that allows the researcher to isolate causal relationships and uncover the mechanisms underlying observed behavior. The extent to which the environmental economists have embraced controlled experimentation as a means to test theory and develop policy relevant estimates of value should thus come as no surprise.

Much of this research has taken the form of laboratory experiments in which volunteers enter a research lab to make decisions in a controlled environment. Over the past decade, economists have increasingly made use of field experiments to explore economic phenomena. Field experiments provide a bridge between laboratory and naturally-occurring data in that they represent a mixture of control and realism usually not achieved in the lab or with uncontrolled data, permitting the analyst to address questions that heretofore were quite difficult to answer.

This book includes a collection of research that makes use of the experimental method to explore key issues within environmental and resource economics. It includes contributions that span both laboratory (lab) and field experiments as well as research utilizing econometric methods to enhance our understanding of controlled data. Table I.1 shows a

Table I.1 A field experiment bridge

	Controlled Data		Naturally-Occurring Data
Lab	AFE	FFE	NFE, NE, PSM, IV, STR

Key

Lab:	Laboratory experiment
AFE:	Artefactual field experiment
FFE:	Framed field experiment
NFE:	Natural field experiment
NE:	Natural experiment
PSM:	Propensity score estimation
IV:	Instrumental variables estimation
STR:	Structural modeling

taxonomy of the variety of research that falls under the rubric of empirical research. On the far left in Table I.1 are lab experiments, which typically make use of randomization to identify a treatment effect of interest. The right-most part of the empirical spectrum in Table I.1 includes examples of empirical models that require making identification assumptions to identify treatment effects from naturally-occurring data.

Between these endpoints are field experiments. The most minor departure from the typical laboratory experiment is the ‘artefactual’ (i.e. artificial, fake, or synthetic) field experiment, which mimics a lab experiment except that it uses non-standard subjects, typically experimental participants from the market of interest. Moving closer to how naturally-occurring data are generated, a ‘framed field experiment’ is the same as an artefactual field experiment, except that it incorporates important elements of the context of the naturally occurring environment with respect to the commodity, task, stakes, and information set of the subjects. Yet, it is important to note that framed field experiments, like lab experiments and artefactual field experiments, are conducted in a manner that ensures subjects understand that they are taking part in an experiment, with their behavior subsequently recorded and scrutinized.

Natural field experiments are those experiments completed in cases where the environment is such that the subjects naturally undertake these tasks and where the subjects do not know that they are participants in an experiment. Therefore, they neither know that they are being randomized into treatment nor that their behavior is subsequently scrutinized. Such an exercise is important in that it represents an approach that combines the most attractive elements of the lab and naturally-occurring data: randomization and realism. In addition, it is difficult for people to respond to treatments they do not necessarily know are unusual, and of course they cannot

excuse themselves from being treated. Hence, many of the limitations cited in the above approaches are not an issue when making inference from data generated by natural field experiments.

With this in mind, we turn to a brief discussion of the contributions within this volume. First, we worked to include several excellent studies exploring econometric methods associated with analyzing experimental data. We view this as frontier work, and appreciate the fact that the studies we received were first rate.

Parmeter and Pope's chapter explores the use of the quasi-experimental approach to estimate hedonic property models. The chapter begins by describing the economic and econometric theory underlying the quasi-experimental hedonic methods. It provides an overview of hedonic theory and its relation to quasi-experiments and outlines the theoretical and empirical assumptions required when moving from a cross-sectional hedonic model to one relying on quasi-experimental treatment over time or space to identify causal impacts.

The chapter next outlines a series of steps that practitioners should take to ensure the reliability of identification strategies combining quasi-experiments and property data. To illustrate the importance of these steps and provide practitioners a blueprint for estimating quasi-experimental hedonics, the chapter concludes by reviewing three recent papers utilizing the methods to explore important issues within the realm of environmental and resource economics. In this regard, the chapter provides a how-to guide for researchers interested in using the approach to quantify the impact of exogenous changes in environmental and urban amenities on housing prices.

The chapter by Moeltner, Murphy, Stranlund and Velez develops a Bayesian framework that allows the researcher to estimate heterogeneous treatment effects in social dilemma games. Using data from a common pool resource game run in two artisanal fishing regions in Columbia, the authors find compelling evidence that subjects respond heterogeneously to treatments introducing harvest quotas and communication possibilities. Importantly, the authors find that the extent of such heterogeneity varies across villages and treatments. As such, it is shown that 'naïve' models that allow for unobservable heterogeneity of a general nature tend to underestimate variability in treatment effects and produce misleading estimates that can lead to sub-optimal institutional choices and related policy recommendations. The empirical approach advanced in this chapter should prove valuable for researchers interested in using experimental methods to test-bed new institutions and environmental policies.

The chapter by Vossler builds upon the recognition that many experimental games involve participants (or groups of participants) making a

sequence of related decisions over several rounds of play. To the extent that subjects learn or adapt play over time, it is plausible to expect serial correlation in outcomes and associated decision errors. Yet, as Vossler aptly notes, most of the empirical approaches utilized when analyzing experimental data fail to account for serial correlation; a fact that is unfortunate as the standard errors for such estimators and any hypothesis test based on them may prove biased.

To address this concern, Vossler proposes the use of heteroskedasticity and autocorrelation consistent (HAC) covariance estimators to analyze experimental data that involve repeated play. The use of HAC estimators allows the researcher to perform hypothesis tests without having to place structure on the correlation in errors across time and/or subjects. Vossler next explores the properties of three different HAC estimators within the context of Monte Carlo experiments that vary both the number of time periods and cross-sectional units. Results from these simulations provide practitioners a guide outlining two conditions under which the use of HAC estimators is preferable to other estimation frameworks: (i) when the number of cross-sectional units is large and (ii) when serial correlation is of an unknown form. In contrast, for experiments that explore data at the group level and therefore contain a small number of cross-sectional units per treatment, the benefits of the HAC estimators are less certain.

Second, we include explorations of issues associated with non-market valuation, which remains the most important policy issue within experimental economics. In the United States, President Clinton's Executive Order 12866, which reaffirmed an earlier executive order from the Reagan Administration, explicitly mandates that all federal agencies consider economic impacts of regulations prior to their implementation. For commodities traded in the marketplace, estimating such impacts is straightforward – market prices provide a direct signal of value. Unfortunately, the task is more daunting when estimating the total benefits of non-market goods as services such as environmental quality. In such instances, policy makers frequently rely on stated preference methods to derive signals of value. Yet, such methods are not beyond reproach and studies exploring the properties of different elicitation formats have become a cornerstone of environmental economics. While this literature is vast and growing, the three studies that touch on non-market valuation should leave a mark on this field.

The chapter by Horowitz, McConnell and Murphy provides a nice summary of the experimental literature exploring two fundamental issues that arise when economists apply stated preference approaches to elicit non-market values: (i) differences in values derived from real and hypothetical surveys and (ii) the gap between measures of willingness to pay

(WTP) and willingness to accept (WTA). A central theme of the review is the interconnection between value elicitation via stated preference techniques and experiments. As the authors aptly note, ‘only for valuation can experiments be said to be indispensable, essential to the conceptualization and understanding of the underlying economic ideas . . . All of what we know about valuation we learn from experiments.’

The chapter should prove a valuable reference for policy makers and academicians interested in non-market valuation. It summarizes what has been learned from the extant experimental literature, presents the conceptual models (both neoclassical and behavioral) that have been posited as a means to explain the WTP/WTA gap, discusses the implications of the experimental evidence for value elicitation, and outlines directions for future research.

Knetsch explores the experimental evidence surrounding the value disparity and what it implies for the choice of welfare measure when evaluating environmental policies. In doing so, the chapter highlights how assessment procedures appear to influence reference states and subsequent values. As such, Knetsch argues that the choice of measure – compensating or equivalent variation – should be based on whether an individual’s reference state reflects the *ex ante* (status quo) or *ex post* state of the world. The chapter concludes by discussing how the choice of measure impacts the way in which we value environmental damages and evaluate policies designed to reduce such damages.

Neilson, McKee and Berrens explore the role of outcome and value uncertainty on WTP and WTA measures. The chapter begins by examining behavior under two classes of utility functions (expected utility with loss aversion and rank-dependent expected utility). Within this framework, the authors show that the two models make very similar predictions concerning the divergence between WTA and WTP and the factors that drive this gap. The chapter concludes by presenting results from a laboratory experiment designed to elicit WTP and WTA measures for an asset with uncertain redemption value and test the underlying theoretical predictions. Experimental results largely conform with theoretical predictions and highlight that the WTA/WTP gap is increasing in value and outcome uncertainty but can be reduced by introducing a re-contracting option that allows subjects to reverse transactions and therefore reduce value uncertainty.

Third, we have interesting explorations of emissions trading and pollution abatement studies. Perhaps there are no issues – normative or positive – as timely as the problems that the authors of these studies approach, but their insights should be cited for years to come. The chapter by Giordana and Willinger provides an overview of the experimental literature on

emissions trading – particularly those studies designed to compare the relative performance of alternate mechanisms. The authors highlight common findings in the literature and outline areas where more research is needed. In this regard, the chapter should prove a must read for policy makers and academicians alike.

The chapter by Gangadharan, Croson, and Farrell presents results from a laboratory experiment designed to explore how the banking of allowances, investments in abatement technology, and uncertainty over future permit allocations impacts the performance of cap-and-trade markets. The chapter extends the extant literature on cap-and-trade programs by exploring the interplay of banking, investment decisions, and uncertain permit allocations. Importantly, the authors find that ‘overbanking’ is significantly reduced when allowing for both investment in abatement technology and banking. This suggests that overbanking may be less of a concern in naturally occurring markets than prior laboratory studies would suggest.

Stranlund, Murphy and Spraggon explore the role of imperfect enforcement on the performance of emissions trading markets and industry welfare. The authors present results from a series of laboratory experiments designed to investigate whether impact market outcomes deviate from theoretical predictions when enforcement is not expected to induce full compliance. Experimental results suggest that concerns regarding imperfect enforcement may be overblown – overall efficiency levels and compliance are higher than theory would predict.

Stoop, van Soest and Vyrastekova present results from a novel laboratory experiment designed to explore the role of peer-to-peer rewards and ‘counter rewards’ on cooperation in social dilemmas. The chapter builds upon a growing literature exploring the impact of various ‘decentralized’ mechanisms on play in social dilemma games and should prove valuable for those who work in this area. Whereas prior work has shown that peer-to-peer rewards are an effective way to promote cooperation, the authors present results calling this work into question. Allowing agents to reciprocate rewards weakens the link between behavior in the social dilemma situation and first-stage rewards – agents who believe their counterparts are reciprocal may dole out rewards independent of choice in the underlying social dilemma. This leads to an ultimate unraveling of cooperation and dramatic reductions in overall market efficiency.

The chapter by Cardenas, Janssen and Bousquet presents results from a series of framed field experiments that introduce ecological complexity (resource dynamics, spatial effects, and non-linearity) into standard public goods and common pool resource games. The authors develop these games as a way to enhance our understanding of human–ecosystem inter-

actions and the results of the study will undoubtedly prove valuable for academicians and policy makers alike. From a sample of 72 sessions conducted in six different villages across Columbia and Thailand, the authors find no evidence that groups are strict money maximizers and behave in a manner that ultimately leads to a 'tragedy of the commons'. Yet, overcoming the temporal and spatial ecological features of the resource dynamics proved to be a difficult task – subtle differences in initial period play led to dramatic differences in subsequent harvests and efficiency levels.

Lopez, Murphy, Spraggon and Stanlund present results from a novel framed field experiment exploring whether the introduction of external regulations impacts self-governance and peer-to-peer monitoring. The experiments were conducted on the southern Pacific coast of Columbia in communities that rely upon the extraction of natural resources such as wood and fish for their existence. The experiments were based on a standard linear public goods game, but were described to the participants in terms of contributions of effort to clean the local mangroves. Across all treatments, subjects had the possibility to take a costly action to sanction others in their group – either by sending a signal of displeasure or imposing a sanction that lowered the recipient's payoff. Formal regulations were introduced as a requirement a minimum contribution level that was enforced through random monitoring. Experimental results suggest that government regulations prove complementary to community enforcement efforts. Although subjects tend to reduce sanctioning efforts once formal regulations are introduced, contributions and earnings were higher in such treatments.

Giordana and Willinger evaluate the effectiveness of three temporally inconsistent policies for regulating the exploitation of a renewable common-pool resource. The authors consider an N -person, discrete-time deterministic game of fixed duration and three policy instruments whose parameters are fixed over all T periods of play: (i) a pigouvian tax, (ii) an ambient tax, and (iii) a mixed flat instrument. Experimental results suggest that the flat tax and mixed instrument are the most effective ways to promote behavior that approximates the optimal extraction path. However, such policies are inefficient. For practitioners, these results suggest that the implementation of time-inconsistent policy instruments must be made very cautiously – stock savings imply forgone earnings that must be cashed in future periods.

Finally, we include public goods work alongside voting experiments to provide a look at the whole cloth of issues that the environmental and resource economist are interested in. These studies not only provide breadth, but are deep in their contributions. The chapter by Zetland reports results from an experiment designed to test whether

water managers drawn from the Metropolitan Water District of Southern California are more cooperative than undergraduate students when facing a social dilemma. Results suggest that there are no differences in average cooperation across these groups, although water managers are more likely to be classified as cooperators and free-riders.

Silz Carson presents results from a series of experiments designed to explore the relative performance of various incentive-compatible mechanisms used to elicit preferences for a public good. Results of these experiments suggest that while it is possible to implement an incentive-compatible mechanism to estimate the demand for an environmental public good, such a mechanism may not necessarily perform better than existing methods to estimate the demand for these goods – particularly when implemented using conditions that mirror those that arise in field settings.

Kroll, List and Mason present results from a laboratory experiment with three different treatments – a standard two-player prisoner's dilemma, a representative democracy, and a direct democracy – designed to compare individual and collective behavior when subjects face variations in the voting institution. The authors find that the propensity to cooperate increases when subjects move from individual games to group games. They also observe an interesting result in the representative democracy regime that is usually not observed in other prisoner's dilemma experiments, namely that cooperation is robust in later periods in the experiment. In this regard, the study can be considered a warning that research using unitary rational actor models in cases where not unitary individuals but groups interact might be inappropriate, and should further the research agenda on the differences between group and individual interactions in environmental economics.

In closing, we must express our gratitude to the authors in complying with our many requests, restrictions, and deadlines. Several of our colleagues also lent fine support in the refereeing process, which is a thankless job, but one which is gratefully appreciated by us.