The impact of experiments on environmental policy and natural resource management

Christian A. Vossler, Timothy N. Cason, James J. Murphy, Paul J. Ferraro, Todd L. Cherry, George Loewenstein, Peter Martinsson, Jason F. Shogren, Leaf van Boven, and Daan van Soest

December 2024

WORKING PAPER #2024-05

WORKING PAPER SERIES
DEPARTMENT OF ECONOMICS
HASLAM COLLEGE OF BUSINESS
https://haslam.utk.edu/economics/

HASLAM COLLEGE OF BUSINESS
THE UNIVERSITY OF TENNESSEE, KNOXVILLE

The impact of experiments on environmental policy and natural resource management

Christian A. Vossler[a,*], Timothy N. Cason[b], James J. Murphy[c], Paul J. Ferraro[d], Todd L. Cherry[e], George Loewenstein[f], Peter Martinsson[g], Jason F. Shogren[e], Leaf van Boven[h], and Daan van Soest[i]

[a] Department of Economics, University of Tennessee, Knoxville, TN 37996
[b] Department of Economics, Purdue University, West Lafayette, IN 47907
[c] Department of Economics, University of Alaska Anchorage, Anchorage, AK 99508
[d] Carey Business School and Department of Environmental Health and Engineering, Johns Hopkins University, Baltimore, MD 21202
[e] Department of Economics, University of Wyoming, Laramie, WY 82071
[f] Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA 15213
[g] Department of Technology, Management and Economics, Technical University of Denmark, 2800 Kongens Lyngby, Denmark
[h] Department of Psychology and Neuroscience, University of Colorado, Boulder, CO
[i] Department of Economics, Tilburg University, 5000 LE Tilburg, Netherlands

December 17, 2024

CAV, TNC, JJM: Conceptualization, Investigation, Writing – Original Draft; **PJF, TLC, GL, PM, JFS, LvB, DvS**: Conceptualization, Investigation

* Direct correspondence to Christian Vossler, Department of Economics, University of Tennessee, Knoxville, TN 37996. E-mail: cvossler@utk.edu. Telephone: 865-974-1699. Fax: 865-974-4601.

The impact of experiments on environmental policy and natural resource management

Abstract

Motivated by the fact that few academic publications document the links between behavioral experiments and public decision making, this paper compiles and describes many studies that were used to inform environmental policy and natural resource management decisions. These experiments informed or changed the designs of emissions trading programs, recreational fishing regulations, conservation auctions, pro-environmental initiatives directed at households, and regulatory enforcement and compliance schemes, and produced nonmarket demand estimates that informed government regulatory analyses. We highlight the context and conditions that produced these experiment-policy links and discuss how researchers can better engage with the policymaking process and increase the impact of experimental research on policy.

JEL Codes: C9; D04; D47; Q28; Q48; Q51; Q58

Introduction

Al Roth famously suggested that an important use of experimentation in economics is to "whisper in the ears of princes" (Roth 1986). He was referring to behavioral experiments designed to test features of real-world public policies or to explore the characteristics of proposed policy options. By "wind-tunneling" or "test-bedding" public policies in a rigorously controlled environment, experimental methods can help evaluate policy proposals and test new institutional designs (Smith 1994). Experiments can also provide insights into behavioral responses to policies (Shogren and Taylor 2008; Carlsson and Johansson-Stenman 2012). Outside the environmental context, high-profile examples of experiments used by US government agencies to inform policy decisions include the Treasury auctions in the 1970s (Smith 1994), payload allocation on NASA's Cassini mission to Saturn (Ledyard et al. 2000), the Securities and Exchange Commission's analysis of short sale price restrictions (US Securities and Exchange Commission 2007), and the design of the Federal Communication Commission's spectrum auctions (Plott 1997; Guala 2001).[1] Evidence-based experimental approaches are required for the adoption of new medical procedures and medications, and are frequently incorporated into the evaluation of social and education programs such as the Oregon Medicaid program (Finkelstein et al. 2012; Baicker et al. 2013) and the Tennessee Student/Teacher Achievement Ratio experiment (Krueger 1999).

In this article, we document evidence on links between experimental research directly used by government agencies and other public organizations to inform the design of a new policy or to amend an existing policy. Few research-policy links are evident from published academic papers or can easily be unveiled through targeted literature searches. As such, much of the

[1] For a description of an economics experiment, a discussion of the different types of experiments and how they can be effective at identifying the causal effects of public policies, see Higgins et al. (2017).

evidence we present is a result of personal communications with researchers at academic institutions and administrators at government agencies, and crowdsourcing efforts through organizational listservs.

Our investigation serves multiple purposes. The first is to aid in instructional efforts. Policymakers, funders, public organizations, students, and researchers are more likely to be convinced of the value of experimentation through many salient and transparent examples. Second, they will also gain a better understanding of the experimental "mindset"—how to combine theory, testing, and tinkering to understand the incentives that underpin and drive environmental policy (e.g., see Duflo 2018; Shogren et al. 2021). Third, policymakers who are either skeptical about experiments with human subjects or need to justify their usefulness will be able to draw upon precedent examples. Finally, although there are many examples in which a link exists between experimental research and policy implementation, only a handful of published papers clearly document these links. This review aims to fill that gap.

The paper proceeds as follows. In the next section, we introduce relevant terms and discuss the general ways in which experiments can inform policies. The subsequent five sections provide tangible examples in specific environmental and natural resource policy settings. We conclude by making recommendations for researchers to increase the policy impact of their experimental work.

Experimental Methods and Sphere of Influence

Formal experimentation is a central component of science, and before proceeding it is important to be clear what we mean by an "experiment" as well as by a "policy link." We define an experiment broadly to mean any scientifically controlled (i.e., a non-naturally occurring) procedure intended to evaluate a hypothesis, measure preferences, or examine the effect of a mechanism, institution, or intervention. This definition of an experiment excludes situations typically referred to as “natural experiments” or “quasi-experiments.” Unlike in these cases, most studies that we document use random treatment assignment (and control groups), which provides a gold standard for causal identification. [2]

Our definition of an experiment encompasses a wide variety of methodologies, including highly stylized experiments conducted in university laboratories using student subjects, lab-in-the-field experiments employing industry professionals or consumers as participants, and field experiments implemented in natural settings.[3] We also include stated preference (SP) survey experiments, which engage representative samples and use randomization methods to measure preferences and predict behavior. The various methodologies have distinct advantages and disadvantages (Charness 2015). For example, online experiments have been growing in recent years due to technological advances (Fréchette, Sarnoff, and Yariv 2022). Although online experiments often employ more representative subjects than laboratory experiments, they also generate considerably noisier data, probably due to weak incentives and inattention (Snowberg and Yariv 2021). The most appropriate experimental methodology and subject pool depends on the questions being asked and types of policy guidance being sought (Cason and Wu 2019; Henrich et al. 2010).

By a “policy link” we are referring to the existence of tangible evidence that experimental research has been used as a direct input into a decision impacting the public.[4] Some policy links

---

[2] For a critical discussion on RCTs and randomization see Deaton (2020).

[3] Harrison and List (2004) provide an influential classification of various types of experiments. Both Fréchette (2015) and Gneezy and Imas (2017) discuss the use of lab experiments with non-student subject pools.

[4] See Normann and Ricciuti (2009) for a review of experiments that were designed to address a specific economic policy or mechanism. Although these examples do not include links to implementation, the paper provides useful insights into the types of policies that are amenable to experimentation and the value of using experiments to inform policymaking.

that we identified are documented in the publications that report the experimental results, whereas other policy links were identified through policy briefs or personal communications with the experimental study's authors. While some experimental results were employed in making specific policy design choices, in far more instances, experiments provided only broad guidance for improving policy design or supporting a policy action.[5]

Examples we discuss illustrate a range of ways that experiments can affect decision-making. One of the most useful roles that experiments can play is in the policy design stage. Policy design can be guided by theory, but empirical evidence provides greater confidence that the policy will accomplish the policymakers' objectives.[6] When considering a novel policy innovation, no "naturally occurring" data exist that indicate whether the policy would work as intended. Experiments can provide the missing empirical guidance. Such "testbed" experiments are particularly influential when the policymakers engage with experimenters at an early stage because the testbed experiments can help refine specific details to be implemented in the policy.

Experiments can also be useful after implementation by determining whether a specific policy feature has led to intended or unintended results, and by suggesting improvements.[7] Making changes after the policy has been implemented can, however, be challenging. It may be difficult legally to make changes, especially when aspects of the initial policy design were guided by enabling legislation, and policy changes often require an extensive notification and

---

[5] For surveys of the literature that use experiments to explore policy-relevant environmental and natural resource issues but did not have a direct impact on the adoption a particular new policy or change of an existing one, see Sturm and Weimann (2006), Ehmke and Shogren (2008), Cárdenas (2009), Friesen and Gangadharan (2013), List and Price (2016), and Brent et al. (2017).

[6] Del Carpio et al. (2021), in the context of a monitoring program in rural Costa Rica aimed at reducing community water use, provide an example of how researchers and policymakers can collaborate during the experimental design phase to ensure a common understanding of what being studied, how outcomes will be measured, and how hypotheses will be tested.

[7] For example, policymakers may rely on engineers' estimates of resource savings. However, if people cannot use the technology correctly, then the actual gain may be much smaller (Alpizar, Del Carpio and Ferraro 2024).

comment period. Perhaps most importantly, implementing policy changes may require an acknowledgement that mistakes were made earlier. Nevertheless, when some aspect of a policy appears to not be working, experiments can provide guidance on what to do next.

While experiments are useful inputs to decision making, we acknowledge that evidence from a single experiment is often inadequate to justify a specific policy.[8,9] An experiment can provide direct evidence of causal links and provide impact estimates, but replication and complementary analyses, such as an assessment of benefits and costs, may be needed prior to implementation. And, even when a causal link between a policy and a desired outcome is clear, this alone does not justify the policy intervention. Many other considerations for policy adoption are intentionally omitted from the controlled setting of an experiment, such as political, ethical, and fiscal concerns. Moreover, some interventions that are successful in small trials may not work when implemented at a larger scale (see Al-Ubaydli, List, and Suskind 2019, Al-Ubaydli et al. 2021, and List 2024 for a discussion of the scaling issue and ways to mitigate this concern). Thus, certain policy interventions may not be appropriate even if the experimental evidence indicates a desired effect. Finally, we acknowledge the broader range of experimental and non-experimental research that make important contributions to knowledge and indirectly influences policy. Examples discussed in this article illustrate the ability of experiments to generate primary data to directly inform policy decisions.

[8] We thank Dale Whittington for emphasizing this point to us.

[9] Well-documented experiments may also form the basis for meta-analyses. In general, and especially for meta-analyses, it is important to report all experimental treatments, whether yielding significant or insignificant results, to avoid selective publication of only significant findings.

Property Rights Trading

One of the most significant contributions of the environmental and resource economics literature has been to establish the value of market-based mechanisms in achieving policy objectives at a lower cost than traditional command-and-control policies. Examples include emissions trading programs for pollutants such as SO_2, NO_x, and CO_2, individual trading quotas in fisheries, and ground water recharge credits. Many of the experiments in this domain focus on mechanism design, such as testing different types of auctions, market structure and price formation, and policy implementation.

In the regulatory toolkit, emissions trading has grown in importance because it can limit a wide variety of pollutants at low cost by harnessing the power of the market. Yet, implementing emissions trading programs requires new market institutions and rules for which there is typically no available observational data to provide guidance on mechanism design. These institutions and rules can, however, be varied in controlled experiments to evaluate their empirical consequences. In other words, laboratory experiments provide a natural testbed to help improve the performance of these nascent markets.

An early example of a testbed experiment is the RECLAIM trading program in Southern California. The South Coast Air Quality Management District (AQMD) used emissions trading to lower NO_x and SO_2 emissions from stationary sources by annually phasing down allocated permits. In trading programs, permits are associated with a specific compliance period that has a start and end date. When too many or too few permits are available late in a compliance period, the permit prices in the market can be volatile. The most common way to limit permit price volatility is by allowing traders to bank permits for later use. This tool, however, was not available to AQMD regulators because the RECLAIM program did not allow unused permits to be banked for use beyond their compliance year. To provide an alternative way of smoothing out price variability without banking, Carlson et al. (1993) developed a set of experiments to test a new idea. In a pair of experiments, the researchers showed that price volatility arising from a uniform permit expiration date could be nearly eliminated by staggering expiration dates. Doing so helped avoid the "squeeze" of excess supply or demand at the uniform expiration date. The AQMD adopted the staggered expiration date design, which continues to this day.

Another early experiment was inspired by regulatory design choices made in the implementation of Title IV of the 1990 Clean Air Act Amendments (CAAA), which was intended to reduce SO_2 emissions. The US EPA implemented a new auction institution with some unusual incentive properties that arose from the agency's interpretation of the CAAA. The auction required individual sellers to receive the bid prices of "matched" buyers, with higher priority and thus higher bid prices given to sellers who offered lower prices.[10] This bid-to-offer matching scheme created incentives for all buyers and sellers to submit offers below their marginal cost of emissions abatement, which can generate downwardly biased price signals and even inefficient trades (Cason 1993).

Cason (1995) found that human subjects in a lab experiment could recognize and bid according to the theoretical incentives of the EPA auction, and Cason and Plott (1996) compared behavior and performance of the EPA auction rules to the more typical uniform-price auction rules (used daily by specialists on the NYSE to set opening prices). Sellers facing EPA auction rules quickly recognized that they could obtain higher prices by offering permits well below cost, and the overall level of value and cost misrepresentation in the trader bids and offers far

[10] Clean Air Act Amendments of 1990 (Public Law 101-549), Sec. 416(d)(2).

exceeded levels seen in other auction institutions. Consequently, prices were less accurate in reflecting the underlying marginal abatement costs.

Based in part on these experiments, the US General Accounting Office recommended a switch to more standard uniform-price rules (US GAO 1994). The EPA determined that the statutory language of the CAAA permitted it to switch to a uniform-price rule, and it solicited comments to a proposal to make that switch because the new price rule "may have the benefit of producing a less confusing price signal" (Federal Register, 1996). The price rule was never changed, however, because the annual permit auction was largely abandoned for inter-firm trading and, in 2010, the SO_2 permits lost all value because of other regulations that limited emissions. Nevertheless, the experimental testbeds did encourage the EPA to rethink their auction pricing rules (Schmalensee and Stavins 2013, 2017).

In other emissions trading programs, regulators engaged with experimentalists to assist in the market design process. For example, to help design their auction for NO_x allowances, Virginia's Department of Environmental Quality (DEQ) commissioned a series of experiments (Porter et al. 2009). The experiments compared revenue and efficiency of three auction mechanisms—a combinatorial auction, a sealed-bid auction, and an English clock auction. The DEQ sought an auction design that would maximize revenue and ensure that permits went to firms with the highest value. The experiment results indicated that the English clock auction generated the highest revenues, which is the format that the Virginia DEQ chose to implement.

In another example of regulators engaging with experimentalists, the New York State Energy Research Development Authority funded an extensive series of auction design experiments for selling CO_2 emission allowances in the Regional Greenhouse Gas Initiative (RGGI) (Holt et al. 2007). The experiments investigated aspects of price discovery, including how realized transaction prices responded to unanticipated shifts in demand. The study recommended a uniform price sealed bid auction, which was adopted for RGGI, as well as subsequently by the EU Emissions Trading System and the California greenhouse gas market (Assembly Bill 32).[11] This recommendation was further supported by experiments in which bidders could tacitly or explicitly collude in uniform and discriminatory sealed-bid and clock auctions (Burtraw et al. 2009).

More recently, two sets of experiments made recommendations for changes to permit auction rules, which were subsequently adopted in emissions trading programs. The first set of experiments tackled a common problem in greenhouse gas emissions trading programs in which permits are over-supplied permits in the early years, leading to extremely low prices (Narassimhan et al. 2018). To address this problem, Burtraw et al. (2022) demonstrated the effectiveness of adjusting allowance auction supply based on previous years' prices and banking. This supply adjustment process was adopted recently both by RGGI and the California market.[12] Another direct way of affecting prices is through price floors, price ceilings, or their combination in a price "collar." Holt and Shobe (2016) report an experiment that compares a price collar to a quantity collar, which adjusts the auction quantity based on the stock of banked permits. The price collar reduces price variability and increases efficiency; nevertheless, a quantity collar ("market stability reserve") was adopted by the EU Emissions Trading Scheme in 2019.

Betz et al. (2010; 2017) report experiments that test different auction designs for CO_2 permit allocation in Australia, comparing sealed-bid and alternative clock formats, as well as simultaneous and sequential bidding. The clock format performed best in simultaneous auctions, but switching to a sequential procedure gave an efficiency advantage to sealed-bid rules.

[11] Personal communication with C. Holt (May 3, 2023).
[12] Personal communication with C. Holt (May 3, 2023).

Recommendations based on the experimental results were followed closely in the subsequent draft legislation creating the CO_2 trading scheme (Australian Government 2013). However, after the 2013 election, a new, conservative-led coalition government abolished the carbon pricing mechanism.

The Australian government also used experiments to help design a credit trading trial to address dryland salinity, comparing uniform price (i.e., same price to all sellers) and discriminant price (i.e., prices match the offers of sellers) rules and different information conditions during bidding (Connor et al. 2008). Results led to adoption of closed book, uniform price rules for field implementation, and benefit-cost analysis estimated net positive benefits for the program overall.

Conservation Auctions

Experiments have also informed conservation auction designs. One important example is the Georgia Irrigation Reduction Auction. In 2000, Georgia passed a law that required the state's Environmental Protection Division (EPD) to use an unspecified "auction-like process" to pay some farmers to suspend irrigation in drought years. In the auction, farmers, who often held multiple irrigation permits, would compete for payments from the EPD to voluntarily cease to irrigate their land from a permitted well for the remainder of the calendar year.

To testbed alternative auction designs, the EPD contracted with experimentalists (Cummings, Holt, and Laury 2004). Given that farmers can communicate before and during the auction, the researchers sought a "collusion-proof" design and allowed face-to-face communication during the experiment. Among the design issues investigated were the auction tie-breaking rule, uniform vs. discriminative pricing rules, and how winning bidders were announced each round. To help EPD officials better understand and trust the results of the experiments, the experimentalists invited the EPD director and staff to attend a laboratory session and to review the study's findings. Based on the experimental results, the EPD implemented the experimenters' recommended procedures: a discriminative price auction with revisions, no maximum accepted price announcement, and a random tie-breaking rule (Cummings, Holt, and Laury 2004). The EPD director, however, did deviate from one recommendation while the auction was running: he did not apply a consistent acceptance rule across revision rounds in order to pursue non-economic objectives.

To help design the BushTender conservation auctions, which aimed to improve biodiversity, the Department of Natural Resources and Environment in Victoria, Australia commissioned a series of laboratory experiments (Cason, Gangadharan, and Duke 2003). In conservation auctions that differentiate offers by their environmental benefits, an important design decision is whether to reveal the benefit information to bidders. In the BushTender experiments, some bidders were randomly assigned to receive their benefit information. When that information was revealed, sellers' offers misrepresented their costs more for high-benefit projects, consistent with rent-seeking incentives. Consequently, when the regulator revealed benefits to sellers, the overall environmental benefits acquired in the laboratory conservation auctions were lower (and seller profits higher). The experimental results thus led to a recommendation to withhold benefit information from bidders, and the government followed this advice when they implemented the policy in Victoria (Stoneham et al. 2003). Analysis of the auction data led to estimates of substantial cost savings relative to fixed-rate payments, which was the prevailing mechanism for subsidizing conservation activities at the time (Schilizzi and Latacz-Lohmann 2007).

Auctions can also be used to encourage spatial coordination, which may increase the ecological benefits of conservations efforts (for a review, see Hanley 2025). Rolfe, McCosker, and Windle (2008) describe pilot tests conducted with landholders. The goal was to create "corridors" protecting native vegetation, which required coordination among adjacent parcels with different owners. Insights from the experiment were used in a full-scale rollout in Queensland, Australia leading to an auction design with multiple bidding rounds to help landholders learn about the process and its risks, and to provide feedback about the locations of other bid proposals (Windle et al. 2009).

Finally, Balmford et al. (2023) use a combination of prior auction theory and lab experiments to persuade an industry partner to change a Payments for Ecosystem Services (PES) auction to reduce nutrient runoff in the field from a discriminative (pay-as-bid) to a uniform pricing rule. Their analysis focuses on a comparison of auction behavior and outcomes following this switch to earlier results using pay-as-bid rules, as well as a "control" catchment that maintained these older pricing rules. The positive results from this natural experiment (lower bids, lower payments to farmers and less strategic behavior) led to continued use of uniform pricing rules in the tested catchment as well as the adoption of these rules as the pricing default as these PES schemes expand geographically.[13]

Environmental Programs Directed at Individuals and Households

Experiments have been important in the context of designing interventions that encourage individuals and households to limit their use of natural resources, and to encourage participation in environment-related programs. Some of these studies develop and evaluate the impact of

[13] A subset of these authors has also conducted experiments to study an innovative pricing rule for two-sided PES auctions in which multiple developers must obtain offsets for their impacts (Balmford, Day, and Lindsay 2024). The novel design addresses the exposure problem facing bidders using package bidding to encourage truthful revelation of costs and values. These promising results have motivated the same industry partner as in Balmford et al. (2023) to pilot this new auction format in the field (personal communication with B. Balmford, September 4, 2023).

information “nudges,” which Thaler and Sunstein (2008, p. 6) define as “any aspect of choice architecture that alters people’s behavior in a predictable way without forbidding options or significantly changing their economic incentives.” For example, Schultz et al. (2007) study the effects of normative and injunctive messaging on residential energy use. They find that only providing information on average electricity usage decreased use among above average energy consumers and increased use with below average consumers. For the latter group, the uptick in electricity use was offset when the normative message was combined with a happy face emoticon. This paper influenced companies that work with public utilities, such as Opower (see Schultz et al. 2018) and VertexOne (previously WaterSmart), to use social norms to encourage conservation behavior. A report by WaterSmart (2014) documents that they scaled up the Schultz et al. (2007) methodology and sent “Home Water Reports” that reduced water consumption by up to 5%; Opower interventions cut energy usage by 2 to 5% (Schultz et al. 2018).

The European Commission Directive 2009/72/EC requires EU member countries to evaluate the roll out of smart electricity meters and, subject to a favorable benefit-cost analysis, to equip at least 80% of customers by 2020. To meet these objectives, Ireland’s Commission for Energy Regulation, in coordination with various industry partners, conducted separate randomized controlled trials (RCTs) for electricity and natural gas to test the effects of detailed energy use statements, billing frequency, smart meters, and variable pricing schemes (CER 2011a, CER 2011b). For both energy sources, the field trials showed that using a combination of a detailed use statement, bi-monthly billing and smart meters resulted in significant reductions. Based on these results, Ireland committed to a nationwide roll-out of smart meter technology.

In principle, improvements in home energy efficiency offer the potential to reduce homeowner’s energy bills while also lowering greenhouse gas emissions, yet most policies designed to improve home energy efficiency have not met expectations (Levinson 2016). As part of the Energy Policy Institute at the University of Chicago's (EPIC) efforts to provide evidence-based recommendations to the US government (EPIC 2021a), Burlig (2021) offers recommendations on how to improve the performance of home energy efficiency programs based in part on experimental evidence by Fowlie, Greenstone, Wolfram (2018) and by Allcott and Greenstone (2023). These recommendations include requiring federal programs to incorporate real-world evaluation to identify which retrofits have the highest return, as well as to identify implementation failures. As documented in EPIC (2021b), the Infrastructure Investment and Jobs Act established a grant program that funds commercial and residential energy audits, as well as upgrades and retrofits. Within a year after the upgrades are made, a final audit is required to assess the total energy savings from the upgrades, reflecting Burlig's recommendation.

The Behavioural Insights Team (BIT), originally created by the United Kingdom (UK) government but now operating as an independent consultancy, applies a behavioral science approach to address public policy problems. One environment-related issue BIT addressed with experiments was the sustainable consumption of electronic and household electric appliances. BIT used laboratory experiments with consumers to help the French government develop the repairability index, which informs consumers about the repairability of electronics and electric products, and then evaluated the most promising options with an RCT (BIT 2024). The repairability index came into force on January 1, 2021, and a post-implementation evaluation suggested that the share of more repairable products sold increased, and producers also responded by building newer models that were more repairable (BIT 2024).

In a second project, the BIT partnered with Chicago's Department of Water Management (DWM) to design and test, with an RCT, methods for increasing return rates for people who had requested a drinking water test kit to detect lead (BIT 2022). One intervention redesigned the test kit to make it easier for people to complete the test, and a second was to send text message reminders to return the kit. The new kit increased return rates by over 20 percentage points, with an additional 3.8 percentage point increase when reminders were sent. The results have motivated the DWM to adopt the new kits and to text reminders (BIT 2022).

In a third example, BIT worked with the City Budget Office in Portland, Oregon to design programs to attract new members to a bicycle sharing program (Kirkman 2019). While framing a $12 discount offer (sent via postcard) as a "discount" rather than as an offer to get services for "free" did not have any impact on take-up rates, the results of an RCT found that the redemption rate was four times higher for households that recently moved to the area relative to existing residents. As a result of the study, the city continues to target new movers with this promotion (Kirkman 2019).

Regulatory Enforcement and Compliance

In the context of regulatory enforcement to promote compliance, several experiments have led to adoption or modification of rules. These changes affected both consumers and producers.

On the consumer side, a recent example concerns enforcement of restrictions on residential outdoor water use during drought conditions (Browne et al. *forthcoming*). Although smart meters lower marginal monitoring costs to near zero and have been available in some areas for over a decade, many communities continue to rely on costly in-person monitoring. In a field experiment, households in Fresno, California were randomly assigned to different combinations of automated and in-person inspections, and fine levels. Automated enforcement increased households' punishment rates from 0.1 to 14%, decreased summer water use by 3%, and reduced violations by 17%. By contrast, higher fine levels had little effect. Automated enforcement also increased customer complaints by over a thousand percent, however, ultimately causing its cancellation and highlighting that political considerations can limit technological solutions to enforcement challenges (Browne et al. *forthcoming*).

In another consumer context, a field experiment was deployed to shed light on the effectiveness of using a threat of punishment to promote household waste sorting (Vollaard and van Soest 2024). The experiment randomized the timing of a warning letter, followed by a one-month period of highly visual and intensive law enforcement, across 65 waste collection routes in Tilburg, Netherlands. These interventions led to a persistent drop in residual waste of 10 to 15 percent. The favorable financial performance of this enforcement program identified by the experiment led to its implementation and expansion to additional areas (van Oirschot 2015).

On the firm side, Duflo et al. (2013) report a two-year experiment in Gujarat, India that varied the incentives of auditors for environmental performance. Half of the auditors for the 473 plants in the study were hired and paid by the plants, which was the status quo, while the other half were paid fixed fees and received incentive payments for accurate reports. Results indicated that the status quo system was clearly corrupted, as auditors systematically reported emissions marginally below the standard, even though actual emissions were generally higher. In the treatment condition auditors reported more truthfully, and this significantly reduced the false compliance rate; consequently, treated plants reduced their pollution emissions. The state used these results to improve monitoring and enforcement procedures, and air and water pollution at plants receiving the new audit format were significantly lower within six months, reducing air pollution by 28 percent.[14]

[14] https://urbanlabs.uchicago.edu/projects/gujarat-auditing-the-air (last accessed July 25, 2024).

Benami et al. (2023) report on an experiment that studies how text message reminders can improve on-time submission of discharge reports for compliance with the US Clean Water Act. Collaborating with the US EPA, the researchers customized messages ahead of reporting deadlines to enhance the firms' perceived probability of violation detection. Although treatment effects were heterogeneous, the pre-deadline reminders increased overall on-time submissions by almost 7 percentage points, translating to a 10 to 20 percent reduction in late submissions relative to baseline. The administrators of the EPA's office responsible for receiving electronic discharge monitoring reports indicated that based on this research, they have added functionality to the system to issue pre-deadline reminders for report submission to enhance compliance.[15]

Nonmarket Demand for Environmental Goods and Services

Environmental policies generate benefits that cannot be measured with market prices, and environmental economists have led the development of nonmarket valuation methods for monetizing these benefits. SP survey experiments represent the primary approach for quantifying both use and non-use values associated with nonmarket goods and services.[16] Nonmarket demand estimates derived from SP surveys mainly enter the policy arena through their inclusion in government benefit-cost analysis and other forms of regulatory analysis.[17,18]

---

[15] Personal communication with E. Benami (September 7, 2023).

[16] Laboratory experiments have long been important to the development of SP methods, e.g., in the context of identifying and mitigating hypothetical bias, and for testing alternative value elicitation mechanisms (see, for example, Harrison 2006).

[17] Most valuation scenarios in SP surveys are hypothetical in the sense that the relevant authorities are not considering the exact policies described. Nevertheless, most participants view their responses to high-quality SP surveys as consequential (e.g., Herriges et al. 2010), and survey results can influence policy decisions as documented in this article.

[18] SP estimates are also used in natural resource damage assessment (NRDA), which is a legal process employed by US federal agencies to determine monetary damages associated with harm to the environment under the Oil Pollution Act and Comprehensive Environmental Response, Compensation, and Liability Act. Kopp and Smith (1989) document two early NRDAs, both of which used results from a SP survey conducted on behalf of the plaintiff of one of the cases (Rowe et al. 1985).

In the US, Circular A-4 is the Office of Management and Budget's guidance document to Federal agencies on how to conduct regulatory analysis (OMB 2023). As broad evidence of policy impact, the Circular states that SP methods "have been widely used in regulatory analyses by Federal agencies, in part because these methods can be employed to address a wide variety of goods and services that are not easy to study through revealed preference methods" (OMB 2023, p. 34). As additional summary evidence, Petrolia et al. (2021) review 49 major rules issued by the EPA between fiscal years 2008 and 2019 for which regulatory impact analyses (RIAs) were conducted. SP surveys represent the most frequently used nonmarket valuation method and figured in 36 of the 49 rules. SP results were used exclusively when estimating benefits associated with decreased morbidity, and when valuing ecological benefits.

In the UK, The Green Book provides guidance issued by the HM Treasury (2022) on policy evaluation methods and includes an extensive supplement on how to conduct and evaluate SP research. Atkinson et al. (2018) evaluate the extent to which environmental valuation has influenced UK policy decisions. In the case of water policy, the authors document extensive use of benefit-cost analysis by the Environment Agency, as well as private water companies, and in this context the "tradition" has been to use SP methods. In some cases, the benefit-cost analysis appears to have merely been undertaken to fulfill policy mandates, while in others the analysis was an input to help determine policy design (e.g., which policy alternative to pursue). In the case of "natural capital" policies, benefit-cost analyses have further influenced government budget allocation decisions (Fenichel 2024; Talis et al. 2024).

When US federal agencies were first charged with using benefit-cost analysis for major or significant proposed regulations in the 1980s, SP research was in an early stage of development. This promulgated the use of either one or a small set of existing studies, or new studies funded specifically for use in regulatory analysis. In the context of early surface water quality regulations, as documented by Griffiths et al. (2012), the 1982 Iron and Steel Manufacturing Effluent Guidelines utilized the analysis of Freeman (1979), and nonmarket benefits for four rules were exclusively based on a nationwide SP survey funded by the EPA (Mitchell and Carson 1984; Carson and Mitchell 1993). Fisher and Raucher (1984) provide a rule of thumb that the EPA followed in several subsequent analyses, specifically that non-use benefits are equal to 50 percent of use benefits.[19]

Several SP studies have served as the exclusive or primary source for estimating nonmarket benefits associated with dams. Loomis (1996) measured the benefits of restoring the Elwha River (Washington state) by removing dams. This study was highlighted in the environmental impact statement on the same issue, where it was emphasized that expected benefits exceed the cost even if nonmarket values were as low as 0.5 percent of the SP estimate (National Park Service 1995). Both dams on the Elwha River have since been removed. Welsh et al. (1995) estimate the total economic value for changes in the operation of the Glen Canyon Dam on the Colorado River. Then Secretary of Interior Bruce Babbitt used the study results to help defend his decision to reduce daily flow fluctuations (US DOI 1996). Information valuing recreation benefits from Adamowicz, Louviere, and Williams (1994) were built into a benefit-cost analysis that supported the Highwood/Little Bow water resource development project in southern Alberta (Canada). The executive summary of the Review Panel (NRCB/CEAA 1997) recommended to the Federal Government that, in part based on the "social" effects of the project (i.e., nonmarket values), the proposed project is in the public interest.

[19] Loomis (2001) provides several early examples of how state and federal agencies have used SP results, including some cases in which individual studies most likely influenced policy decisions.

SP studies have been used as inputs into decisions involving species protection. A survey by Jones & Stokes Associates (1994), as discussed by Loomis (2000), helped inform the decision to increase the water flows into Mono Lake while reducing the water rights of Los Angeles to protect migratory birds. The California State Resources Board (1994) issued a decision that the lake must rise to 6392 feet, which coincides closely with the conclusion of the SP study that the 6390 feet elevation yielded the highest benefits. As a second example, an earlier working paper version of Entem et al. (2022) is heavily referenced as providing the benefit-cost analysis of the proposed action plan for the "south of the divide" threatened species case in Saskatchewan, Canada (Environment and Climate Change Canada 2017). The RIA statement (Canada Gazette 1998) in support of the Woodland Caribou (Rangifer tarandus caribou) Boreal Population Order, which was later put into force in June 2019, highlighted willingness-to-pay estimates for Caribou conservation adapted from SP studies by Tanguay, Adamowicz, and Boxall (1995) and Adamowicz et al. (1998).

There are yet other examples of individual studies that were potentially influential through their inclusion in benefit-cost analyses. Bateman et al. (1995) estimate the value of protecting the Norfolk Broads wetland area (England) from flooding from the North Sea. The postscript to Bateman, Langford, and Rasbash (2001) states, "Given our findings, the National Rivers Authority chose to use the downwardly biased [open ended] WTP measure in their attempt to argue the case for funding from the UK Treasury…. As a consequence the principle of funding flood defences in Broadland has now been accepted and the construction of new defences has recently commenced." A study conducted in Rhode Island (Johnston et al. 2012; Zhao, Johnston, and Schultz 2013) was used to determine total economic benefits associated with the Final 2014 Rule for Existing Electric Generating Plants and Factories (US EPA 2014).

Similarly, a wetland valuation study conducted by Blomquist and Whitehead (1998) informed the analysis underlying the 2020 Navigable Waters Protection Rule (US EPA and DA 2020). Finally, DEFRA commissioned a study to provide benefit estimates of a drinks container deposit return scheme in the UK in terms of the reduced dis-amenity of litter and to help inform the question of which types of containers should be included (eftec 2020). Following several delays, UK's bottle and deposit return scheme is currently set to begin October 2027 (Comerford 2024).

In most cases it is not practical to conduct a case-specific SP study to inform a regulatory analysis. As documented by Griffiths et al. (2012), Atkinson et al. (2018), Petrolia et al. (2021), and Moore et al. (2023), the trend has been to use multiple existing studies to estimate benefits via a benefit transfer approach. For example, many government agencies use a central estimate of the value of a statistical life (VSL) when determining benefits associated with decreased mortality rates. While this mutes the importance of a single study within the context of a particular regulatory analysis, in most cases this means that a single study is used to inform myriad analyses. For example, the SP study by Viscusi, Magat, and Huber (1991) was included in the RIAs associated with 33 of the 49 major rules examined by Petrolia et al. (2021). In the UK, The National Water Environment Benefit Survey (NERA-Accent 2007; Metcalfe et al. 2012) is another example of a "workhorse" study used in various analyses involving program appraisals and assessments, due in part to the fact that estimated values can be adapted to fit different geographic and ecological scales, and are based on SP valuation scenarios congruent with the European Union Water Framework Directive.

Many valuation tools are now available to aid decision makers, which has further increased the use of SP experiments. The Environment Valuation Reference Inventory (EVRI), which is administered by Environment and Climate Change Canada, characterizes and provides summaries for over 4500 primary studies, of which over 70 percent involve SP methods. The EPA has developed platforms to aid the estimation of human health impacts from air quality changes (BenMAP), and to quantify the benefits associated with changes in water quality indicators (HAWQS and BenSPLASH). In the UK, the costs and benefits associated with land-use change can be estimated using the online software NEVO. Estimates obtained from SP surveys inform each of these, and several other, decision tools.

SP experiments have also influenced public decisions outside of the specific benefit-cost context. Lee, Steinback, and Wallmo (2017) describe a bioeconomic model that is currently used by the New England Fishery Management Council to inform recreational fishing regulations for Gulf of Maine cod and haddock. The recreation demand model is parameterized using a SP survey that is updated every 3-4 years. In 2023, the same modeling approach was used to set fishing regulations for summer flounder, black seas bass, and scup in the Northeastern states. NOAA Fisheries is currently working to expand this modeling approach to other US fisheries.[20]

Aadland et al. (2012) conducted a nationwide study to estimate the WTP for a new interagency pass to replace the National Parks Pass for recreational access to public lands. They estimated that, among households that had previously purchased a National Parks Pass, the average household would pay more than $80 to replace the Pass they previously purchased. Based on this, the federal agencies set the price of the new pass (now called the America the Beautiful National Parks and Federal Recreational Lands Pass) at $80 (Aadland et al. 2012).

[20] Personal communication with Scott Steinback (July 27, 2023).

The Future Experimental-Environmental Policy Nexus

While we have documented many experiment-policy links, experimental evidence is not routinely used to inform the design of environmental policies, and it is rare for experimentation to be embedded into environmental programs for the purpose of quantifying their impacts (Ferraro et al. 2023). Indeed, our outreach efforts most often led to communications about promising experimental results that were effectively ignored. As one of many examples, Ferraro and Price (2013) partnered with the Cobb County Water System on an RCT designed to examine the effectiveness of norm-based messages on managing residential water demand. Even though social comparison messaging lowered water use by about 4% during droughts, the County never used those comparisons during its later droughts.[21]

The good news for researchers is that demand for experimental research appears to be increasing. In the US, Executive Order 13707 encourages agencies to identify policies where applying behavioral science may substantially improve welfare, and, where possible, to evaluate the impact of these applications (US President 2015). The Foundations for Evidence-Based Policymaking Act of 2018 promulgates the need for data-driven policy decisions and fosters a culture of experimentation both directly and indirectly (Ferraro et al. 2023). Nevertheless, opportunities for policy impact do not necessarily translate into links themselves. We conclude this article with recommendations for how we can better cultivate demand for, and increase the impact of, experimental research.

Two key recommendations are for researchers to actively seek out and document policy links with their research, and for public decision makers to better acknowledge the influence of research findings. In our various communications with researchers, it was clear that most did not

[21] Personal communication with Paul Ferraro (July 31, 2024).

already make a habit of documenting impacts. Instead, many of the experiment-policy links we describe resulted from investigative work among those we contacted, either through those in the policy sphere digging through their meeting notes or internal documents, or researchers asking their industry or government contacts how their past work was influential. In other cases, the individuals we communicated with were convinced their research was influential, but ultimately were unable to document links; this was mainly due to losing contact with project officers or public decision makers or failing to keep older communications (e.g., emails, reports, etc.). Overall, this highlights the importance of researchers building and maintaining relationships with policymakers. This would enable economists to provide policymakers with insights into how experiments can be used not only to evaluate the success or failure of a specific policy, but also to better understand the mechanism behind it. Such collaboration would most likely encourage policymakers to involve economists during the early design phase of new policies, allowing for robust experimentation before implementation. Furthermore, a stable relationship with policymakers would allow for investigation into how the policy works, particularly in terms of unintended behavior among people. This could refine future design and implementation of policies.

Academic researchers have not traditionally been incentivized to produce, let alone document, policy impacts from their research. Nevertheless, the landscape has begun to change, and there are likely to be dividends for researchers who pivot sooner than later. In the 2021 Research Excellence Framework (REF) in the UK, 25% of an institution's evaluation was based on impact, defined as "an effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia" (REF 2021, p. 2). The Engagement and Impact assessment in Australia also encourages greater collaboration between universities and users of research. As documented in Durrant and MacKillop (2022), this emphasis on the social impact of research has, in turn, motivated some universities to implement systems for measuring, assessing, and reporting impact.

A third recommendation is for researchers to communicate their work (and that of others) in ways that are often undervalued by academia but are conducive to policy impact. One mechanism is through survey articles that disseminate key insights substantiated from multiple studies. Prime examples of this include Noussair and van Soest (2014), List and Price (2016), Stranlund (2017), Bouma (2021), and Banerjee (2022), all of which communicate important experimental findings in accessible ways, e.g., in terms of "lessons" for policymakers. Other avenues include non-technical policy briefs, and various forms of internet media (podcasts, blogs, etc.) that can readily be uncovered through search engines. Some of the experiment-policy links we reference arose from these alternative forms of communication. For example, the websites of the BIT and the University of Chicago Urban Labs provided easily accessible and digestible information on research projects and their policy impacts.

Fourth, we recommend that researchers interested in informing policy not only identify emerging policy issues but generate the right data to address policy needs. In some cases, those working in government agencies have directly collaborated with researchers in identifying and disseminating information on research needs (e.g., Moore et al. 2023). Generating the right data further requires considering scalability issues. As articulated by List (2024), researchers can produce policy-relevant data by imagining what a fully implemented intervention looks like and testing accordingly. This mindset has important implications for what data to generate and in what sequence. Moreover, many governments now require that regulatory analyses consider distributional effects (e.g., see OMB 2023). Experimental studies that clearly document welfare impacts on different groups can help policymakers appreciate the political feasibility of alternative policies.

Last, we encourage research that helps identify how to best garner the attention of decision makers about research findings (e.g., emails versus tweets), and in turn what approaches ultimately lead to policy impact. As one example, Scott et al. (2023) find that a new policy engagement method, which involves sending timely and relevant emails in tandem with short fact sheets and briefs, increased US state legislatures' use of technical research language and research-based concepts in social media posts about the COVID-19 pandemic. We believe that experimentation in this domain represents an essential, but untapped, use of experiments for those who endeavor to inform environmental policy and natural resource management decisions.

References

Aadland, David, Bistra Anatchkova, Burke D. Grandjean, Jason F. Shogren, Benjamin Simon, and Patricia A. Taylor. 2012. Valuing access to US public lands: A pricing experiment to inform federal policy. *Social Science Quarterly 93* (1): 248-269.

Adamowicz, W., J. Louviere, M. Williams. 1994. Combining Revealed and Stated Preference Methods for Valuing Environmental Amenities. *Journal of Environmental Economics and Management* 26 (3): 271-292.

Adamowicz, Wiktor, Peter Boxall, Michael Williams, and Jordan Louviere. 1998. Stated Preference Approaches for Measuring Passive Use Values: Choice Experiments and Contingent Valuation. *American Journal of Agricultural Economics* 80 (1): 64-75.

Al-Ubaydli, Omar, John A. List, and Dana Suskind, D. 2019. The science of using science: Towards an understanding of the threats to scaling experiments (No. w25848). National Bureau of Economic Research.

Al-Ubaydli, Omar, Min Sok Lee, John A. List, Claire L. Mackevicius, and Dana Suskind. 2021. How can experiments play a greater role in public policy? Twelve proposals from an economic model of scaling. *Behavioural Public Policy* 5 (1): 2-49.

Allcott, Hunt, and Michael Greenstone. 2023. Measuring the welfare effects of residential energy efficiency programs. NBER WP No. w23386.

Alpizar, Francisco, María Bernedo Del Carpio, and Paul J. Ferraro. 2024. Input efficiency as a solution to externalities and resource scarcity: A randomized controlled trial. *Journal of the Association of Environmental and Resource Economists* 11(1), 171-211.

Atkinson, Giles, Ben Groom, Nicholas Hanley, and Susana Mourato. 2018. Environmental Valuation and Benefit-Cost Analysis in U.K. Policy. *Journal of Benefit Cost Analysis* 9 (1): 97-119.

Australian Government. 2013. Clean Energy Act (Auction of Carbon Units) Determination 2013. https://www.legislation.gov.au/F2013L00759/latest/text.

Baicker, Katherine, Sarah L. Taubman, Heidi L. Allen, Mira Bernstein, Jonathan H. Gruber, Joseph P. Newhouse, Eric C. Schneider, et al. 2013. The Oregon experiment—effects of Medicaid on clinical outcomes. *New England Journal of Medicine* 368 (18): 1713-1722.

Balmford, Ben, Joseph Collins, Brett Day, Luke Lindsay, and James Peacock. 2023. Pricing rules for PES auctions: Evidence from a natural experiment. *Journal of Environmental Economics and Management* 122: 102889.

Balmford, Ben, Brett Day, and Luke Lindsay. 2024. Designing double-sided markets for bundled environmental goods. Working paper, University of Exeter.

Bateman, Ian J., Ian H. Langford, R. Kerry Turner, Ken G. Willis, and Guy D. Garrod. 1995. Elicitation and truncation effects in contingent valuation studies. *Ecological Economics* 12 (2): 161-179.

Bateman, Ian J., Ian H. Langford, and Jon Rasbash. 2001. Willingness-To-Pay Question Format Effects in Contingent Valuation Studies. In *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries*, Ian J. Bateman and Kenneth G. Willis (eds.), p. 511-539. Oxford University Press.

Banerjee, Simanti. 2022. Use of Experimental Economics in Policy Design and Evaluation: An Application to Water Resources and Other Environmental Domains. In *The Oxford Research Encyclopedia of Environmental Science*, H. Shubart (ed.). Oxford University Press.

Benami, Elinor, Nathanael Jo, Elizabeth Ragnauth, and Daniel E. Ho. 2023. Drop a Line, Submit on Time? Randomized Tailored Reminders Improve Pollution Reporting Timeliness. Working paper, available at SSRN: http://dx.doi.org/10.2139/ssrn.4355984.

Betz, Regina, Karl-Martin Ehrhart, Ben Greiner, Andreas Ortmann, Sascha Schweitzer, and Stefan Seifert. 2010. Experimental testing of possible designs for the Australian carbon pollution permit allocation auction. Report prepared for: Department of Climate Change and Energy Efficiency. https://www.ceem.unsw.edu.au/sites/default/files/uploads/publications/PittSherryReport-20110525-PDF.pdf.

Betz, Regina, Ben Greiner, Sascha Schweitzer, and Stefan Seifert. 2017. Auction Format and Auction Sequence in Multi-item Multi-unit Auctions: An Experimental Study. *Economic Journal* 127 (605): F351-F371.

BIT (Behavioural Insights Team). 2024. Leveraging behavioural insights to design and test the repairability index in France. Report. March 13, 2024. https://www.bi.team/publications/leveraging-behavioural-insights-to-design-and-test-the-repairability-index-in-france/.

BIT (Behavioural Insights Team). 2022. Increasing lead kit return rates among Chicago residents. Report. August 19, 2022. https://www.bi.team/wp-content/uploads/2023/09/2022-08-WWC-Chicago-DWM_Lead-Testing-Kits-Final-Report.pdf.

Blomquist, Glenn C., and John C. Whitehead. 1998. Resource quality information and validity of willingness to pay in contingent valuation. Resource and Energy Economics 20 (2): 179-196.

Bouma, Jetske. 2021. Evaluation environmental policy: The use and usefulness of experiments. *Journal of Environmental Economics and Policy* 10 (4): 468-480.

Brent, Daniel A., Lana Friesen, Lata Gangadharan, and Andreas Leibbrandt. 2017. Behavioral insights from field experiments in environmental economics. *International Review of Environmental and Resource Economics* 10 (2): 95-143.

Browne, Oliver R., Ludovica Gazze, Michael Greenstone, and Olga Rostapshova. Man vs. Machine: Technological Promise and Political Limits of Automated Regulation Enforcement. *The Review of Economics and Statistics*, forthcoming.

Burlig, Fiona. 2021. Making Energy Efficiency Work. In *U.S. Energy & Climate Roadmap: Evidence-based Policies for Effective Action*, edited by Energy Policy Institute at the University of Chicago.

Burtraw, Dallas, Jacob Goeree, Charles A. Holt, Erica Myers, Karen Palmer, and William Shobe. 2009. Collusion in Auctions for Emission Permits: An Experimental Analysis. *Journal of Policy Analysis and Management* 28 (4): 672-691.

Burtraw, Dallas, Charles A. Holt, Karen Palmer, and William Shobe. 2022. Price-responsive allowance supply in emissions markets. *Journal of the Association of Environmental and Resource Economists* 9 (5): 851-884.

California State Water Resources Control Board. 1994. Decision 1631, Decision and order amending water right licenses to establish fishery protection flows in streams tributary to Mono Lake and to protect public trust resources at Mono Lake and the Mono Lake Basin. Sacramento, CA: California Water Resources Control Board. https://www.waterboards.ca.gov/waterrights/board_decisions/adopted_orders/decisions/d1600_d1649/wrd1631.pdf

Canada Gazette, Part I, Volume 152, Number 49: Critical Habitat of the Woodland Caribou (Rangifer tarandus caribou) Boreal Population Order.

Cárdenas, Juan Camilo. 2009. Experiments in environment and development. *Annual Review of Resource Economics* 1 (1): 157-182.

Carlson, Dale, Charles Forman, Nancy Olmstead, John O. Ledyard, Charles R. Plott, David P. Porter, and Anne Sholtz. 1993. *An analysis and recommendation for the terms of the RECLAIM trading credit*. Report to the South Coast Air Quality Management District, Diamond Bar, CA.

Carlsson, Fredrik, and Olof Johansson-Stenman. 2012. Behavioral economics and environmental policy. *Annual Review of Resource Economics* 4 (1): 75-99.

Carson, Richard T., and Robert Cameron Mitchell. 1993. The Value of clean water: The public's willingness to pay for boatable, fishable, and swimmable quality water. *Water Resources Research* 29 (7): 2445-2454.

Cason, Timothy N. 1993. Seller incentive properties of EPA's emission trading auction. *Journal of Environmental Economics and Management* 25: 177-195.

Cason, Timothy N. 1995. An experimental investigation of the seller incentives in the EPA's emission trading auction. *American Economic Review* 85 (4): 905-922.

Cason, Timothy N., and Charles R. Plott. 1996. EPA's new emissions trading mechanism: A laboratory evaluation. *Journal of Environmental Economics and Management* 30 (2): 133-160.

Cason, Timothy N., Lata Gangadharan, and Charlotte Duke. 2003. A laboratory study of auctions for reducing non-point source pollution. *Journal of Environmental Economics and Management* 46 (3): 446-471.

Cason, Timothy N., and Steven Y. Wu. 2019. Subject pools and deception in agricultural and resource economics experiments. *Environmental and Resource Economics* 73: 743-758.

CER (Commission for Energy Regulation). 2011a. Smart Metering Information Paper 4: Results of Electricity Cost-Benefit Analysis, Customer Behaviour Trials and Technology Trials. Information Paper, CER11080, May 16, 2011. https://www.ucd.ie/issda/t4media/cer11080.pdf.

CER (Commission for Energy Regulation). 2011b. Smart Metering Information Paper: Gas Customer Behaviour Trial Findings Report. Information Paper, CER11180a, October 11, 2011. https://www.ucd.ie/issda/t4media/Gas%20Customer%20Behaviour%20Trial%20Findings%20Report.pdf.

Charness, Gary. 2015. The Hammer and the Screwdriver. In *Handbook of Economic Methodology*, Guillaume R. Fréchette and Andrew Schotter (eds.), p. 197-199. Oxford Academic Press.

Comerford, Ruth. 2024. UK cash-for-bottle deposit scheme delayed until 2027. BBC News. 25 April 2024. URL: https://www.bbc.com/news/uk-68898109.

Connor, Jeffrey D., John Ward, Craig Clifton, Wendy Proctor, and Darla Hatton MacDonald. 2008. Designing, testing and implementing a trial dryland salinity credit trade scheme. *Ecological Economics* 67 (4): 574-588.

Cummings, Ronald G., Charles A. Holt, and Susan K. Laury. 2004. Using laboratory experiments for policymaking: An example from the Georgia irrigation reduction auction. *Journal of Policy Analysis and Management* 23 (2): 341-363.

Deaton, Angus. 2020. Introduction: Randomization in the Tropics Revisited, a Theme and Eleven Variations, in Florent Bédécarrats, Isabelle Guérin, and François Roubaud (eds), *Randomized Control Trials in the Field of Development: A Critical Perspective* (Oxford; online edn, Oxford Academic), https://doi.org/10.1093/oso/9780198865360.003.0002.

Del Carpio, Maria Bernedo, Francisco Alpizar and Paul J. Ferraro. 2021. Community-based monitoring to facilitate water management by local institutions in Costa Rica. *Proceedings of the National Academy of Sciences of the United States of America* 118(29): e2015177118

Duflo, Esther. 2018. The economist as a plumber. *American Economic Review* 107(5): 1-26.

Duflo, Esther, Michael Greenstone, Rohini Pande and Nicholas Ryan. 2013. Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India. *The Quarterly Journal of Economics* 128 (4): 1499–1545.

Durrant, Hannah, and Eleanor MacKillop. 2022. University policy engagement bodies in the UK and the variable meanings of and approaches to impact. *Research Evaluation* 31 (3): 372–384. https://doi.org/10.1093/reseval/rvac015.

eftec (economics for the environment). 2020. Amenity Value Benefits of a Deposit Return Scheme for Drinks Containers. Technical Report, Department for Environment Food and Rural Affairs, July 2020. Defra EQ0126.

Ehmke, Mariah D., and Jason F. Shogren. 2009. Experimental methods for environment and development economics. *Environment and Development Economics* 14 (4): 419-456.

Entem, Alicia, Patrick Lloyd-Smith, Wiktor (Vic) Adamowicz, Peter C. Boxall. 2022. Using inferred valuation to quantify survey and social desirability bias in stated preference research. *American Journal of Agricultural Economics* 104 (4): 1224–1242.

Environment and Climate Change Canada. 2017. Action Plan for Multiple Species at Risk in Southwestern Saskatchewan: South of the Divide. Species at Risk Act Action Plan Series. Environment and Climate Change Canada, Ottawa. xi + 127 pp.

EPIC (Energy Policy Institute at the University of Chicago). 2021a. *U.S. Energy & Climate Roadmap: Evidence-based Policies for Effective Action*. https://epic.uchicago.edu/wp-content/uploads/2021/02/EPIC-Energy-and-Climate-Roadmap.pdf.

EPIC (Energy Policy Institute at the University of Chicago). 2021b. Evidence-based Policies to Tackle Today's Energy and Climate Challenges. Insight / Impact Take-Away, December 13, 2021. https://epic.uchicago.edu/insights/evidence-based-policies-to-tackle-todays-energy-and-climate-challenges/.

Federal Register, Environmental Protection Agency (40 CFR), Part 73, "Acid Rain Program, SO2 Allowance Auction and Electronic Allowance Transfer: Advanced Notice of Proposed Rulemaking," pp. 28995-28998 (June 6, 1996).

Ferraro, Paul J., Todd L. Cherry, Jason F. Shogren, Christian Vossler, Hilary Byerly Flint, Jacob P. Hochard, Olof Johansson-Stenman, Peter Martinsson, et al. 2023. Create a Culture of Experimentation in Environmental Programs. *Science* 381 (6659): 735-737.

Ferraro, Paul J., and Michael K. Price. 2013. Using nonpecuniary strategies to influence behavior: Evidence from a large-scale field experiment. *The Review of Economics and Statistics* 95 (1): 64-73.

Fenichel, Eli. P. 2024. A New Era of Economic Measurement for the Environment and Natural Capital. *Review of Environmental Economics and Policy* 18(2): 321-330.

Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, Katherine Baicker, Oregon Health Study Group. 2012. The Oregon health insurance experiment: evidence from the first year. *The Quarterly Journal of Economics* 127 (3): 1057-1106.

Fisher, A., and R. Raucher. 1984. Intrinsic benefits of improved water quality: Conceptual and empirical perspectives. In *Advances in applied micro-economics: A research manual*, V.K. Smith, ed. Greenwich, CT: JAI Press.

Fowlie, Meredith, Michael Greenstone, Catherine Wolfram. 2018. Do energy efficiency investments deliver? Evidence from the weatherization assistance program. *Quarterly Journal of Economics* 133 (3): 1597-1644.

Fréchette, Guillaume R. 2015. Laboratory Experiments: Professionals Versus Students. In *Handbook of Experimental Economic Methodology*, G.R. Fréchette and A Schotter, eds. New York: Oxford University Press.

Fréchette, Guillaume R., Kim Sarnoff, and Leeat Yariv. 2022. Experimental Economics: Past and Future. *Annual Review of Economics* 14: 777-794.

Freeman, A. Myrick III. 1979. The benefits of air and water pollution control: A review and synthesis of recent estimates. Washington, DC: Council on Environmental Quality.

Friesen, Lana, and Lata Gangadharan. 2013. Environmental markets: What do we learn from the lab?. *Journal of Economic Surveys* 27 (3): 515-535.

Gneezy, Uri, and Alex Imas. 2017. Lab in the field: Measuring preferences in the wild. In *Handbook of Economic Field Experiments*, vol. 1, A.V. Banerjee and E Duflo, eds. Amsterdam: North-Holland.

Griffiths, Charles, Heather Klemick, Matt Massey, Chris Moore, Steve Newbold, David Simpson, Patrick Walsh, and William Wheeler. 2012. US Environmental Protection Agency Valuation of Surface Water Quality Improvements. *Review of Environmental Economics and Policy* 6(1): 130-146.

Guala, Francesco. 2001. Building economic machines: The FCC auctions. *Studies in History and Philosophy of Science Part A* 32 (3): 453-477.

Hanley, Nick. 2025. Economic incentives for biodiversity conservation on farmland. Working paper, University of Glasgow.

Harrison, Glenn W. 2006. Experimental evidence on alternative environmental valuation methods. *Environmental & Resource Economics* 34: 125-162.

Harrison, Glenn W., and John A. List. 2004. Field Experiments. *Journal of Economic Literature* 42 (4): 1009-1055.

Henrich, Joseph, Steven J. Heine, and Ara Norenzayan. 2010. The Weirdest People in the World. *Behavioral and Brain Sciences* 33: 1-23.

Herriges, Joseph, Catherine Kling, Chih-Chen Liu, and Justin Tobias. 2010. What are the consequences of consequentiality? *Journal of Environmental Economics and Management* 59 (1): 67-81.

Higgins, Nathaniel, Daniel Hellerstein, Steven Wallander, and Lori Lynch. 2017. *Economic Experiments for Policy Analysis and Program Design: A Guide for Agricultural Decisionmakers*. ERR-236, US Department of Agriculture, Economic Research Service, August 2017.

HM Treasury. 2022. *The Green Book: Central Government Guidance on Appraisal and Evaluation*. London: HM Treasury.

Holt, Charles A., and William M. Shobe. 2016. Price and quantity collars for stabilizing emission allowance prices: Laboratory experiments on the EU ETS market stability reserve. *Journal of Environmental Economics and Management* 80: 69-86.

Holt, Charles, William Shobe, Dallas Burtraw, Karen Palmer, and Jacob Goeree. 2007. Auction design for selling CO2 emission allowances under the regional greenhouse gas initiative. RFF Report. Final Report October 2007.

Johnston, Robert J., Eric T. Schultz, Kathleen Segerson, Elena Y. Besedin, and Mahesh Ramachandran. 2012. Enhancing the Content Validity of Stated Preference Valuation: The Structure and Function of Ecological Indicators. *Land Economics* 88 (1): 102-120. https://doi.org/10.3368/le.88.1.102

Jones & Stokes Associates. 1994. Final Environmental Impact Report for the Review of Mono Basin Water Rights of the City of Los Angeles.

Kirkman, Elspeth. 2019. Free riding or discounted riding? How the framing of a bike share offer impacts offer-redemption. *Journal of Behavioral Public Administration* 2 (2): 1-10.

Kopp, Raymond J., and V. Kerry Smith. 1989. Benefit Estimation Goes to Court: The Case of Natural Resource Damage Assessments. *Journal of Policy Analysis and Management* 8 (4): 593-612.

Krueger, Alan B. 1999. Experimental estimates of education production functions. *The Quarterly Journal of Economics* 114 (2): 497-532.

Ledyard, John, David Porter, and Randii Wessen. 2000. A market-based mechanism for allocating space shuttle secondary payload priority. *Experimental Economics* 2: 173-195.

Lee, Min-Yang, Scott Steinback, and Kristy Wallmo. 2017. Applying a Bioeconomic Model to Recreational Fisheries Management: Groundfish in the Northeast United States. *Marine Resource Economics* 32 (2): 191-216.

Levinson, Arik. 2016. How Much Energy Do Building Codes Save? Evidence from California Houses. *American Economic Review* 106 (10): 2867-2894.

List, John A. 2024. Optimally generate policy-based evidence before scaling. *Nature* 626: 491-499.

List, John A., and Michael K. Price. 2016. The Use of Field Experiments in Environmental and Resource Economics. *Review of Environmental Economics and Policy* 10 (2): 206-225.

Loomis, John B. 1996. Measuring the economic benefits of removing dams and restoring the Elwha River: Results from a contingent valuation survey. *Water Resources Research* 32 (2): 441-447.

Loomis, John B. 2000. Environmental valuation techniques in water resource decision making. *Journal of Water Resources Planning and Management* 126 (6): 339-344.

Loomis, John B. 2001. Contingent valuation methodology and the US institutional framework. In *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries*, Ian J. Bateman and Kenneth G. Willis (eds.), p. 613-628. Oxford University Press.

Metcalfe, Paul J., William Baker, Kevin Andrews, Giles Atkinson, Ian J. Bateman, Sarah Butler, Richard T. Carson, Jo East, Yves Gueron, Rob Sheldon, and Kenneth Train. 2012. An assessment of the nonmarket benefits of the Water Framework Directive for households in England and Wales. *Water Resources Research* 48: W03526.

Mitchell, Robert C., and Richard T. Carson. 1984, Willingness to Pay for National Freshwater Quality Improvements. Draft report prepared for US Environmental Protection Agency, Washington, D.C. Resources for the Future, October 1984.

Moore, Chris C., Joel Corona, Charles Griffiths, Matthew T. Heberling, Julie A Hewitt, David A. Keiser, Catherine L. Kling, et al. 2023. *Proceedings of the National Academy of Sciences* 120 (18): e2120247120.

Narassimhan, Easwaran, Kelly S. Gallagher, Stefan Koester, and Julio Rivera Alejo. 2018. Carbon pricing in practice: A review of existing emissions trading systems. *Climate Policy* 18 (8): 967-991.

National Park Service. 1995. Final Environmental Impact Statement, Elwha River Ecosystem Restoration, Olympic National Park, Washington.

NERA-Accent. 2007. The benefits of Water Framework Directive programmes of measures in England and Wales, Rep. CRP 4b/c, Dept. for Environment, Food and Rural Affairs, London, U.K.

Normann, Hans-Theo, and Roberto Ricciuti. 2009. Laboratory experiments for economic policy making. *Journal of Economic Surveys* 23 (3): 407-432.

Noussair, Charles N. and Daan P. van Soest. 2014. Economic Experiments and Environmental Policy. *Annual Review of Resource Economics* 6: 319-337.

NRCB/CEAA (Natural Resources Conservation Board / Canadian Environmental Assessment Agency). 1998. Report of the NRCB/CEAA Joint Review Panel. Application #9601 – Alberta Public Works, Supply and Services, May 1998. Little Bow Project/Highwood Diversion Plan, Application to Construct a Water Management Project to Convey and Store Water Diverted from the Highwood River. https://www.nrcb.ca/download_document/2/182/5498/decision-report.

OMB (Office of Management and Budget). 2023. *OMB Circular No. A-4*. Washington, D.C.: Executive Office of the President, Office of Management and Budget.

Petrolia, Daniel R., Dennis Guignet, John Whitehead, Cannon Kent, Clay Caulder, and Kelvin Amon. 2021. Nonmarket valuation in the environmental protection agency's regulatory process. *Applied Economic Perspectives and Policy* 43 (3): 952-696.

Plott, Charles R. 1997. Laboratory experimental testbeds: Application to the PCS auction. *Journal of Economics & Management Strategy* 6 (3): 605-638.

Porter, David, Stephen Rassenti, William Shobe, Vernon Smith, and Abel Winn. 2009. The design, testing and implementation of Virginia's NOx allowance auction. *Journal of Economic Behavior and Organization* 69 (2): 190-200.

REF (Research Excellence Framework). n.d. REF 2021 Key Facts. URL: https://2021.ref.ac.uk/media/1848/ref2021_key_facts.pdf. Accessed July 21, 2024.

Rolfe, John, Juliana McCosker, and Jill Windle. 2008. Identifying the incentives that graziers in central-western Queensland need to manage land for conservation outcomes. *The Rangeland Journal* 30(3): 297-303.

Roth, Alvin E. 1986. Laboratory Experimentation in Economics. *Economics & Philosophy* 2 (2): 245-273.

Rowe, Robert D., William D. Schulze, Brian Hurd, and Douglas Orr. 1985. Economic assessment of damage related to the Eagle Mine facility. Boulder, CO: Energy and Resource Consultants, Inc.

Schilizzi, Steven, and Uwe Latacz-Lohmann. 2007. Assessing the Performance of Conservation Auctions: An Experimental Study. *Land Economics* 83 (4): 497-515.

Schmalensee, Richard, and Robert N. Stavins. 2013. The SO2 Allowance Trading System: The Ironic History of a Grand Policy Experiment. *Journal of Economic Perspectives* 27 (1): 103-122.

Schmalensee, Richard, and Robert N. Stavins, 2017. Lessons learned from three decades of experience with cap and trade. *Review of Environmental Economics and Policy* 11 (1): 59-79.

Schultz, P. Wesley, Jessica M. Nolan, Robert B. Cialdini, Noah J. Goldstein, and Vladas Griskevicius. 2007. The constructive, destructive, and reconstructive power of social norms. *Psychological Science* 18 (5): 429-434.

Schultz, P. Wesley, Jessica M. Nolan, Robert B. Cialdini, Noah J. Goldstein, and Vladas Griskevicius. 2018. The constructive, destructive, and reconstructive power of social norms: Reprise. *Perspectives on Psychological Science* 13 (2): 249-254.

Scott, J. Taylor, K. Megan Collier, Jessica Pugel, et al. 2023. SciComm Optimizer for Policy Engagement: A randomized controlled trial of the SCOPE model on state legislators' research use in public discourse. *Implementation Science* 18: 12.

Shogren, Jason F., Jacob Hochard, Katherine D Lee, and Leticia Varelas Henderson. 2021. Experimental mindset for environmental challenges: The puzzling case of public good contributions. *European Review of Agricultural Economics* 48(4): 785–804

Shogren, Jason F., and Laura O. Taylor. 2008. On behavioral-environmental economics. *Review of Environmental Economics and Policy* 2 (1): 26-44.

Smith, Vernon L. 1994. Economics in the Laboratory. *Journal of Economic Perspectives*, 8 (1): 113-131.

Snowberg, Erik, and Leeat Yariv. 2021. Testing the Waters: Behavior across participant pools. *American Economic Review* 111 (2): 687-719.

Stoneham, Gary, Vivek Chaudhri, Arthur Ha, and Loris Strappazzon. 2003. Auctions for conservation contracts: An empirical examination of Victoria's BushTender trial. *Australian Journal of Agricultural and Resource Economics* 47 (4): 477-500.

Stranlund, John K. 2017. The Economics of Enforcing Emissions Markets. *Review of Environmental Economics and Policy* 11 (2): 227-246.

Sturm, Bodo, and Joachim Weimann. 2006. Experiments in environmental economics and some close relatives. *Journal of Economic Surveys* 20 (3): 419-457.

Tallis, Heather, Eli P. Fenichel, Laura Petes, Solomon Hsiang, Phillip S. Levin, Hila Levy, and Jane Lubchenco. 2024. Mainstreaming nature in US federal policy. *Science 385*(6708), 498-501.

Tanguay, Mark, Wiktor L. Adamowicz, and Peter Boxall. 1995. An Economic Evaluation of Woodland Caribou Conservation Programs in Northwestern Saskatchewan. Project Report 95-01.

Thaler, Richard H., and Cass R. Sunstein. 2008. *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT, US: Yale University Press.

US DOI (US Department of Interior). 1996. Record of Decision, Operation of Glen Canyon Dam, Final Environmental Impact Statement. October 1996.

US EPA (US Environmental Protection Agency). 2014. Benefits Analysis for the Final Section 316(b) Existing Facilities Rule. Report EPA-821-R-14-005. US Environmental Protection Agency, Office of Water, Washington, DC.

US EPA and DA (US Environmental Protection Agency and Department of the Army). 2020. Economic Analysis for the Navigable Waters Protection Rule: Definition of "Waters of the United States", January 22, 2020.

US GAO (US General Accounting Office). 1994. Allowance Trading Offers an Opportunity to Reduce Emissions at Less Cost. Report to the Chairman of the Environment, Energy, and Natural Resources Subcommittee, Committee on Government Operations, US House of Representatives, Washington, D.C.

US President. 2015. Executive Order 13707 of September 15, 2015, Using Behavioral Science Insights To Better Serve the American People. *Federal Register* 80 (181): 56365-56367.

US Securities and Exchange Commission, Office of Economic Analysis. 2007. Economic Analysis of the Short Sale Price Restrictions Under the Regulation SHO Pilot. https://www.sec.gov/news/studies/2007/regshopilot020607.pdf, accessed July 16, 2024.

van Oirschot, Piet. 2015. Handhaven op afvalscheiding Terugblik na 1 jaar, handreiking voor anderen. Municipality of Tilburg, BAT department, December 2015. https://vang-hha.nl/publish/pages/112042/2016-05handhavenopafvalscheidingintilburgpvo.pdf

Viscusi, W.K., W.A. Magat, and J. Huber. 1991. Pricing Environmental Health Risks: Survey Assessments of Risk-Risk and Risk-Dollar Trade-Offs for Chronic Bronchitis. *Journal of Environmental Economics and Management* 21 (1): 32-51.

Vollaard, Ben, and Daan van Soest. 2024. Punishment to promote prosocial behavior: A field experiment. *Journal of Environmental Economics and Management* 124: 102899.

WaterSmart. 2014. Tapping into the Power of Behavioral Science: Insights & Opportunities for Water-Use Efficiency. WaterSmart Software. https://www.watersmartsoftware.com/wp-content/uploads/2014/11/WSS_TappingintoBehaviorScience.pdf.

Welsh, Michael P., Richard C. Bishop, Marcia L. Phillips, and Robert M. Baumgartner. 1995. Glen Canyon Dam, Colorado River Storage Project, Arizona—Non-Use Values Final Study Summary Report. Madison, Wisconsin: Hagler Bailly Consulting, Inc., September 8. Springfield, Virginia: National Technical Information Service, NTIS No. PB98-106636.

Windle, Jill, John Rolfe, Juliana McCosker, and Andrea Lingard. 2009. A conservation auction for landscape linkage in the southern Desert Uplands, Queensland. *The Rangeland Journal* 31 (1): 127-135.

Zhao, Minjuan, Robert J. Johnston, and Eric T. Schultz. 2013. What to Value and How? Ecological Indicator Choices in Stated Preference Valuation. *Environmental Resource Economics* 56: 3-25.