Environment and Development Economics: page 1 of 38 © 2008 Cambridge University Press
doi:10.1017/S1355770X08004592

Experimental methods for environment and development economics

MARIAH D. EHMKE
Assistant Professor, Department of Agricultural and Applied Economics, University of Wyoming

JASON F. SHOGREN
Stroock Professor, Department of Economics and Finance, University of Wyoming, 1000 University Ave, Dept 3985, Laramie, WY 82071, USA and the King Carl XVI Gustaf Professor, Umeå University.
Email: jramses@uwyo.edu

ABSTRACT. Many poor countries remain trapped in a cycle of poverty and environmental degradation. Understanding how people react to existing and proposed solutions most likely can be improved using the methods of experimental economics. Experiments provide researchers a method to test theory, look for patterns of behavior, testbed economic institutions and incentives, and to educate people. Herein we explore how experimental economics has been used and could be used to help guide decision making to increase prosperity without overexploiting the resource base and environmental assets needed for basic survival.

1. Introduction

Why are people in some countries with abundant natural resources ensnared in both poverty and environmental degradation?[1] In theory, natural resources should generate economic rents through production and trade. These rents can then be invested in the physical and human capital needed to increase productivity and protect the environment, which in turn

Our thanks to the reviewers, Stephan Kroll, and other colleagues for their helpful comments. Ehmke thanks the following for their help with data collection: Tomas Nilsson, Cole Ehmke, Tony Torres, Glynn Tonsor, Sean Fox, Virginia Moxley, Carmel White, Min Wang, Jia Lv, Zuhui Huang, Tahirou Abdoulaye, Germain Ibro, Abdoulaye Mohammed, Moustpha Moussa, Bokar Moussa, Stephané Robin, and Bernard Ruffieux. The survey work was made possible through funding from the USDA-CSREES National Needs program, Purdue Research Foundation, Kappa Kappa Gamma Foundation, and the American Association of University Women. Shogren thanks Umeå University, the Norwegian University of Life Sciences, Swedish University of Agricultural Sciences (SLU), GREQAM, USDA-ERS, and NSF for their hospitality and support.

[1] See for instance the five volume report from the Millennium Ecosystem Assessment Board (2005), and the detailed synthesis and global/multiscale assessment reports online at www.millenniumassessment.org.

increases the net wealth of a nation and the well-being of its people (see, e.g., Fogel, 2004). In practice, however, resource-rich poor countries remain poor and polluted because these resource rents are inefficiently captured, imprudently spent, poorly reinvested, or wasted in rent-seeking conflicts (e.g., Pearce *et al.*, 1990; Barbier, 2005). A *resource curse* is said to exist, in which resource-based development fails to trigger broader economy-wide innovations due to several factors including overexploitation due to open access commons, weak institutions, initial factor endowments, frontier expansion on fragile lands, insecure property rights, and corruption.

How can societies design policies and incentives to encourage more development and prosperity while still protecting natural resources and the environment? Addressing this question requires insight into how people behave within institutions and natural environments. In response, economists offer theories and working rules about how people behave in markets and non-market settings. They provide rules to compare gains against losses under alternative policy scenarios and institutional arrangements, account for prosperity across time and space, show how trade can move resources from low-valued uses to high-valued uses, and consider whether policies balance incremental gains per cost across all policies actions (Hanley *et al.*, 2007). Traditionally, these rules have been framed within the economists' model of rational choice theory – rational behavior within alternative institutional and property rights structures (Dasgupta and Mäler, 1997).

But over the past three decades, many observers have questioned whether rational choice theory is the best guide for policy, both development and environmental. The traditional idea that people are 'poor but efficient, or at least rational' is being challenged by the notion that they are 'poor and bounded in their rationality' (Duflo, 2005). This argument holds with greater force for environmental and development issues that fall outside the market forces which act as catalysts for consistent choices (Smith, 2003). While most economists appreciate how rationality is a useful fiction that helps frame choices within exchange institutions, some observers believe we must recognize its limits for guiding environmental and development policy (Knetsch, 1997).

Relying on rational choice theory to guide policy makes more sense if people in developing countries make, or act as if they make, consistent and systematic choices toward certain and risky events. Anomalous behavior like overestimating low-probability high-severity events, preference reversals, endowment effects, loss aversion, context-dependent preferences over goods, risk, time and space undercut the rational underpinnings of environmental and development policy. But we also know these anomalies weaken or even disappear if people make their choices in a market setting (e.g., Shogren, 2006a). People become more rational when they make decisions within reach of other people or institutions that reward rational choices and punish irrational ones (e.g., Crocker *et al.*, 1998). These competing mindsets suggest the decision about whether rational choice theory is a reasonable guide for environmental and development policy cannot be made in the abstract. Rather researchers would benefit from a method that can be used to control and isolate

basic behavioral characteristics to test the predictive power of alternative analytical theories.

Experimental economics is one such method. Over two decades ago, Vernon Smith (1982) argued that experimental micro-economies are real economic systems richer in behavior than the systems parameterized in most abstract theories. He argued economics as a discipline will advance as a science once theorists become less own-literature oriented, take seriously the data and disciplinary function of experiments, and advance their own theories as potential generators of testable hypotheses (also see the pioneering work collected in Kahneman and Tversky, 2000). Experimental economics has re-introduced the role of institutions into the forefront of economic reasoning.[2] Today economists view the experimental method as another tool to isolate and test questions within environmental and natural resource economics (e.g., Shogren and Nowell, 1992; Sturm and Weimann, 2006; Cherry *et al.*, 2008).[3]

Directing Smith's line of reasoning toward environment and development economics (EDE), we also believe the challenges within this field can benefit from the experimental mindset.[4] This paper reviews the

[2] Experiments have also promoted the idea of the joint determination of institutions and choice. The three components underlie most experiments: the initial endowment of resources defining initial human and natural environments, the exchange institution or mechanism that allocates benefits and costs, and the actual behavior of people given their endowments and institutions. The environment includes endowments like preferences, natural resources, environmental conditions, technology, physical constraints, property rights, and information structure. The exchange institution specifies the rules to aggregate information and to coordinate actions, and it outlines the rules of exchange and its consequences. Repeated exposure to the institution is common practice in the lab so that people have the opportunity to gain experience with the institution, new information, and their own mind to better understand their endowments and preferences. People send messages (e.g., auction bid) to the institution. Based on the set of messages received, the institution allocates resources and costs based on the rules of exchange. Researchers then observe the choices people make, and how this behavior matches up with a specific performance criterion like demand revelation or efficiency.

[3] The experimental method has been used in environmental economics starting with the initial valuation work of pioneers like Peter Bohm in the 1970s and the Brookshire, Coursey, and Schulze team in the 1980s, but it has had a distinctly *developed nation* perspective (see Shogren, 2006b). Today researchers focus on institutional and property rights design, market-based incentive design, measuring preferences for non-market goods, and understanding elements of conflict and cooperation, primarily in places like the United States, Australia, and Europe. Experiments are either designed to address environmental policy questions in countries like the United States or Australia; or the citizens of these countries are subjects of experiments intended to effect developing country policy.

[4] As of this date, the journal of *Environment and Development Economics* has not published an experimental economic paper that involves decision making, given real economic commitments. Hypothetical surveys eliciting responses on non-market values, risk and time preferences are the closest application of the method. See for example Holden *et al.*'s (1998) survey on risk and time preferences in

role that the experimental method and mindset can play in EDE. Our goal herein is to examine how experimental methods can be expanded to EDE (also see Carpenter and Cardenas, 2006). We use an organizing framework on experimental mindsets to capture the key areas of on-going research, and how they relate to EDE applications.

Experimental methods can help guide policy by providing insight into how a proposed policy change in developing countries could affect behavior. By supplying information on the behavioral link between incentives, values, and choice, experiments can help us understand how policy might work. Working with controlled environments can improve our understanding of the underlying assumptions and incentives that drive behavioral responses to policy.[5] Experimental control complements the contextually rich social experiments, natural events, and field data traditionally used in development economics (e.g., Townsend, 1994).[6]

This work reveals the on-going challenge and tension of external validity – finding the right balance of experimental control and real-world social context. One way to understand this balancing act is to recognize that different people choose to frame the role of experiments differently. For many years, many economists agreed implicitly with Samuelson's textbook view that economics was not an experimental science. Experiments could not capture the workings of and decisions made within a modern economy. But over the last four decades, other economists reframed the debate such that 'simplicity' was in fact a virtue – researchers could control the economic environment such that new complexities could be added one by one into a design to see which complexities mattered and which did not (see Plott, 1989). Recently, some economists have revisited the idea that experimental *control* could be an illusion if people in a lab behave differently than they normally would because they are in the lab (e.g., Levitt and List, 2008, who argue field experiments reduce this risk; see Plott and Zeiler, 2007, who rebuff this argument).[7] Basically, this argument restates the classic

Zambia; Gebremedhin *et al.*'s (2003) survey on natural resource use in Ethiopia; Kirigia's (1998) Delphi approach to assess the risks of schistosomiasis control strategies in Kenya; Alpizar and Carlsson's (2003) choice *experiment* on travel mode choice; and Whittington's (2002) overview and examples on the application of the contingent valuation method in developing countries.

[5] See Kagel and Roth (1995) for an introduction into experimental economics; see Bohm (2003) and Shogren (2006b) for overviews on how experiments are used in environmental and resource economics.

[6] Thanks to a reviewer for stressing this point.

[7] Plott and Zeiler (2007: 1451) discuss the confounding factors that arise in a field setting. It is worth quoting their concerns at length: 'While these [field] experiments provide insights about theory robustness, they do not serve well as tests of competing theories. The problem is that observed asymmetries cannot be attributed convincingly to endowment effect theory, given the existence of competing alternative theories related to uncontrolled field variables. Testing endowment effect theory against classical preference theory in the field with all controls needed to make a convincing case, while at the same time maintaining an unperturbed field environment, would add complexities that likely would make it impossible to identify the theory at work due to the various extraneous

Hawthorne effect – the act of observing how a person behaves affects his or her behavior.[8] Others have countered that this argument is a straw man (see Roth, 2006). One should view the experimental method for what it is: a tool to understand economic phenomena that works alongside theory and econometrics. But this healthy debate illustrates how far the experimental method has come over the years in economics. Rather than reject the tool out of hand, today the *experimental mindset* exists and persists in modern economics, even if different researchers choose to frame its validity differently. Experiments do not create a 'people vacuum' – in the lab or the field. Economic theory should still be robust enough to capture the fundamental human behaviour it purports to explain across a variety of contexts within and outside active exchange institutions. That said, generalizing context-driven findings beyond the context they were created does not serve the method or the mindset well (see e.g., a context-driven bargaining experiment in Shogren, 1989; the handy overview in Harrison and List, 2004). The experimental mindset should work for researchers, not the other way around.

2. Experiments as a mindset

We divide the experimental mindset into three broad and overlapping areas: market, no market/social, and missing market/environmental experiments (see figure 1). To the extent the areas are delineated, a traditional division of research labor exists as well. While some blur exists, the economists' mindset has focused on market-centered experiments, psychologists and behavioral economists have explored no-market or social experiments, and environmental/resource economists have examined non-market or missing market experiments. We now consider the traditional role within each category, and then concentrate on the overlap between categories and the importance to environmental applications in developing countries.

2.1. Individual categories of experiments

We do not provide an extensive overview of the individual categories here because other surveys exist that serve this purpose (see for example Kagel and Roth, 1995). Market-centered experiments frequently examine

forces triggered by entitlement creation. Fortunately, laboratory experimental procedures can be structured to test the predictions of competing theories more easily, given the ability to control the environment. By peeling away the previously unrecognized complexities and using several subtle variations of controls and replications that would be very difficult if not impossible to implement in the field, we are able to identify the theory that better explains observed exchange asymmetries.'

[8] See Bardsley (2007) for example, who discusses the Hawthorne effect in his innovative experiment on inconsistent behaviour in dictator games. Also see for instance Karlan (2005), who is up-front about the Hawthorne effect when discussing his experimental design. He notes 'much care had to be taken to ensure that participants understood the game. The transactions for both parties were done face to face (and privately) with the game administrator. This risked that our presence influenced their decision, but provided us the opportunity to confirm that each individual understood the rules.'

Figure 1. *Experimental mindsets.*

how people behave in the context of an active exchange institution like the market with well-defined property rights (Smith, 1982). The market allows a rational person to exploit non-rational choices, which can help reinforce rational decision-making throughout the population, depending on the degree of arbitrage. Experiments have focused on the properties of general equilibrium, multiple markets, industrial economics, incentive design, growth and trade, auction design, and asset markets (Davis and Holt, 1993).

Psychological/behavioral experiments focus on individual behavior, in which decisions are made outside the realm of exchange institutions (see Kahneman and Tversky, 2000; Levine and Norenzayan, 1999). These isolated choices are compared to the benchmark rational choice theory, which not surprisingly is an imperfect predictor of actual behavior given the existence of social preferences (see e.g., Smith, 2003). Topics of interest include entitlement perception, heuristic based decision-making, anomalous behaviors and decision biases, belief formation, preference reversal, time preferences, risk avoidance/acceptance, endowment effects, fairness expectations, and framing effects.

Environmental/resource economic experiments confront the idea that, while markets do exist in general, they are absent for the specific good or service in question (e.g., air pollution, biodiversity protection). Here experiments are designed to evaluate the rules that define new exchange institutions and incentive systems like tradable emission permits; efficiency impacts of property rights security, resource depletion, and access to resources; conventions creating conflicts and cooperation over resources, e.g. Coase theorem, and valuing the demand for a public good like risk reduction.

2.2. Overlapping categories

No perfect separation exists between categories. Overlap arises because people make their decisions within market and non-market contexts simultaneously, and this overlap of exchange institutions can affect their decisions. This overlap requires the researcher interested in environment and development economics to explore the insight within each experimental literature because they are intertwined. This points to the power of the experimental method – the method allows one to focus on and isolate key parameters and incentives created by alternative economic and cultural institutions and laws of nature. Consider now examples of four research areas created by the overlapping categories, and how they relate to environment and development economics.

Market-behavioral overlap

This is the most active area of experimental overlap in developed countries, yet lacks application in developing contexts. In developed country literature, traditional neo-classical economics has collided with recent developments in so-called behavioral economics.[9] The debate focuses on the degree to which economists should account for social preferences, cultural, norms, and individual traits in market and non-market transactions. Smith (2003, p. 466) expresses the sentiment well: '[m]arkets economize on the need for virtue, but do not eliminate it'. These traits are embedded in and shaped by the structure and political economy of the marketplace (see Bowles, 1998). The need for market transactions depends on the efficiency of social and cultural norms to facilitate cooperation and endowment allocations across different parties. Market transactions are less vital if efficient cultural and social rules dominate resource allocation, which can be the case in developed and developing countries (Mullainathan, 2004).

The intersection of social preference measurement and market economics raises issues related to Smith's (2003) notion of *ecological rationality* and Bowles's (1998) argument that social preferences are shaped by the institutional transactions. While many individual decision-making outcomes may appear irrational or inconsistent in isolation, successful markets bring together decisions so that rational decisions dominate in aggregate (Camerer *et al.*, 1989). The market provides feedback to its participants, and it defines what behavior evolves as rational and utility optimizing. How markets shape behavior depends on the ability to decrease the transaction costs of social exchanges or make visible the opportunity costs of irrational decisions.

Despite these observations, there is little application of market-behavioral overlap experiments with regard to EDE issues. This gap implies many unexplored opportunities exist to implement such experiments in developing country contexts. Experiments can be a useful tool to define which transactions require greater market intervention and which ones are regulated through the adaptive symbols created by existing social

[9] See for example Mullainathan and Thaler (2000) who survey the state of behavioral economics; also see Shogren and Taylor (2008) who examine its intersection with environmental economics.

preference mechanisms (see Alvard, 2004). Currently, conclusions that can be drawn from the experimental economics literature lay the groundwork to clarify the role that preferences play in market interactions and vice versa. For example, on a cross-cultural context level, research shows greater market integration in the real world can lead to less rational behavior. Henrich *et al.* (2001) find behavior in the ultimatum bargaining game moves further away from the Nash prediction (i.e., little given by the proposer and little accepted by the respondent) in locations closer to markets.

But market formation around otherwise non-market situations can also lead to more rational and efficient public good resource allocations. Experiments resulting in such outcomes have been conducted mainly in developed countries and show how market experience and market structure shape subject behavior. Experienced people more familiar with certain market procedures and structure behave differently from the inexperienced. A specific example related to Henrich *et al.* (2001) is found in Tracer (2004). He examines the question of market integration, reciprocity, and fairness in rural Papua New Guinea. He uses the now classic ultimatum game to explore whether people with more integration with markets behave more as rational choice theory predicts. The ultimatum game is an experiment in which one person offers to split some resources with another person. If he accepts, they both receive the offered split; if he rejects the offer, both receive nothing. Theory predicts the person will accept any positive offer, such that the person will offer up a 99 per cent (proposer) to 1 per cent (respondent) split. Experimental evidence has not been kind to theory, however, as people usually reject such low offers. Most offers are closer to a 60–40 per cent split. Tracer runs the ultimatum game in two villages – Anguganak and Bogasip – which differ in market integration through cash cropping, education, and acculturation. His results suggest a positive relationship exists between the amount offered and market integration – more contact with the market as in Anguganuk, the less behavior goes as rational choice theory predicts. Rather the people in the more isolated village Bogasip made lower offers more in line with *Homo economicus* (also see Henrich, 2000).

Missing market–market overlap

The missing market–market overlap takes a strong rational choice perspective. Following Arrow's notion that market failure can be defined as a problem of missing markets, experiments study the evolution of the rational solution to pollution. Many classic environmental economic and market experiments studies illustrate the overlap between missing market and market experiments. The lab provides an environment wherein economists can test economic and market incentives to control environmental production (also see Bohm, 2003). Economists have long promoted control systems as cost-effective alternatives to technological restrictions and other forms of inflexible command-and-control environmental regulations. Economic incentives through prices or quantity rationing increase cost of environmental shirking, and provide more flexibility to find the least-cost pollution control strategy.

Economists promote market-based mechanisms like tradable permits as cost-effective methods to control pollution (e.g., Crocker 1966). Here the

experimental challenge is *economic design* – how can one set the rules of a pollution trading market to hit a given emissions target at least cost. Plott (1983) used the first experiments to test the efficiency of economic-based incentive systems relative to the more traditional regulatory approach of command-and-control. Given the wedge between the private and social optima, Plott considered three corrective policies: taxes, standards, and tradable permits. Theory also predicts that economic price-based incentives like Pigovian taxation and tradable permits are usually more cost-effective than command-and-control systems.

Plott designed a competitive market of buyers and sellers who trade a valuable good. The institutional trading mechanism was a double-oral auction in which buyers make bids that increase in value, and sellers call out offers that decrease in cost. The externality was constructed by reducing a seller's profits from trades as a function of the total number of trades in the market. Plott first established that the competitive market with an externality converges to the private optimum; traders ignored social costs. The results suggest taxes and tradable permits could work to equate private incentives with social costs, and they were significantly more efficient than the command and control approach. Plott concluded the standard models are 'amazingly accurate' in the lab setting.

Another example is Cason's (1995) lab evaluation of the US Environmental Protection Agency's (EPA) sulfur dioxide (SO_2) emission trading program. He designed an experimental market to capture the design of the initial SO_2 auction: buyers and sellers submit bids and offers for emission permits given their induced preferences; and the intersection of the implied supply and demand curves determines the quantity of permits traded. The EPA set the market price discriminatively off the demand curve. Their scheme matches the buyer with the highest bid to the seller with the lowest offer; the final price is the bid price of the buyer. The matching continues, the second highest bidder to second lowest offer and so on, until the equilibrium quantity is reached. Using an inverted scenario (matching the highest bidder with the lowest offer and having her pay the offer price), Cason showed that such an institutional design provides buyers incentives to overstate their willingness-to-pay. Likewise, in the actual EPA market, sellers have an incentive to understate their willingness-to-accept – a lower value will increase the chance they are matched with a high bidder to earn a higher price for their permits. The results of this experiment suggest incentive mechanisms depend on the rules defining the institution, and changing the trading rules would make the SO_2 markets more efficient (see Cason *et al.* (2003) on the design of mechanisms for non-point source pollution).

Another example of the market-missing market overlap is the experimental literature on the Coase theorem. Recall the theorem says if a regulator can implement an efficient Pigovian tax, transaction costs must be low. If costs are this low, the disputing parties could just as easily bargain to an equally efficient outcome, regardless of who is assigned property rights over the resource (Coase, 1960). Here the basic lab design considers two parties in each bargain, the demand and supply for pollution control or pollution emissions depending on which party has property rights, perfect

knowledge of payoffs, zero transaction costs, perfect contract enforcement, and no wealth effects. These assumptions create two testable outcomes: two parties will bargain to an efficient and self-interested agreement. The early experimental tests found efficient but equitable bargaining (Hoffman and Spitzer, 1982). Coase experiments now examine the key issue on how different types of economic friction affected bargaining behavior. Evidence suggests bargaining efficiency was robust to large groups and uncertain payoffs, but falls with imperfectly enforced contracts, delay costs, tournament incentives, risk of third party intervention, and insecure property rights (e.g., Rhoads and Shogren, 2003).

Behavioral and missing-market overlap

The further one moves from markets, the less acceptable is the idea of rationality and rational choice. One expects locally defined social norms now matter more for individual behavior within the group than global ideas of market efficiency based on price transactions. This is potentially the richest area for new research within the field of EDE. The intersection between social preferences and non-market allocation is at least as productive as that between social preferences and market economics.

Evaluating behavior in common property institutions and towards public goods contributions is probably the most active experimental research related to EDE issues. One example of an experimental evaluation of common property resources is Walker and Gardner's (1992) conflict game. They examine how people might exploit the rents of a common resource without over-exploitation. Their experimental design captures the idea that a natural regeneration process exists to create a range of exploitation in which the probability of destruction is zero. When one goes beyond the 'safe yield,' the resource now faces probabilistic destruction. High levels of economic activity would destroy the resource with certainty; low levels would not extract the optimal level of rent. The key is to find the balance between maximize economic gain and preserving the resources. Framing their experiments as non-cooperative games with multiple Nash equilibria, they create two treatments: the safe zone has a single point or an interval. They find when the safe zone is a single point, the resource was rapidly destroyed. When the zone was an interval, groups found but could not sustain the social optimum (see Ostrom *et al.*, 1994). In contrast, Mason and Phillips (1997) observed increased common pool cooperation within a group given they had repeated interactions with each other – they knew each other and had established a reputation and trust. The expectation for the Nash equilibrium is low or unsustainable cooperation and it is sensitive to the length of the game and other dynamics like entry/exit of players. Cooperation was undercut, however, when more people (firms) were allowed to enter the common pool.

Public good provision is another active research area within this intersection. The traditional economic view on public goods is that rational people have the incentive to free-ride. Economic experiments have been used extensively to test this prediction. The most common form of public

goods experiment is the Voluntary Contribution Mechanism (VCM).[10] The expectation is for each person to contribute nothing to the public account relative to the social optimum in which everyone contributes their endowment to the public account to maximize group surplus. Experimental findings show people contribute an average of 40 to 60 per cent of their endowments to the group account (Ledyard, 1995).

The actual level of contributions, however, depends on the groups making the contribution and use of peer enforcement in the group. First, we discuss the role of group characteristics, then peer enforcement. Contributions given different social and cultural characteristics, especially in developing country environments, are sensitive to the ethnic and gender composition of subject groups (e.g., Barr and Kinsey, 2002). In her experiments in a Nairobi slum, for instance, Greig (2005) finds gender plays an important role in signaling trust and cooperation among people. Men in mixed gender groups expect women to be much more forthcoming with public investment funds than they are. Women decrease their public investment in mixed gender groups (compared to all female groups) due to distrust of men. Another example is Gurven (2004); he found high voluntary contributions levels in public good games among the people living in five Tsimane' villages in the Amazonian portion of Bolivia.

Additional laboratory-based literature has emerged around the use of peer enforcement in the VCM. Ostrom *et al.* (1992) originally introduced peer sanctioning as a means to increase cooperation in the Common Pool Resource game; Fehr and Gächter (2000) introduced sanctioning into the VCM. Group members are allowed to use monetary fines to sanction or punish those people who contribute less than the group average contribution to the VCM. The fines do incur a cost for both the punished and the punisher (i.e., the punisher pays the experimenter a fee to impose a fine on the punished). They considered two experimental treatments: people are either randomly paired with an anonymous partner each round, or are paired with a continuous partner from round to round. Group behavior evolves toward Pareto optimum equilibrium in the partner treatment and cooperation is the most frequent action.

Masclet *et al.* (2003) show similar outcomes can be achieved using non-monetary enforcement. By allowing students to express disapproval with each other's contributions through a points system, which did not affect individual's earnings, they find that, like monetary punishment, non-monetary punishment results in higher VCM contributions than no-punishment treatments. Both Masclet *et al.* and Fehr and Gächter find sanctions increase, either monetary or non-monetary, with a punished subject's negative deviation from average group giving. Occasionally, people are also punished for giving 'too much' relative to the group

[10] In one VCM design, people grouped into two to four are given an endowment which they either keep or invest into a public account. The experimenter collects the contributions to the group account in an envelope, doubles them, and then equally divides the outcome amongst the group regardless of whether they contributed to the public account. The game theoretic prediction of the VCM is classic free-riding – no one will contribute anything to the collective.

average. Carpenter (2007) develops a typology of punishment behavior: Hypocritical free-riders, principled cooperators, and principled free-riders. The principled cooperators are most likely to punish despite the cost of doing so to their own personal income.

The past application of these experiments in a field context for EDE issues is limited. The VCM experiment using peer enforcement, however, has many possible applications including social, educational, and environmental problems. Past work has also discussed the role of peer enforcement. Miguel and Gugerty (2005), for instance, examine social sanctioning effects in public good provision in Western Kenya. They investigate non-monetary sanctioning within and across ethnic groups for two public goods: education and water wells. They find social sanctions produce less free-riding within than across ethnic groups. Their investigation uses extensive secondary field data.

Voting schemes with third party enforcement have also been proposed as a mechanism to increase cooperation (see Walker *et al.*, 2000). But in some contexts related to development, such third party enforcement schemes are unavailable or too costly. The voting mechanism is non-binding, unless people can institute costly self-enforcement. The challenge is that punishment is a public good in its own right. In theory, each person should want someone else to spend his or her private resources to punish the defector, which benefits the entire group. But Kroll *et al.* (2007) find that such within-group punishment can work to improve the performance in non-binding voting mechanisms. When punishment was available, they observed relatively more cooperation in an otherwise inefficient voting environment.

An additional approach to understanding the complexity of the problem and possible policy to decrease free-riding in ethnically diverse communities is to implement the VCM experiment with economic peer sanctioning. Abbink *et al.* (2006) look at free-riding of loan repayment in microfinance institutions using an economic experiment. Their subjects were German university students. They test the effects of increasing the repayment burden and loan interest rates to decrease free-riding within borrower groups. One adaptation of this experiment is to test the effect of peer sanctions of free-riders in the micro-lending environment. Such an experiment could be extended further through field implementation in relevant and randomly selected cultural contexts rather than with developed country undergraduate students.

We next consider research on social capital, social networks, and trust. Traditional economic analysis assumes resources are organized around the principal of profit maximization with individual ownership of capital. The primary focus on profit maximization, however, overlooks many non-profit objectives involved in resource allocation and social preferences beyond greed. Recent growth in the social capital literature identifies social networks and trust as key components of successful economic exchange systems (e.g., Knack and Keefer, 1997; Barr, 2002, 2003). The importance of these social capital components has been accentuated in economies making rapid switches from communism to planned capitalism (e.g., the former Soviet Union and Eastern Europe). If the appropriate behavioral

and institutional conditions are not present or do not evolve in conjunction with market reform, the evolution of successful market operations may be thwarted. The unforeseen corruption (especially organized crime) within the new systems halts the anticipated growth from market reform (Sen, 1999).

A strict focus on profit maximization and own-utility maximization does not always suffice to explain individual's behavior in non-market or public good settings. In addition, market solutions focusing on greedy behavior alone might not suffice. This holds for the case of non-point source pollution and mechanism design. Although market incentives may be initiated to encourage pollution reduction, multi-dimensional principal-agent problems complicate implementation and increase transaction costs (see Horan *et al.*, 2002). In addition, in low-resource and alternative institution contexts, it is important to understand how social networks and trust promote (or discourage) economic growth by reducing uncertainty and revealing market information (Barr, 2002).

Main components of social capital from the literature that are considered for this paper include trust and social networks. Social network analysis includes identifying and understanding the nature and strength of relationship links between individuals and groups of individuals. The importance of society with regard to individual behavior is part of a nearly age-old argument within the social sciences about the effect of societal structure on individual behavior. Social preferences omission from economic analysis often leads to criticism from other social scientists that the assumption of individual rationality lacks relational context (Degenne and Forsé, 1999). Social networks may have many unintended consequences on individual behavior. But as Hayek (1967) noted, an unintended consequence is not necessarily unpredictable (Sen, 1999). As economists strive to refine the predictability of an economic policy alternative for environmental and development policy, it is important to consider the direction and predictive power of social network measurements. This holds for environmental and development policy that may be developed and implemented across various social networks.

Trust, social capital, and social networks are intertwined in natural resource use and governance. When one person trusts another, they are more likely to put a resource in the care of this other person. How much people trust each other depends on their expectations about the trustworthiness of another party. As one party's trustworthiness expectations for another party decrease, both exchanges and transfers become riskier. Trust matters for the successful operation of an exchange economy, and it helps overcome principal-agent problems in missing markets.

Ensminger (2001) uses ethnographic methods to document the value of social networks and trust, focusing on the Orma society in Kenya. One challenge faced by Orma is common to nearly all agrarian land and capital owners in the world. Without constant monitoring, workers may act in ways that harm current production or the long-term value. Orma cattle herd owners traditionally rely on their sons to travel with hired herdsman and watch over their stock. One of the sons' primary roles is to monitor workers

and reduce moral hazard concerns. In recent history, however, the Orma have become more sedentary. Sons in wealthy, herd-owning families go to school rather than herd stock, leaving their families to rely on non-family members to care for their stock in distant grazing lands. These changes have created principal-agent problems as the stock caretakers have less ownership in the stock and more motivation for opportunistic behavior. To deal with the new monitoring challenges, wealthy Orma families went from a traditional, balanced reciprocity relationship (e.g., one party gives aid or is generous to another, but expects equivalent aid or generosity in return) to generalized reciprocity (e.g., unidirectional transfers of housing and food from herd-owner to herdsman and herdsman family) through fictive adoptions of the herdsmen into the wealthy household's family. Herders that perform well are rewarded with deepened integration into the herd-owner's family. These kinsmen-like arrangements create stronger bonds and reduce monitoring costs for the herd-owners, providing adequate capital conservation.[11]

Few studies consider the role of social networking in natural resource use or allocation in developing nations. There are, however, studies that do find social networking to be a significant economic growth predictor in developing country context. Barr (2002) investigates the role of social networking in social capital spillovers within the Ghanian manufacturing sector. The study addresses the role social networking plays in an environment defined by missing or dysfunctional markets for credit and insurance and a lack of clear property rights. She finds that small firms with less access to formal markets for credit and insurance rely more heavily on social networks for market information and uncertainty reduction. Smaller firms engaged in more networking activity and put greater reliance on the network in difficult times. This study shows how the value of social capital derived from networks may be valuable in low resource economies relative to richer economies.

Experimental economic methods offer opportunities to measure trust within groups and its resulting influence on social networks within societies. An original and oft-replicated trust experiment comes from Berg *et al.* (1995). They developed an experimental mechanism which extends fairness and generosity experiments to measure trusting and trustworthiness among individuals within groups. In their original experiment, people were directed into two separate rooms, rooms A and B. All subjects were paid a $10 show-up fee. The people in room A were asked to decide how much of their $10 show-up fee they would send people in room B. Meanwhile, the room B people were allowed to pocket their show-up fee. The amount sent by a room A subject to room B was tripled before it reached a room B subject. After a subject in room B received the money from his/her anonymous partner in room A, he or she then decided

[11] Ensminger (2004) returned to the Orma village and examined societal levels of reciprocity using voluntary contribution, ultimatum, and dictator games. She found their reciprocity and altruism to be greater than levels found in other developing countries, but less than that found in US studies, which are primarily run with students.

how much of the money to return to room A. The room A subject was the 'trustor'; the room B subject was the 'trustee'. Trust was defined by the trustor giving the trustee an opportunity to make a decision with the money. The trustworthiness of the trustee was then measured by his/her response, which affected the outcomes for both the trustor and trustee.

Berg *et al.* test the null hypothesis that people make decisions consistent with the subgame equilibrium and give each other nothing (beginning with a test of the hypothesis only for room A people). They test the effect of social history on subject behavior via two different treatments. In the first one, people are given no transfer history on their partner or anyone else in the experiment. In the second treatment, people receive summary information about the transfer behavior of people in the first experiment. Results reject the null hypotheses that trustors will send nothing. Although some do send nothing or $1, nearly one half send more than $1. The amount of money returned to the trustee is positively correlated with the trustor's initial offer. More knowledge about each other increases the amount and frequency of money trustees send trustors and the amount of money trustors send trustees.

Barr (2003) replicated the trust experiment in a developing country context. She conducted the experiments across 24 communities in Zimbabwe. Eighteen of these communities were non-traditional, resettled communities with fewer internal kinship connections. The remaining six communities were traditional, long-term communities with deep family and kinship relationships. The experiments allowed her to examine how pure altruism, in the case of the resettled communities, and genetic closeness, in the case of the traditional communities, motivate trusting behavior. She found greater variance in trustworthiness in resettled communities. Due to their general aversion to risk, trusting behavior fell in resettled communities relative to traditional communities. A lower trustor giving rate in the resettled communities reflects less trusting based on pure altruism compared to genetic closeness in the traditional communities. This is despite strong preferences for and activities related to community building in the resettled communities. While Barr does not directly relate her findings to an applied policy problem, they hold important implications for environmental and development policy. In the Zimbabwean case, trust differences may imply that group incentives to contribute to public goods may need to be adjusted across villages which have resettled instead of traditional populations.

But before taking this policy advice, one must first address the open question of whether the trust game behavior is driven by trust or by the propensity to gamble, or both. Karlan (2005), for instance, ran the trust game in Peru with members of a non-profit microfinance organization. He used behavior in the trust game to predict a person's real finance decision making, i.e. default on a microfinance loan and voluntary savings. He found that the behavior of people previously assumed to be 'trusting' was better explained by risk preferences – in the real world, these people defaulted more and saved less, which suggests they were *gamblers*. In a related experiment run in Paraguay, Schechter (2007) also found risk preferences better organized data in the trust game. She designed a treatment aimed at disentangling

risk preferences from trust in the trust game. She used a two step design: (1) each person was given some money (8,000 Guaranies) and asked how much he or she would choose to bet (0, 2,000, 4,000, 6,000, or 8,000) in a specific gamble in which the roll of a 6-sided die determined the pay back (0, 0.5, 1, 1.5, 2, 2.5 times his bet); and (2) he or she played the standard trust game. The ordering of the design, however, could have created its own confound (gambling first could lead to extra gambling in the trust game). More understanding on how context affects trust, risk, and social networks may enable us to give better policy advice in the future (see, e.g., Johansson-Stenman *et al.*, 2008).

Another critical line of experimental work is the elicitation of preferences – for goods and services, for risk, and for time. An accepted practice is to measure *ex ante* benefits of public policy and *ex post* economic losses from environmental damage using the set of stated preference methods, e.g., contingent valuation method (CVM). A complicating factor with this method is hypothetical bias or preference reversal, e.g., a person states a hypothetical willingness-to-pay (WTP) for a project, but then changes his or her WTP for a real economic commitment.

Experimental evidence from Ehmke *et al.* (forthcoming) suggests this may not be sufficient in developmental contexts. Ehmke *et al.* use economic experiment techniques to determine the degree to which hypothetical bias is consistent across geographically and culturally isolated groups. In experiments conducted in Indiana, Kansas, Niger, China, and France, they find greater heterogeneity in hypothetical bias in non-Western cultures. While American and French subjects frequently overstate their preferences, people in Niger are more likely to understate their preferences (or vote 'No' in a hypothetical referendum and then 'Yes' in the real referendum). Meanwhile, people in China are most likely to not change their vote at all across hypothetical and real referendums. Through the use of Hofstede's Value Survey Module and econometric modeling, they conclude that people in more individualistic cultures display a greater tendency toward preference reversal or hypothetical bias (either an overstatement or understatement of preferences) in a dichotomous choice mechanism.[12] This work supports the view that cross-cultural valuation work should address differences in collective versus individualistic cultural orientation for calibration and benefit transfer exercises.

The economic implications of such differences can be measured by applying Ehmke *et al.*'s hypothetical bias measures to an environmental policy problem: global warming. In the original experiment, they asked people in China, France, Niger, and the United States to vote in a hypothetical and then real referendum to purchase water. An overview of the referendum results, presented in table 1, reveal that people in different cultures do not change their votes from hypothetical to real referendums to varying degrees across locations. People in the US and France tend to overstate their WTP in the hypothetical referendum while those in

[12] Hofstede (1980, 1991) identified five dimensions by which culture may vary: power distance, individualism, masculinity, uncertainty avoidance, and long-term orientation, which are measured by the Value Survey Module.

Table 1. *Experimental referendum voting results*

Location	Number of people	Experiment referendum		
		Hypothetical	Real	Change
		Percent 'Yes' votes		
China	96	77	65	−12
France	70	37	30	−7
Indiana	63	75	25	−50
Kansas	57	61	39	−22
Niger	60	30	55	25

Source: Ehmke *et al.* (forthcoming).

China and, especially, Niger frequently understate their preferences in the hypothetical referendum.

In addition to this experiment, Ehmke *et al.* administered a survey to the same group of people, eliciting their WTP to support policy to reduce global warming by paying an annual $US 20 tax (or its purchasing power parity equivalent in non-United States locations). Appendix A presents the US version (English) of the survey referendum. In the hypothetical survey referendum in each location between 70 to 89 per cent voted for paying $20 annually to support global warming prevention. The expected WTP for the policy can be calculated using the Turnbull estimate based on the hypothetical survey referendum results. Using the hypothetical referendum results, WTP estimates range from $13.97 to $17.71 per year.[13] The WTP measures are then adjusted using the hypothetical bias measures from table 2. In the experiment referendum, Indiana subjects decreased their support for the referendum by 50 per cent from the hypothetical to the real referendum. If they behave consistently in the survey referendum, then their WTP would decrease by $10 or from $13.97 to $3.97. The Nigerian subjects, in contrast, would actually pay more in the end, increasing their real payment to $20.33, $5.00 above their hypothetical statement of value. Culture influences entitlement expectations, and affects preference expression, especially between hypothetical and real valuation situations.

Time and risk behavior measurements matter in environmental policy and development, especially for cost–benefit analysis (e.g., Hanley and Spash, 1993). Economic theory assumes both time preference and risk aversion are functions of wealth and little else. Using Samuelson's discounted utility theory, time preference is largely a function of present wealth and future earning expectations. Expected utility theory assumes rational people equally weight expected outcome probabilities when

[13] The Turnbull estimate of the lower bound of willingness to pay for water, obtained by multiplying the percentage of subjects voting *Yes* to the vote and the price of the water, in Indiana (Kansas) changes from $0.23 in the hypothetical referendum to $0.08 in the real referendum ($0.18 to $0.12) per 24 oz. bottle of water (Haab and McConnell, 2002).

Table 2. *Hypothetical survey referendum voting results and hypothetical and calibrated (WTP) to support global warming policy*

			$US		
Location	*Number of people*	*Voting 'Yes' for global warming policy (%)*	*WTP based on hypothetical global warming vote*	*Calibrated WTP for global warming policy*	*Difference*
China	96	89	17.71	15.31	−2.40
France	70	89	17.71	16.31	−1.40
Indiana	63	70	13.97	3.97	−10.00
Kansas	57	70	14.04	9.64	−4.40
Niger	60	77	15.33	20.33	5.00

making decisions. These presumptions occur, however, despite arguments to the contrary by economists and non-economists who point out all the other individual characteristics and social experiences that influence behavior toward risk and time. While using wealth as the primary driver of behavior can be useful in some contexts, it becomes more problematic in development contexts.

Consider first time preferences. Over a century ago, John Rae discussed the influence of time preference on savings and consumption behavior. He discussed the role socialization plays in forming time preference, which was excluded in early discounted utility theory (Maital and Maital, 1977). Across the social sciences, especially psychology, societal context matters in how a person forms his or her time preference. Psychologists find that behavioral differences in time preferences for non-monetary gratification arise as a part of the human development process. One's ability to delay gratification depends on role modeling observed from family members and other social role models. Like strategic ultimatum bargaining behavior, economic experiments have shown that each type of behavior is much more complex than theory predicts. Understanding the role culture and individual characteristics play in time value of money matters because, as Adam Smith pointed out, societal differences affect how different societies accumulate capital (Frederick *et al.*, 2002).

Isolating a person's decisions based on time preference, however, is a challenge. Some experiments reinforce the notion that individual characteristics beyond income matter. For Denmark, Harrison *et al.* (2002) find several individual characteristics beyond income influence individuals' time preference. They find that education, life stage (e.g., retirement), unemployment status, and one's credit rating are significant characteristics affecting time preference. The open question is whether these results can be replicated across cultures due to the sensitivity of results to payment amounts and the institutional framework. If one is doing cross-cultural experiments, translating monetary endowments across cultures to maintain purchasing power parity can be a challenge. In addition, different cultures

have different preferences for the institutions design, e.g., all male versus all female investment group.

Regarding risk preference, a classic is Binswanger's (1980) experiment designed to measure Indian farmers' risk aversion coefficients. The Binswanger study shows that people display higher levels of risk aversion as the maximum payoff levels increase. In the study, the relationship between risk aversion and wealth is tested in regression analysis. The wealth measurement coefficient associated with the gross value of physical assets has a negative coefficient, but it is not significant. If significant, wealthier people were more likely to gamble. But this study was unable to find a significant relationship between wealth and risk aversion (also see Eggert and Martinsson, 2004; Weber and Hsee, 1998; and Holt and Laury, 2002).

In addition, experiment results also show that a person's time and risk preferences are subject to contextual and framing effects. For example, experiments that ask people to compare delayed rewards over varying time lines find that people have a declining rate of time preference or people display hyperbolic discounting (Frederick *et al.*, 2002). Kahneman and Tversky (2000) show that choice framing has a significant effect on the degree of risk aversion people display. People are more risk averse over choices involving gains and risk seeking over those involving losses. Applying this knowledge to developmental contexts may be useful to explain anomalies to expected utility model outcomes. Understanding how framing affects people can be useful to help better understand behavioral choices involving environmental losses versus gains.

Market, behavior, and missing market experimental overlap

The overlap between all three areas is potentially the most important and least understood and studied topic of research. The overlap of these areas creates a dynamic nexus of incentives. Property rights, social preferences, and exchange institution definition bind together the three areas. The power of the experimental method to explore the interaction is strong when moving across economic and cultural/societal contexts. We now briefly consider two examples – rationality spillovers and corruption.

First, we explore the idea of a *rationality spillover*. In EDE problems, resource and environmental assets can lack well-defined exchange institutions, e.g. open access. But people make choices every day in some mixture of market and non-existent institutions, which can serve to blur the distinction between private and social preferences. The question is whether the consistency created by arbitrage in a market will spill over to choices in non-market settings. If markets make people more consistent in one context, do they become more consistent in another non-market context?

Cherry *et al.* (2003) address this question for the case of preference reversals. Recall, a preference reversal occurs when a person says he prefers lottery A to B, but assigns a greater monetary value to B than A. Their experimental design simulates two simultaneous but independent settings – a market with arbitrage and a non-market setting. Treatments they considered are: (1) real versus hypothetical choices, (2) arbitrage versus no-arbitrage decisions, and (3) monetary versus environmental payoffs, e.g., chances to see a grizzly bear in a park. With arbitrage, people who reversed

their preferences were arbitraged in three steps. The market sold its least preferred/highest valued lottery (say lottery B for $12); next it traded its most preferred for the least preferred lottery (trade B for A); and, finally, it bought back the most preferred/lowest valued lottery (A for $5). The subject owned neither lottery, and was $7 poorer (the difference between the values for the two lotteries, $-\$7 = \$5 - \$12$).

Their results found that relative to the no-arbitrage treatment of 25–30 per cent preference reversals, once arbitrage was introduced, preference reversals decline significantly in all treatments. More importantly, their findings suggest a *rationality spillover* did exist. Preference reversals declined in the non-market setting – even though no direct arbitrage took place, the non-market choices were hypothetical, and the lotteries became environmental experiences in a national park. People transferred what they learned in the market setting, and applied it in the environmental setting. The rationality spillover results provide some behavioral support for models assuming rational choice to motivate EDE policy. In addition, the rationality spillover result raises the intriguing question about whether there are preference spillovers, institutional spillovers, equity spillovers, or rationality spillovers with and between other cultures and contexts more related to developing country issues. This in turn raises the question of whether we should be trying to explain biases in behavior versus developing active exchange institutions that induce rationality spillovers which tempers these biases. The rationality spillover results suggest more effort spent on developing parallel market and non-market exchange institutions that generate the rationality spillovers might be interesting.

A second set of experiments we consider are on corruption. Corruption is a key challenge to environment and development policy in developing countries. Otherwise successful policies can easily be thwarted by corrupt behavior. Corruption exists in many different forms around the world and in different contexts, e.g. fraud, bribery, extortion, embezzlement, tax evasion, and kickbacks. While corruption has been overlooked to remain politically correct (Klitgaard, 1988), behavioral and experimental economists have begun to reconsider corruption as an area of research (Barr *et al.*, 2004; Cameron *et al.*, 2005; Bertrand *et al.*, 2006). Experimental economic methods allow empirical investigation of corrupt behavior and its definition across cultural contexts. As this process evolves, we can better identify cases in which the Pareto optimality of public good situations is decreased due to higher public good provision costs, and efficiency is lost in market allocation mechanisms. Economic experiments offer researchers the ability to understand the cognitive and cultural contexts that underlie inefficient behavior. While they can be used to measure the group's tendency toward say embezzlement, they can be effective to explain its social acceptability in different cultures and whether the behavior is detrimental on both economic and moral grounds.

Consider for example experiments on the ultimatum bargaining game, which have been run in a variety of developmental and cultural contexts (e.g., Roth *et al.*, 1991; Henrich, 2000; Henrich *et al.*, 2001). The game theoretic expectation of bargaining outcomes is that *Proposers* will give *Respondents* a small offer rate. The experimental results show how different

cultures have different concepts of entitlement. Some groups have high expectations of equal allocations of an endowment; other groups find it much more acceptable for the proposer to keep a higher proportion of the initial endowment.[14] Stretching beyond the experiment, this suggests people in some cultures might find it more acceptable by both the governing and the governed for those in power to retain a greater percentage of societal wealth. Such relationships have not been tested outright, but experimental methods could be useful in defining *corrupt* behavior across cultural and political contexts. Cameron *et al.* (2005) demonstrate the degree to which corrupt behavior is socially acceptable and punishable across cultural groups, and the challenges in using experimental methods to understand complex societal contexts.

Within public goods provision, embezzlement of the actual good can be detrimental to both market and non-market allocation mechanism. Barr *et al.* (2004) demonstrate how experiments can measure embezzlement within groups and be used to test policy mechanisms designed to decrease this tendency. They run their experiment with Ethiopian nursing students to increase the external validity. Ethiopia suffers from embezzlement of health supplies for sale to the private/black market. The experiments were designed to test the influence of employment wage, employee product endowment, and public monitoring regimes on embezzlement behavior. Eight people participated in each experiment session. One subject played the role of the health care worker, another was a monitor, and the remaining people were voting members of the community.

Four different treatments used a 2×2 design (the health care worker is randomly selected in all treatments). The monitor was either randomly selected or elected from the community; the health care worker was either paid a high amount (60 Birr per round) or a low amount (20 Birr per round). The game begins with the health care worker sitting behind a blind to hide his/her activity from the rest of the group. The worker rolled a six-sided die to determine how many 'valuable' tiles he or she receives, worth 60 Birr each. He or she is given additional 'worthless' tiles (worth 20 Birr each) so that he/she has 10–18 tiles in total. The health care worker decides which six of the tiles he or she will give to the group. The tiles are given to the group in a bag without revealing the tile type. The monitor can decide whether he or she wants to inspect the bag's contents. The monitor can inspect up to four tiles at a cost of 5 Birr per tile. The cost is deducted from the Monitor's total salary (60 Birr). If the monitor finds a 'worthless' tile in the bag, the health care worker is not compensated. After inspection, the bag is passed to the community and each community member blindly draws one tile from the bag, receiving 60 Birr for a 'valuable' tile and nothing for a 'worthless' tile.

[14] Early experiments, conducted in developed countries, however, reveal that subjects expect the Proposer to give an average of 40 to 50 per cent of the endowment to the Respondent (e.g., Roth *et al.*, 1991). There have been observations of lower offer rates (e.g., 26–27 per cent) in certain indigenous groups, such as the Machiguenga in Peru, Hadza in Tanzania, and Quichua in Ecuador (Henrich, 2000; Henrich *et al.*, 2001). These findings suggest cultural experience significantly impacts subjects' perceptions of entitlement and endowment distribution.

The results reveal how embezzlement can occur in the lab. The different treatments reveal how several factors affected embezzlement – the health care workers' pay, the monitor selection method (random or democratic), and the number of valuable tiles initially received. The higher-paid health care workers retained fewer valuable tiles, reflecting their pay and how monitors scrutinized higher-paid workers with greater frequency. The monitors were more active when they were elected rather than randomly selected. Health care workers kept more valuable tiles when they were endowed with more tiles. The example shows us how embezzlement can be replicated in the experimental lab, and what conditions make it more or less rife within a group. While this experiment was conducted within a health care industry context, one could develop similar designs to address environmental and resources policy in developing countries.

3. Applying the experimental method

People with the experimental mindset use the experimental method for four reasons: to test theory, look for empirical patterns, testbed new institutions, and to educate people about economic principles. We now illustrate these various applications of experimental methods with regard to EDE issues around rents. Following Barbier (2005), poverty persists when the rents generated from resources are less than optimal due to poor technology or education (Rent Capture), and when these suboptimal rents are spent on current consumption rather than re-invested in new technology/education (Rent Use), invested ineffectively (Rent Reinvestment), or are wasted in fights over the rents (Rent Seeking). Table 3 summarizes a matrix of the rationale for using the experimental method and EDE research opportunities. We discuss a subset of research presented in the table, with the idea that many of the items listed are either self-explanatory or speculative. Also see the summary tables in Carpenter and Cardenas (2006) on cooperation, trust, altruism, fairness, and preferences in developing country contexts.

3.1. Tests of economic theory

Theory gives coherent meaning to empirical results. Experiments are useful to test theoretical predictions about behavior, rational or otherwise. Researchers can use the method to test the predictive power of a theory – reject or not reject null hypotheses of predicted choices and reactions to changes in exogenous parameters. Experiments provide a sterile environment to test theory by controlling confounding factors. One interesting issue in EDE work is on the security of property rights. Most economists believe a secure property rights system is crucial for economic development since it provides incentives for people to invest hard work and capital to create and store private wealth (see, e.g., Robinson, 2005). With insecure rights, rents can be appropriated by others, which either triggers a race to overdevelop or over harvest open-access resources or to provide less capital than socially optimal. There is a downside to secure property rights, however, when considered in the lens of the recent enthusiasm for collaborative decision making. Many experts promote Coasean-style

Table 3. *Matrix of experimental rationales and environmental and development areas*

	Environmental and development areas			
Experimental rationale	*Rent capture*	*Rent use*	*Rent reinvestment*	*Rent seeking*
Testing theory	Common pool resources; Open access models; Property rights security	Consumption vs Savings decisions; human capital decisions	Hartwick Rule of sustainable consumption	Conflict models and gaming behavior; Corruption models
Pattern recognition	Risk and time preference elicitation	Reaction to credit markets; Trust and reciprocity	Portfolio diversification; Credit markets & asymmetric information	Degree of rent dissipatation; Trust games
Economic design	Institutional design of common property resources	Micro-finance institutions	Smart subsidies for conservation	Institutional design and changes
Education	Open access games; Coasean bargaining	Trust games	Coordination & Trust games	Normal form Contests; VCM

processes at the local level as a tool to resolve environmental disputes and develop resource management plans (e.g., Susskind *et al.*, 2000).

One relevant question in conflict resolution is how property rights security affects bargaining efficiency given positive transaction costs (see Amman and Duraiappah, 2004). The answer is ambiguous. Secure property rights avoid costly conflict over who owns what, but they also imply an owner might want to avoid otherwise socially efficient bargaining if it is too expensive to him- or herself. In contrast, insecure rights trigger costly conflict but all parties are also forced to go to the bargaining table to capture the rents. The dominant effect is unclear – avoiding costly conflict or bargaining gains.

Cherry and Shogren (2005) design a bargaining experiment to help understand which theoretical explanation has more behavioral support. Consider a Coasean bargain between two players negotiating over lottery tickets defining the likelihood of winning a monetary payoff. All bargaining has positive transaction costs based on offers, evaluations, and counter-offers. They assume one player is the *controller* – the player with property rights – who has a positive probability of winning the large reward when he exercises the outside option if this property right is upheld when challenged. Solving the bargaining problem yields the Nash solution for the split of lottery tickets based on the outside option, the impact of insecure

property rights on player A's tickets, and the effect of transaction costs. The player with property rights should prefer to bargain since taking the outside option only becomes attractive when transaction costs are too large. The experimental design considered face-to-face, ten minute bilateral negotiation over lottery tickets, reflecting chances of winning a large reward.

The results suggest the counter-intuitive result that greater property rights security lowered average bargaining efficiency given transaction costs. Rather than pay the transaction costs associated with bargaining, many property rights owners took their secure outside option. Future work on bargaining should explore a broader perspective in which bargainers choose between private and socials gains from protecting new investments in capital to create wealth versus the social losses that arise when avoiding costly bargaining.

3.2. Pattern recognition and preference measurement

People also use the experimental method to look for patterns of behavior. One can explore how people reveal their demand, construct preferences and beliefs, observe breakdowns from rationality, examine how contextual frames affect behavior, determine reactions to information, and consider how people coordinate actions given multiple equilibria. Pattern recognition can provide the motivation for theorists to develop alternative models based on *ex post* rationalizations of observed behavior. Raiffa's (1982) work on the art and science of negotiation is a classic example of using experiments for pattern recognition. He created a classroom 'quasi-laboratory' in which students discussed how the heuristic insight gained from observing bargaining behavior in the classroom might be applied to the real world.

One useful application of the experimental method to EDE is to examine the underlying preferences for time and risk of the poorest populations who see frontier expansion on to fragile lands as their best chance to increase wealth, e.g. deforestation for agriculture (see Barbier, 2005). While many experiments have been conducted in developing countries to measure risk and time preferences, few have used the method to help guide EDE policy. Binswanger's (1980) experiment measuring risk preferences of Indian farmers was an exception; follow up experiments mainly consider preferences without social context or are in a non-environmental or agriculture setting (e.g., Wik *et al.*, 2004).

One recent exception is Godoy *et al.* (2001), who use experimental methods to test the role of time preference in individual and household's resource use in the Bolivian lowlands. Assuming hyperbolic discounting, they do not find a significant relationship between the experimental measure of time preference and actual resource use behavior. More work on time preference experiments and the application of time preference measures to environmental modeling is needed (Botelho *et al.*, 2005). Research applying discount rate measures to actual resource use behavior could be a catalyst to develop better experimental measures of time preference, which reflect and explain cultural contextual issues.

Ehmke *et al.* (forthcoming) demonstrates the role experiments can play to increase our ability to perform benefits transfer across a variety of cultural and geographic contexts. Ehmke *et al.* discuss the difficulty hypothetical bias adds to benefits transfer between countries and locations. The basic difficulty of transferring benefits from developed to developing countries is also demonstrated in Barton and Mourato's (2003) work. After controlling for socio-demographic variables, they find little reliability in WTP estimates transferred from Portugal to Costa Rica. Other issues related to CVM use may differ or need attention in a developing context. Starting point bias is an issue that has been identified and measured in a developing country. Bhatia (2005) finds it present in respondents' WTP for mosquito netting to reduce malaria. Surprisingly, both benefit transfer and starting point bias have not been addressed for environmental goods in developing countries. Experiments like the one borrowed from List and Shogren (1998) can be used to measure hypothetical bias across a greater number of goods and for more direct application to a certain issue.

Another use of the method for pattern recognition is to measure cultural differences in *shallow* market behavior. Many experiments, motivated by anthropologic questions, are used to measure cultural and societal differences or tendencies in social norms, especially negotiation and bargaining behavior (Roth *et al.*, 1991; Cason *et al.*, 2002; Croson and Buchan, 1999). Applying these experiments to environmental–development challenges is scarce. Examples of experiments trying to bring out the culture and developing context of their experimental subjects include Carpenter *et al.* (2004), Greig (2005), and Cameron *et al.* (2005).

Carpenter *et al.* (2004) use a version of the Voluntary Contribution Mechanism to explain the role of trust and cooperation in the social capital and social dilemmas faced by residents of Southeast Asian urban slums. Greig builds a more direct application of cooperation measures from a public goods experiment to the formation of women's microfinance groups (ROSCAs) in Kenyan slums. She finds much more cooperation in women-only compared to mixed gender groups. This experiment provides insight into the effectiveness of women-only ROSCAs in the person's everyday environment. Finally, Cameron *et al.* explore the mechanisms to explain cultural differences in corruption and punishment behavior. While these experiments do not have applications to environmental goods, one can imagine possible applications. The mechanisms may be adjusted to reflect local environmental conditions and localized policy options.

These experiments raise the general issue of external validity and replication beyond the lab and field experiment. In general, the power of the experiment method is control – control over preferences, technology, endowments, and institutions – such that behavioral changes can be understood when the economic circumstances change (e.g., price increase, new technology). Social preferences and social context, however, play important roles in setting the institutional stage, and the lack of understanding or the elimination of these contexts can confound any experimental results. The transferability of one experiment to a policy setting becomes questionable at best. Lab experiments can be too sterile relative to the real-world application if key contextual elements are not

addressed. Field experiments that move closer to reality are no panacea either if the sample selected is a 'convenience sample' – examining the behavior of people which are easy to have participate in an experiment, i.e. people who already chose, for whatever reason, to go to some nearby memorabilia show. In general, as has been well-known for over a century in the broader scientific literature on the use of experimental methods, external validity is most likely achieved by randomization – randomly selecting a community of people, and then given these people, randomly selecting people with the treatment and control groups (also see Carpenter and Cardenas, 2006; Duflo *et al.*, 2006).

3.3. Economic design

Over the years, economists have designed new economic instruments or institutions to increase allocative efficiency given missing markets or market failure. This process of economist as engineer has come to be called *economic design*. The experimental method has emerged as a popular way to testbed these new institutions to provide evidence of success before actual implementation. One prominent example of a testbed was experiments examining the efficiency of the auctions designed to sell off the rights to the spectrum in the United States (see e.g., Plott, 1994). Examples of testbed experiments in environmental economics include market-based emission trading and water trading. These experiments examine how flexibility, imperfect information, market size, and market power affect the potential efficiency of trading (e.g., Bohm and Carlen, 1999; Dinar *et al.*, 2000). These experimental results provide evidence about how people could behave in a real-world application. Additional real-world complexity can be added into the experiment environment in a controlled fashion to identify likely conditions that might cause a mechanism to fail in the wild.

One relevant question for EDE is how institutions can be designed to help people in developing countries deal with debt and savings. There is a fear that high indebtedness of the nation and low savings rates of people can increase resource exploitation (see Neumayer, 2005). A good example of research exploring such questions is Ashraf *et al.*'s (2006) experiment designed to testbed a commitment savings instrument in a developing country. They also examine how hyperbolic discounting may affect the investment selection decisions. The authors give people the option to invest in a commitment savings instrument designed so people invest in a savings account with limited opportunities to withdraw funds. They can withdraw money at a specific time in the future or when a savings objective is met, but not beforehand. The authors test whether people who practice hyperbolic discounting know their tendencies and will be more likely to open such a commitment savings account. Further, they test whether opening such an account will increase savings for these people.

The experiment begins with a hypothetical time preference questionnaire administered to 1,777 households in the Philippines. The researchers used the questionnaire to distinguish households with hyperbolic discounting preferences from those without such preferences. After the survey step was complete, households were randomly selected from the pool to participate in a natural experiment. Half selected were offered a commitment savings

account with a local banking institution. This account was limited in access to deposits only until the client either met pre-specified savings time requirement or investment goals. The other households were randomly divided in half again and placed into a control group or marketing group. The control group did not receive any further offers. The marketing group received an additional marketing visit in which they were encouraged to invest in traditional (non-commitment) savings products.

The survey results were used to classify clients as those with or without hyperbolic discounting preferences based on their choices over sets of monetary and non-monetary payment and satisfaction delay options. Clients were classified as having hyperbolic preferences if they wanted immediate rewards in the short term, but were willing to wait for a higher amount in the long term. About 27 per cent of the people in the survey displayed hyperbolic preferences. People dissatisfied with current savings habits are more likely to be hyperbolic. Empirical analysis reveals people who are time inconsistent, or displaying hyperbolic discount preferences, are most likely to take up the commitment savings account. This is significant for non-married females. The group offered the commitment savings account displayed higher rates of savings.

Another key question is incentive design for conservation (O'Connor, 1999; Winters *et al.*, 2004).[15] Experiments provide a tool to testbed proposed instruments. Cason and Gangadharan (2004), for instance, examine how information about public benefits and the market clearing mechanism affect the efficiency of the Bush Tender auctions designed to conserve land in Australia. Their experimental results reveal an interesting pattern: people who did not know the environmental benefits provided by their private land were less likely to bid strategically in a conservation auction. Private ignorance reduces public expenditures. Based on this they suggest a provocative policy – a regulator might restrict the biological information publicly provided to landowners prior to running the auction. Aside from the ethics of a government keeping its citizens in the dark, a landowner may or may not rely on the government for his or her sole source of information. He or she has the option to hire a private biologist to appraise the environmental benefits on his or her land. The landowner has an outside option, one not addressed in their experiment. Future work can add this extra complexity to test the robustness of these initial findings. If landowners know or will pay to learn about the public benefits on their own lands, the Australian policymakers guided by experimental results for conservation auctions should know how such an outside option could affect auction behavior.

Another example of testbedding is Parkhurst and Shogren's (2007) agglomeration bonus/smart subsidy coordination game experiments. The idea here is to testbed an incentive scheme that can induce private landowners to create contiguous protected areas voluntarily. They use experimental economic methods to testbed a *smart subsidy* proposal relative

[15] O'Connor (1999: 106) discusses 'experimentation' in terms of the real-world application of a new instrument at a local level to determine its potential acceptability before any actual implementation at a national level.

to two standard policy options: compulsion and a standard fixed-fee subsidy. A smart subsidy creates an explicit link between neighboring landowners by paying an additional agglomeration bonus when they retire land adjacent to other conserved parcels, both their own and their neighbors. Whereas earlier incentives treated each landowner as an isolated entity, the smart subsidy creates incentives for landowners to 'cooperate' by retiring land along their adjoining fence lines.

Their experimental design had six structural elements – landscape and landowners, policy treatments and subsidy design, players and rounds, game strategies, calculator and communication, and information and history. They represented the landscape with a 10×10 land grid, divided into four private 5×5 landholdings. Some parcels are highly valued for development or agricultural production, others are not; some parcels provide better habitat for the species than others. The smart subsidy divides a landowner's payment into four distinct parts: (1) a $20 flat fee per cell retired; (2) a $50 own-border bonus for each common border shared between two of their own retired cells; (3) a $24 row-border bonus for each shared border with his or her row-neighbor; and (4) a $22 column-border bonus for each shared border with his or her column-neighbor. The amount of each bonus payment depends on the productive values and desired configuration and location of the habitat, and can be positive, negative, or zero.

Each subject had a 10×10 grid calculator on the computer screen to assist him or her in calculating profits. People could send one message per round to the group. People had two minutes to send messages, use the calculator, and send their choices. After all brown-out choices were submitted, the resulting land grid was shown to the group. People had common knowledge regarding payoffs and strategies. Each person's individual payoffs and accumulated payoffs were private information. The entire 10×10 grid showing the configuration of brown cells and the payoffs for each subject within the group then appeared in the history box. People had record sheets and the history box to help them keep track of their own and the other group members' choice of strategies and associated payoffs in previous rounds.

The empirical question is whether people can coordinate their land retirement decisions to find the single *first-best* equilibrium outcome (contiguous targeted habitat) out of a set of over six thousand potential *second-best* equilibria (fragmented habitat). Their results suggest the smart subsidy outperforms the alternative policies at creating the desired contiguous habitat. While success in the lab does not guarantee the same in actual settings, what these experimental results do is contribute to better evidence-based policy making. The favorable testbed results suggest policy makers might consider adding the smart subsidy idea to their list of pilot studies exploring approaches for voluntary conservation on private lands and across public–private landholdings.

Experiments can be used to testbed regulation policy for a variety of public good and common pool resource problems. One challenge faced by governments is to identify the correct level of (a) punishment/fines for free-riding or non-cooperative behavior, and (b) monitoring to detect such

behavior. In addition, the effectiveness and efficiency of enforcement policy depends on the range of individual risk preferences within a population; assuming risk neutrality, punishment and detection are perfect substitutes (see Becker, 1968). The governance of public resources can be improved if governments can determine *a priori* the most efficient or effective level of punishment and monitoring of illegal or non-cooperative behavior around public resources. If policies are implemented, it remains difficult to detect non-contributors in the field, whereas such variables are easily detected and measured in the laboratory. A number of pre-existing experimental designs lend themselves to such regulatory policy test-bedding. But as is the case with other experiments, their current application to EDE contexts is limited.

Briefly, we examine some experimental studies measuring the tradeoff between punishment levels and monitoring. In a developed country context, experiments have explored the trade-off between punishment levels and detection across several topics including public good contributions, industrial collusion, tax evasion, and parental delay in picking up pre-school children. Evidence suggests non-contributors who are not caught free-riding are more likely to continue to free ride in the future than those who are caught (see Anderson and Stafford, 2003). Another study explored the difference between students and prisoners. They observed students were deterred more by the punishment size (e.g., fine) than the probability of detection. In contrast, prisoners were more sensitive to the probability of detection than punishment size because they were relatively more risk-seeking (see Block and Gerety, 1995).

The role of punishment, ethical obligation, and social contracts has also been explored. For instance, Feld and Tyran (2002) gave people the opportunity to vote on a regulation to contribute to the good before making contributions to the good. Evidence suggests people contribute more to goods that are *legitimized* through the social contract created by majority-peer-support voting outcomes. Another study considered how punishment crowds out ethical obligations implied by the social contract. Focusing on daycare providers in Israel, Gneezy and Rustichini (2000) imposed a fine on parents who were late picking up their children. They observed the fine *increased* the number of late parents, suggesting the fine changed the implicit social contract in the institution, indicating it was 'okay' to be late. The problem with this result, however, is that the fine was extremely low – the cost of leaving a child at daycare was only a few dollars per hour, which is a cheap 'babysitting service', which should and did increase the demand for babysitting in the form of late pick up of children (see Rubinstein, 2006).

In a developing country context and an environmental good, Cardenas *et al.* (2000) examine punishment within a common pool resource game. This study is an excellent example of how to use an experiment to test-bed government policy around punishment. They examine different strategies to encourage efficient and sustainable firewood collection in a rural Columbian village. In their base treatment, people decide how much time they will spend collecting firewood, out of eight months. The more time they all spend collecting firewood, the more deforestation. In the base treatment, theory predicts people will spend too much time collecting firewood relative to the social optimum (one month). Treatments 2 and 3 introduced

government regulation (punishment) and peer-to-peer communication to give incentives for socially optimal firewood collection. In the regulation treatment, the Nash Equilibrium for a risk neutral subject is five months with and six months without regulation. In the communication treatment, no formal prediction is made; rather they observe and compare outcomes to Treatment 2. They find regulation *increases* the average time spent collecting firewood; whereas the communication treatment produces the most socially efficient outcomes. This experiment illustrates how economic design can capture social context variables – i.e., the complex role of communication within the group – that goes beyond the scope of most theoretic predictions.

The effectiveness of punishment, however, should not be viewed in isolation. Rewards are also a classic incentive designed to promote cooperative behavior in social dilemma games. Regulators can choose between punishments versus rewards to generate more efficient behavior. Early experiments observed that rewards were less effective than punishment. Andreoni *et al.* (2003) revisit this question with a straightforward game of proposer–responder given separate and joint reward and punishment schemes. Their results support earlier work in that rewards alone were insufficient in generating relative cooperative behavior – but rewards in combination with the threat of punishment generated the most generous offers. Rewards complemented by punishment worked best together, even if punishment was rarely used.[16]

These testbed examples illustrate that economic design is ripe with opportunity to learn about the role of institutions and incentives for environmental and development economics. We now consider how people use experiments as teaching tools.

3.4. Experiments as teaching tools

Classroom experiments allow instructors to illustrate and explain the uniqueness of developing country environmental economic issues in any undergraduate classroom. Murphy and Cardenas (2004) provide a good example of a classroom experiment related to a developing country environmental and development problem. Their experiment can be used to teach concepts such as public good, common pool resource, efficiency, private welfare, social welfare, and Nash equilibrium using a developing country example.

The common pool resource is firewood. Students must decide how much time they will use to collect firewood in a year. An environmental tradeoff exists between the firewood collected and water quality. The experiment is run under three different treatments. The first treatment has no regulation or communication between students; the second has weak governmental

[16] Previous reward-driven experiments used a *transfer reward* system in which there is a one-for-one transfer of wealth, i.e., he pays $1 to reward her; she receives a $1 reward. In contrast, Andreoni *et al.* used a 'net postive' reward scheme: he pays $1 to reward her; she receives a $1.50 reward. In turn, Vyrastekova and van Soest (2008) tested whether the reward system itself affected its effectiveness. They found the transfer reward to be less effective relative to the 'net positive' reward in generating cooperative behavior.

regulation; and the third has no regulation with communication between students (a proxy for local self-government). The experiment has been run in developed country classrooms and in field labs using Columbian villagers. Their results are consistent across location. They find the most frequent result is the social welfare optimizing level of resource extraction (one month of firewood collection) is attained using the communication treatment. The experiment can be used to stimulate classroom discussion about the role of self government versus external regulation in achieving economic efficiency.

The experiment can also be useful to undergraduate classes in developing countries as it helps students develop methods to deal with local problems. But this experiment can be run in developed countries to expose students to environmental problems developing countries face and to expose students to questions they might not have considered relevant to their lives (e.g., firewood collection) to improve their understanding on how both the market and governments can fail the people they are supposed to serve. The experiment also illustrates how non-obvious policies may be more efficient than traditional policies to address an environmental problem.

Classroom experiments can also be used to help policy-makers overcome another challenge of implementing a program like tradable permits: teaching the economic agents involved in the quota system how it works. Households and local businesses affected by a quota could learn about its mechanics through a field classroom experiment. Due to the complexity of this and other policy solutions, active participation in field classroom experiments may be the most effective way for some citizens, especially illiterate citizens, to understand new policies and regulations. Field classrooms may be useful to help implement tradable permits on a local and regional scale in non-traditional geographic locations.

4. Conclusions

Ten years ago, Dasgupta and Mäler (1997: 1) noted 'until very recently official development economics did not acknowledge their existence [environmental resources], in that you would not find them in any recognized survey article, or text, or treatise on the subject'. Today one could make a similar observation for the use of the experimental method in environment and development economics, although this is changing with vigor. The experimental method has proven itself a useful tool to frame questions and gather insight into the behavioral underpinnings of environmental policy in developed nation contexts (e.g., Murphy *et al.*, 2000). Over the next decade, the method will inevitably be used more widely to further our understanding of environmental and development policy. The relative benefits of using experiments for resource policy in developing countries may even be greater than in the developed world. Due to the low capital resource constraints, the rewards may be significant in using cost-effective experimental research relative to pilot projects. As our survey suggests, we expect this mindset to expand over the next decade as researchers continue to ask questions that can be best addressed using the tools provided by experimental economics.

References

Abbink, K., B. Irlenbusch, and E. Renner (2006), 'Interest rates in group lending: a behavioural investigation', *Pacific Economic Review* **11**: 185–199.

Alpizar, F. and F. Carlsson (2003), 'Policy implications and analysis of the determinants of travel mode choice: an application of choice experiments to metropolitan Costa Rica', *Environment and Development Economics* **8**: 603–620.

Alvard, M. (2004), 'The ultimatum game, fairness, and cooperation among big game hunters', in J. Henrich *et al.* (eds), *Foundations of Human Society: Economic Experiments and Ethonographic Evidence from Fifteen Small-scale Societies*, Oxford: Oxford University Press, pp. 413–435.

Amman, H. and A. Duraiappah (2004), 'Land tenure and conflict resolution: a game theoretic approach in the Narok District in Kenya', *Environment and Development Economics* **9**: 383–408.

Anderson, L. and S. Stafford (2003), 'Punishment in a regulatory setting: experimental evidence from the VCM', *Journal of Regulatory Economics* **24**: 91–110.

Andreoni, J., W. Harbaugh, and L. Vesterlund (2003), 'The carrot or the stick: rewards, punishment, and cooperation', *American Economic Review* **93**: 893–902.

Ashraf, N., D. Karlan, and W. Yin (2006), 'Tying Odysseus to the mast: evidence from a commitment savings product in the Philippines', *Quarterly Journal of Economics* **121**: 673–697.

Barbier, E. (2005), *Natural Resources and Economic Development*, Cambridge: Cambridge University Press.

Bardsley, N. (2007), 'Dictator game giving: altruism or artefact?', *Experimental Economics* (in press).

Barr, A. (2002), 'The functional diversity and spillover effects of social capital', *Journal of African Economics* **11**: 90–113.

Barr, A. (2003), 'Trust and expected trustworthiness: experimental evidence from Zimbabwean villages', *Economic Journal* **113**: 614–630.

Barr, A. and B. Kinsey (2002), 'Do men really have no shame?', CSAE WPS/2002-05, University of Oxford, Oxford, England.

Barr, A., M. Lindelow, and P. Serneels (2004), 'To serve the community or oneself: the public servant's dilemma', World Bank Policy Research Working Paper 3187, Washington, DC.

Barton, D. and S. Mourato (2003), 'Transferring the benefits of avoided health effects from water pollution between Portugal and Costa Rica', *Environment and Development Economics* **8**: 351–371.

Becker, G. (1968), 'Crime and punishment: an economic approach', *Journal of Political Economy*. **76**: 169–217.

Berg, J., J. Dickhaut, and K. McCabe (1995), 'Trust, reciprocity, and social history', *Games and Economic Behavior* **10**: 122–142.

Bertrand, M., S. Djankov, R. Hanna, and S. Mullainathan (2006), 'Obtaining a driving license in India: an experimental approach to studying corruption', Paper presented at the 2006 Allied Association Meetings, Boston, MA.

Bhatia, M. (2005), 'From evidence to calibration for starting point bias: willingness to pay for treated mosquito nets in Gujarat, India', *Applied Economics* **37**: 1–7.

Binswanger, H. (1980), 'Attitudes toward risk: experimental measurement in rural India', *American Journal of Agricultural Economics* **62**: 395–407.

Block, M. and V. Gerety (1995), 'Some experimental evidence on differences between student and prisoner reactions to monetary penalty and risk', *Journal of Legal Studies* **24**: 123–138.

Bohm, P. (2003), 'Experimental evaluations of policy instruments', in K.-G. Mäler and J. Vincent (eds), *Handbook of Environmental Economics*, vol. 1, Amsterdam: North-Holland, pp. 437–460.

Bohm, P. and B. Carlén (1999), 'Emission quota trade among the few: laboratory evidence of joint implementation among committed countries', *Resource and Energy Economics* **21**: 43–66.
Botelho, A., G. Harrison, L.M. Costa Pinto, E. Rutström, and P. Veiga (2005), 'Discounting in developing countries: a pilot experiment in Timor-Leste', Working Paper, University of Central Florida.
Bowles, S. (1998), 'Endogenous preferences: the cultural consequences of markets and other economic institutions', *Journal of Economic Literature* **36**: 75–111.
Camerer, C., G. Loewenstein, and M. Weber (1989), 'The curse of knowledge in economic settings: an experimental analysis', *Journal of Political Economy* **97**: 1232–1254.
Cameron, L., A. Chaudhur, N. Erkal, and L. Gangadharan (2005), 'Do attitudes towards corruption differ across cultures? Experimental evidence from Australia, India, Indonesia, and Singapore', Working Paper, University of Melbourne, Melbourne.
Cardenas, J.C., J. Stranlund, and C. Willis (2000), 'Local environmental control and institutional crowding-out', *World Development* **28**: 1719–1733.
Carpenter, J.P. (2007), 'The demand for punishment', *Journal of Economic Behavior and Organization* **62**: 522–542.
Carpenter, J. and J.C. Cardenas (2006), 'Behavioral development economics: lessons from field labs in the developing world', *Journal of Development Economics* (forthcoming).
Carpenter, J., A. Daniere, and L. Takahashi (2004), 'Cooperation, trust, and social capital in Southeast Asian urban slums', *Journal of Economic Behavior and Organization* **55**: 533–551.
Cason, T. (1995), 'An experimental investigation of the seller incentives in the EPA's emission trading auction', *American Economic Review* **85**: 905–922.
Cason, T. and L. Gangadharan (2004), 'Auction design for voluntary conservation programs', *American Journal of Agricultural Economics* **86**: 1211–1217.
Cason, T., L. Gangadharan, and C. Duke (2003), 'A laboratory study of auctions for reducing non-point source pollution', *Journal of Environmental Economics and Management* **46**: 446–471.
Cason, T., T. Saijo, and T. Yamato (2002), 'Voluntary participation and spite in public good provision experiments: an international comparison', *Experimental Economics* **5**: 133–153.
Cherry, T., T. Crocker, and J. Shogren (2003), 'Rationality spillovers', *Journal of Environmental Economics and Management* **45**: 63–84.
Cherry, T., S. Kroll, and J. Shogren (eds.) (2008), *Environmental Economics, Experimental Methods*, London: Routledge.
Cherry, T. and J. Shogren (2005), 'Costly Coasean bargaining and insecure property rights', *Environmental and Resource Economics* **31**: 349–367.
Coase, R. (1960), 'The problem of social cost', *Journal of Law and Economics* **3**: 1–44.
Crocker, T. (1966), 'The structuring of atmospheric pollution control systems', in H. Wolozin (ed.), *The Economics of Air Pollution*, New York: Norton, pp. 61–86.
Crocker, T., J. Shogren, and P. Turner (1998), 'Incomplete beliefs and nonmarket valuation', *Resource and Energy Economics* **20**:139–162.
Croson, R. and N. Buchan (1999), 'Gender and culture: international experimental evidence from trust games', *American Economic Review* **89**: 386–391.
Dasgupta, P. and K.-G. Mäler (1997), 'The resource basis of production and consumption: an economic analysis', in P. Dasgupta and K.-G. Mäler (eds), *The Environment and Emerging Development Issues*, Oxford: Oxford University Press, pp. 1–32.
Davis, D.D. and C.A. Holt (1993), *Experimental Economics*, Princeton: Princeton University Press.

Degenne, A. and M. Forsé (1999), *Introducing Social Networks*, Thousand Oaks: Sage Publications.
Dinar, A., R. Howitt, J. Murphy, S. Rassenti, and V.L. Smith (2000), 'The design of "smart" water market institutions using laboratory experiments', *Environmental and Resource Economics* **17**: 375–394.
Duflo, E. (2005), 'Field Experiments in Development Economics', Working paper, MIT, Cambridge, MA.
Duflo, E., R. Glennerster, and M. Kremer (2006), 'Randomization in development economic research: a toolkit', Working paper, MIT, Cambridge, MA.
Eggert, H. and P. Martinsson (2004), 'Are commercial fishers risk-lovers?', *Land Economics* **80**: 550–560.
Ehmke, M., J. Lusk, and J. List (forthcoming), 'Is hypothetical bias a universal phenomenon? A multi-national investigation', *Land Economics*.
Ensminger, J. (2001), 'Reputations, trust, and the principal agent problem', in K. Cook (ed.), *Trust in Society*, Volume II, New York: Russell Sage Foundation.
Ensminger, J. (2004), 'Market integration and fairness', in J. Henrich *et al.* (eds), *Foundations of Human Society: Economic Experiments and Ethonographic Evidence from Fifteen Small-scale Societies*, Oxford: Oxford University Press, pp. 356–381.
Fehr, E. and S. Gächter (2000), 'Cooperation and punishment in public goods experiments', *American Economic Review* **90**: 980–994.
Feld, L. and J.-R. Tyran (2002), 'Tax evasion and voting: an experimental analysis', *Kyklos* **55**: 197–222.
Fogel, R. (2004), *The Escape from Hunger and Premature Death, 1700–2100: Europe, America, and the Third World*, Cambridge: Cambridge University Press.
Frederick, S., G. Loewenstein, and T. O'Donoghue (2002), 'Time discounting and time preference: a critical review', *Journal of Economic Literature* **40**: 351–401.
Gebremedhin, B., J. Pender, and G. Tesfay (2003), 'Community natural resource management: the case of woodlots in northern Ethiopia', *Environment and Development Economics* **8**: 129–148.
Gneezy, U. and A. Rustichini (2000), 'A fine is a price', *Journal of Legal Studies* **26**: 1–17.
Godoy, R., K. Kirby, and D. Wilkie (2001), 'Tenure security, private time preference, and use of natural resources among Lowland Bolivian Amerindians', *Ecological Economics* **38**: 105–118.
Greig, F. (2005), 'Gender and the voluntary provision of public goods: evidence from a slum in Nairobi, Kenya', Paper presented at the Regional Economic Science Association Meetings, Tucson, Arizona.
Gurven, M. (2004), 'Does market exposure affect economic game behavior? the ultimatum game and the public goods game among the Tsimane' of Bolivia', in J. Henrich *et al.* (eds), *Foundations of Human Society: Economic Experiments and Ethonographic Evidence from Fifteen Small-scale Societies*, Oxford: Oxford University Press, pp. 194–231.
Haab, T. and K. McConnell (2002), *Valuing Environmental and Natural Resources*, Northampton, MA: Edward Elgar Publishing, Inc.
Hanley, N. and C. Spash (1993), *Cost–Benefit Analysis*, Northampton, MA: Edward Elgar Publishing, Inc.
Hanley, N., J. Shogren, and B. White (2007), *Environmental Economics: Theory and Practice*, 2nd edition, London: Palgrave.
Harrison, G., M. Lau, and M. Williams (2002), 'Estimating individual discount rates in Denmark: a field experiment', *American Economic Review* **92**: 1606–1617.
Harrison, G. and J. List (2004), 'Field experiments', *Journal of Economic Literature* **42**: 1013–1059.
Hayek, F. (1967), *Studies in Philosophy, Politics, and Economics*, Chicago: University of Chicago.

Henrich, J. (2000), 'Does culture matter in economic behavior? Ultimatum game bargaining among the Machieguenga of the Peruvian Amazon', *American Economic Review* **90**: 973–979.
Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, and R. McElreath (2001), 'In search of Homo economicus: behavioral experiments in 15 small-scale societies', *American Economic Review* **91**: 73–78.
Hoffman, E. and M. Spitzer (1982), 'The Coase theorem: some experimental tests', *Journal of Law and Economics* **25**: 73–98.
Hofstede, G. (1980), *Cultures Consequences*, Beverly Hills: Sage Publications.
Hofstede, G. (1991), *Cultures and Organizations: Software of the Mind*, United Kingdom: McGraw-Hill.
Holden, S., B. Shiferaw, and M. Wik (1998), 'Poverty, market imperfections and time preferences: of relevance for environmental policy', *Environment and Development Economics* **3**: 105–130.
Holt, C. and S. Laury (2002), 'Risk aversion and incentive effects', *The American Economic Review* **92**: 1644–1655.
Horan, R., J. Shortle, and D. Abler (2002), 'Ambient taxes under m-dimensional choice sets, heterogeneous expectations, and risk aversion', *Environmental and Resource Economics* **21**: 189–202.
Johansson-Stenman, O., M. Mahmud, and P. Martinsson (2008), 'Trust and religion: experimental evidence from Bangladesh', *Economica* (forthcoming).
Kagel, J. and A. Roth (eds) (1995), *Handbook of Experimental Economics*, Princeton, NJ: Princeton University Press.
Kahneman, D. and A. Tversky (eds) (2000), *Choices, Values, and Frames*, Cambridge: Cambridge University Press.
Karlan, D. (2005), 'Using experimental economics to measure social capital and predict financial decisions', *American Economic Review* **95**: 1688–1699.
Kirigia, J. (1998), 'Cost–utility analysis of schistosomiasis intervention strategies in Kenya', *Environment and Development Economics* **3**: 319–346.
Klitgaard, R. (1988), *Controlling Corruption*, Los Angeles: University of California Press.
Knack, S. and P. Keefer (1997), 'Does social capital have an economic payoff? A cross-country investigation', *Quarterly Journal of Economics* **112**: 1251–1288.
Knetsch, J. (1997), 'Evaluation and environmental policies: recent behavioural findings and further implications', in A. Dragun and K. Jakobsson (eds), *Sustainability and Global Environmental Policy: New Perspectives*, Cheltenham: Edward Elgar Publishing, pp. 193–212.
Kroll, S., T. Cherry, and J. Shogren (2007), 'Voting, punishment and public goods', *Economic Inquiry* **45**: 557–570.
Ledyard, J. (1995), 'Public goods: a survey of experimental research', in J.H. Kagel and A.E. Roth (eds), *The Handbook of Experimental Economics*, Princeton, NJ: Princeton University Press.
Levine, R. and A. Norenzayan (1999), 'The pace of life in 31 countries', *Journal of Cross-Cultural Psychology* **30**: 178–205.
Levitt, S. and J. List (2008), 'Homo economicus evolves', *Science* **319**: 909–910.
List, J. and J. Shogren (1998), 'Calibration of the differences between actual and hypothetical valuations in a field experiment', *Journal of Economic Behavior and Organization* **72**: 615–625.
Maital, S. and S. Maital (1977), 'Time preference, delay of gratification and the intergenerational transmission of economic inequality: a behavioral theory of income distribution', in O. Ashenfelter and W. Oates (eds), *Essays In Labor Market Analysis*, New York: Halstead Press.

Masclet, D., C. Noussair, S. Tucker, and M.-C. Villeval (2003), 'Monetary and nonmonetary punishment in the voluntary contributions mechanism', *American Economic Review* **93**: 366–380.
Mason, C. and O. Phillips (1997), 'Mitigating the tragedy of the commons through cooperation: an experimental evaluation', *Journal of Environmental Economics and Management* **34**: 148–172.
Miguel, E. and M. Gugerty (2005), 'Ethnic diversity, social sanctions, and public goods in Kenya', *Journal of Public Economics* **89**: 2325–2368.
Millennium Ecosystem Assessment Board (2005), *Ecosystems and Human Well-Being, Findings of the Condition and Trends Working Group*, vols. I–IV, Washington, DC: Island Press.
Mullainathan, S. (2004), 'Development economics through the lens of psychology', World Bank Working Paper, Report No. 28974.
Mullainathan, S. and R. Thaler (2000), 'Behavioral economics', MIT Department of Economics Working Paper No. 00–27.
Murphy, J. and J.-C. Cardenas (2004), 'An experiment on enforcement strategies for managing a local environment resource', *Journal of Economic Education* **35**: 47–61.
Murphy, J., A. Dinar, R. Howitt, S. Rassenti, and V.L. Smith (2000), 'The design of 'smart' water market institutions using laboratory experiments', *Environmental and Resource Economics* **17**: 375–394.
Neumayer, E. (2005), 'Does high indebtedness increase natural resource exploitation?' *Environment and Development Economics* **10**: 119–125.
O'Connor, D. (1999), 'Applying economic instruments in developing countries: from theory to implementation', *Environment and Development Economics* **4**: 91–110.
Ostrom, E., J. Walker, and R. Gardner (1992), 'Covenants with and without a sword? Self-governance is possible', *American Political Science Review* **86**: 404–417.
Ostrom, E., R. Gardner, and J. Walker (1994), *Rules, Games, and Common-Pool Resources*, Ann Arbor, MI: University of Michigan Press.
Parkhurst, G. and J. Shogren (2007), 'Spatial incentives to protect ecosystem services', *Ecological Economics* **64**: 344–355.
Pearce, D., E. Barbier, and A. Markandya (1990), *Sustainable Development: Economics and Environment in the Third World*, London: Edgar Elgar.
Plott, C. (1983), 'Externalities and corrective policies in experimental markets', *Economic Journal* **93**: 106–127.
Plott, C. (1989), 'An updated review of industrial organization applications of experimental methods', in R. Schmalensee and R. Willig (eds), *Handbook of Industrial Organization*, vol. II, Amsterdam: Elsevier Science Publishers, pp. 1111–1176.
Plott, C. (1994), 'Market architectures, institutional landscapes and testbed experiments', *Economic Theory* **4**: 3–10.
Plott, C. and K. Zeiler (2007), 'Exchange asymmetries incorrectly interpreted as evidence of endowment effect theory and prospect theory?', *American Economic Review* **97**: 1449–1466.
Raiffa, H. (1982), *The Art and Science of Negotiation*, Cambridge, MA: Belknap Press of Harvard University Press.
Rhoads, T. and J. Shogren (2003), 'Regulation through collaboration: final authority and information symmetry in environmental Coasean bargaining', *Journal of Regulatory Economics* **24**: 63–89.
Robinson, E. (2005), 'Reassessing the interaction between investment and tenure uncertainty', *Environment and Development Economics* **10**: 143–157.
Roth, A. (2006), 'Laboratory vs. field experiments: what can we learn?', Comments on Levitt–List paper, Roundtable discussion, ASSA meetings, Boston, available at: http://kuznets.fas.harvard.edu/~aroth/critiques%20of%20experimental%20econ.html

Roth, A., V. Prasnikar, M. Okuno-Fujiwara, and S. Zamir (1991), 'Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: an experimental study', *American Economic Review* **81**: 1068–1095.

Rubinstein, A. (2006), 'Comments on behavioral economics', in R. Blundell, W. K. Newey, and T. Persson (eds), *Advances in Economic Theory*, 2005 World Congress of the Econometric Society, vol. II, Cambridge University Press, pp. 246–254.

Schechter, L. (2007), 'Traditional trust measurement and the risk confound: an experiment in rural Paraguay', *Journal of Economic Behavior and Organization* **62**: 272–292.

Sen, A. (1999), *Development as Freedom*, New York: Alfred A. Knopf.

Shogren, J. (1989), 'Fairness in bargaining needs a context: an experimental examination of loyalty', *Economics Letters* **31**: 319–323.

Shogren, J. (2006a), 'A rule of one', *American Journal of Agricultural Economics* **88**: 1147–1159.

Shogren, J. (2006b), 'Experimental methods and valuation', in K.-G. Mäler and J. Vincent (eds), *Handbook of Environmental Economics*, vol. **II**, Amsterdam: North-Holland, pp. 969–1027.

Shogren, J. and C. Nowell (1992), 'Economics and ecology: a comparison of experimental methodologies and philosophies', *Ecological Economics* **5**: 101–126.

Shogren, J. and L. Taylor (2008), 'On behavioral environmental economics', *Review of Environmental Economics and Policy* (forthcoming)

Smith, V.L. (1982), 'Microeconomic systems as an experimental science', *American Economic Review* **72**: 923–955.

Smith, V.L. (2003), 'Constructivist and ecological rationality in economics', *American Economic Review* **93**: 465–508.

Sturm, B. and J. Weimann (2006), 'Experiments in environmental economics and some close relatives', *Journal of Economic Surveys* **20**: 419–457.

Susskind, L., P. Levy, and J. Thomas-Larmer (2000), *Negotiating Environmental Agreements*, Washington, DC: Island Press.

Townsend, R. (1994), 'Risk and insurance in village india', *Econometrica* **62**: 539–591.

Tracer, D. (2004), 'Market integration, reciprocity, and fairness in rural Papua New Guinea: results from a two-village ultimatum game experiment', in J. Henrich *et al.* (eds), *Foundations of Human Society: Economic Experiments and Ethonographic Evidence from Fifteen Small-scale Societies*, Oxford: Oxford University Press, pp. 232–259.

Vyrastekova, J. and D. van Soest (2008), 'On the (in)effectiveness of rewards in sustaining cooperation', *Experimental Economics* **11**: 53–65.

Walker, J. and R. Gardner (1992), 'Probabilistic destruction of common-pool resources: experimental evidence', *Economic Journal* **102**: 1149–61.

Walker, J.M., R. Gardner, A. Herr, and E. Ostrom (2000), 'Collective choice in the commons: experimental results on proposed allocation rules and votes', *Economic Journal* **110**: 212–234.

Weber, E. and C. Hsee (1998), 'Cross-cultural differences in risk perception, but cross-cultural similarities in attitudes towards perceived risk', *Management Science* **44**: 1205–1217.

Whittington, D. (2002), 'Improving the performance of contingent valuation studies in developing countries', *Environmental and Resource Economics* **22**: 323–367.

Wik, M., T. Aragie Kebede, O. Bergland, and S. Holden (2004), 'On the measurement of risk aversion from experimental data', *Applied Economics* **36**: 2443–2451.

Winters, P., C. Crissman, and P. Espinosa (2004), 'Inducing the adoption of conservation technologies: lessons from the Ecuadorian Andes', *Environment and Development Economics* **9**: 695–720.

Appendix A

Environmental Policy Referendum
Directions: Please read the following statement about an environmental issue. After reading about the issue, a proposed policy will be described. You are then asked to vote for or against the proposed policy.

Issue
Greenhouse gases occur naturally and are part of a natural process that traps heat on the earth's surfaces to warm the planet. Human activities produce greenhouse gases and as economies have industrialized, more and more greenhouse gases have been released into the atmosphere. This has increased the natural level of greenhouse gases to create a global warming effect. This has resulted in a 1 degree Fahrenheit warming trend across the earth's surface since the end of the 19th century. The warming accelerated over the last two decades of the 20th century. Of the different greenhouse gases produced by industry, carbon dioxide (CO_2) is the one that causes the most warming overall. Increased global temperatures from global warming can lead to a melting of polar ice caps, increased desertification of areas closer to the equator, rising sea levels, and increased rainfall in higher latitude countries. A 4.5 degree Fahrenheit increase in temperature by 2100 is estimated to cost global markets 1.5 per cent of their gross domestic product, on average.

Policy Proposal
The United Nations (U.N.) is currently considering a policy that will reduce global emissions of CO_2 to 1990 levels around the world over the next five years. This means some businesses will have to reduce output to meet this goal. Others will have to use new technologies and different fuel sources to lower their CO_2 emissions. An annual tax has been proposed by your government that would require all citizens to pay a tax to help businesses lower CO_2 emissions. The money from this tax will be transferred to businesses to help implement new technologies to reduce their carbon emissions. The tax will cost you $20 per year for the next five years (a total of $100 from the beginning of 2005 to 2010). The alternative to this program is to do nothing and to let CO_2 emissions levels go unchanged or continue to rise. If a vote were being held today and the cost to you is $20 per year for five years, would you vote for the program to reduce global emissions of CO_2 emissions to 1990 levels?

- Yes, I would vote for the proposed policy which would increase my yearly taxes by $20 per year for the next five years.
- No, I would vote against the proposed policy.
- I don't know.