

Social Learning about Environmental Innovations: Experimental Analysis of Adoption Timing

Julian Jamison¹, David Owens² and Glenn Woroch^{3*}

¹*Mind, Behavior, and Development Team, The World Bank, Washington, DC 20552, USA*

²*Department of Economics, Haverford College, 370 Lancaster Ave, Haverford, PA 19041, USA*

³*Department of Economics, University of California, Berkeley, CA 94720, USA, 510-642-4308, woroch@berkeley.edu*

ABSTRACT

We conduct laboratory experiments to investigate how private and public information affect the selection of an environmental innovation and the timing of its adoption. The results reveal behavioral patterns underlying the “energy-efficiency gap” in which consumers and firms delay adoption of cost-effective energy and environmental innovations. Our subjects choose between competing innovations with freedom to select the timing of their adoption, relying on private signals and possibly on observation of their peers’ actions. When deciding whether to make an irreversible choice between a safe and a risky technology, roughly half of subjects delay adoption beyond the time prescribed by equilibrium behavior — pointing to a possible behavioral anomaly. When they do adopt, subjects give proportionately more weight to their private signals than to their peers’ actions, implying that they do not ‘herd’ on the actions of

*Financial support for this research was provided by France Telecom Research & Development, LLC. We thank the participants of the conferences at: the French Economic Association (Lyon, France), the Florida State Conference on Experimental

their peers. Nevertheless, when subjects observe their peers' decisions, they accelerate the timing of their adoptions, but do not necessarily imitate their peers. This occurs even when payoffs are statistically independent as though observing prior adoptions exerts "peer pressure" on the subjects to act. The experimental results suggest that rapid dissemination of information of peer actions can speed up diffusion of innovations that save energy and protect the environment, and improve selection from among competing technologies.

Keywords: Social learning; herding; endogenous timing; behavioral economic policy; diffusion of technology; environmental innovations

JEL Codes: Q55, C91, D83, O33

Introduction

We conduct a series of laboratory experiments that are designed to gain a deeper understanding of the timing of adoption of innovations that save energy or protect the environment relative to the *status quo*. In particular, the experiments attempt to isolate how private and social information affect the pace of diffusion of superior energy and environmental technologies.

The results promise to shed light on an empirical regularity known as the "energy efficiency gap" (EEG).¹ The EEG takes different forms, but the most common example occurs when a consumer prefers a new energy-using appliance or vehicle that has a lower purchase price but higher operating cost, making it more costly than an efficient alternative over its expected lifetime. Many empirical studies have estimated the

Game Theory, the Industrial Organization Workshop (Lecce, Italy), CEDEX (Nottingham, UK), and the European School of Management and Technology (Berlin, Germany). Lastly, we are grateful to Anita Sonawane, Ryan Fackler and Tasmia Rahman for their research assistance.

¹See, *e.g.*, Allcott and Greenstone (2012) and Jaffe and Stavins (1994a).

implicit discount rates supporting consumer decisions finding they are often much higher than any reasonable market discount rate.²

Another dimension of the EEG is the timing of these adoption decisions.³ Besides the selection from among competing energy-using (or pollution-generating) technologies, the timing when consumers make their adoption decisions is also important. While a consumer may adopt the efficient technology in the end, significant time may elapse before this happens.⁴ The timing of adoption of environmental innovations is an aspect of EEG that has been explored far less in the field and the laboratory.⁵

Our experiments are designed to replicate many of the key features of the decision problem underlying the EEG. First, subjects in our treatments choose one of two innovations, with the default technology being an inferior *status quo*. The two innovations consist of a “safe” technology that has a higher payoff than the *status quo* with certainty, and a “risky” technology that may pay off more than the *status quo* and the safe technologies, or it may be worse than both of these alternatives. This menu of options is designed to capture the common choice between a mature, conventional technology and a new one with which decision makers personally have little or no experience.

An example of this type of decision confronts consumers when they consider purchasing a new car: they can keep their conventional gasoline-powered vehicle, choose one using gas–electric hybrid technology which now has well-established performance record, or take a chance on an unproven technology such as a vehicle powered by a hydrogen fuel cell. A similar auto purchase problem was considered by Gaker *et al.* (2010) in a laboratory experiment. Another illustration is the choice of light bulb type, with the traditional incandescent bulb being the *status quo*

²Stadelmann (2017) (Table 1) reports a range of implicit discount rates that were estimated for each of seven categories of consumer durables.

³See Gerarden *et al.* (2015) for a simple formulation of the static consumer purchase problem.

⁴Van Soest (2005) models the adoption delay that occurs before firms install a new energy-efficient technology, and predicts the impact on this delay of different policy interventions.

⁵Stadelmann (2017) singles out the timing of appliance replacement as an important decision that has not received sufficient attention.

technology, a compact florescent bulb the safe innovation, and an LED bulb the risky innovation in terms of payback.⁶

Second, unlike the vast majority of laboratory tests of observational learning, our experiments elicit the timing of consumer adoption decisions. Our subjects are free to choose which of the eight rounds they select a technology. When they do select a technology, their decision is irreversible.

The timing of adoption of environmental innovations is a question that has not often been addressed in the literature.⁷ In the vehicle purchase example, for instance, the consumer may prefer to wait until more information becomes available about the hydrogen car. In our treatments, the cost of delay arises naturally because postponing adoption amounts to remaining with the *status quo* and earning a lower return than the safe option. The cost of delay was made clear to subjects. In each round of the experiments, subjects who remain undecided about which innovation to select receive a private signal that is correlated with the true return of the risky innovation. In this way we maintain, *ex ante*, parity in the amount of private and public information when subjects choose to delay their adoption decisions.

Finally, the central objective of our lab experiments is to test the role that information plays in adoption timing, and especially the observation of the actions of peers. As we will discuss below, field studies often find that a social observation encourages adoption by nonadopters.

Our three experimental treatments are designed to assess how subjects balance private and public information when making their adoption decisions. The first treatment serves as a baseline in which subjects do not observe peers' actions. In the second treatment, subjects observe prior decisions made by members of their group but individual payoffs are statistically independent and, hence, observation conveys no useful information. The third treatment also allows for observation of prior actions, but those actions now are informative since the payoff of the risky innovation is perfectly correlated across members of each group.

⁶See Allcott and Taubinsky (2015).

⁷Van Soest (2005) models the adoption delay that occurs before firms install a new energy-efficient technology, and predicts the impact on this delay of different policy interventions. Stadelmann (2017) singles out the timing of appliance replacement as an important decision that has not received sufficient attention.

To generate hypotheses about adoption timing, we attempt to characterize Bayesian Nash Equilibrium (“BNE”) strategies in the non-cooperative games underlying the three treatments. We find that a risk-neutral expected-payoff maximizer in the first two treatments should adopt at her first opportunity, choosing the innovation that is favored by her first private signal. BNE is less conclusive for the third, ‘common-values’ treatment. We show in that case that immediate adoption by all players is not an equilibrium, as the opportunity to delay and observe others’ adoption decisions is too valuable.

The results from our experiments confirm that subjects tend to delay adoption of either one of the superior technologies — contrary to theoretical predictions of optimizing behavior but consistent with the energy-efficiency gap. This adoption pattern is robust across treatments and suggests the presence of a behavioral anomaly. For instance, subjects may display a “*status quo bias*” in which they persist with the default technology despite the availability of a superior alternative.

One plausible explanation for this delay is that subjects seek to take advantage of social learning. In fact, the experimental results confirm that, on average, the speed and accuracy of adoption improve when subjects can observe their peers’ actions. While this result is not surprising when observation conveys payoff-relevant information, it is unexpected that mere observation also promotes faster, more profitable decisions even when the true underlying values are statistically independent. It appears that the opportunity to observe others’ decisions exerts “peer pressure” on subjects, driving them to reach their own decisions more quickly.

Turning to the accuracy of adoption decisions, we find that subjects do better than randomizing between the two innovations, *i.e.*, they pay attention to private and/or public information. On average, subjects show a slight preference for the safe innovation over the risky one, consistent with risk aversion. We do find that when payoffs are correlated delays beyond the first round are rewarded with improved accuracy in the selection of an innovation.

While the lab results show our subjects take into account both kinds of information, their actions confirm that they place greater weight on their private signals. To see this, we classify an innovation as “popular” when members of the peer group choose it more often than the other

innovation. We classify as “leading” the innovation that is favored by the history of a subject’s private signals. The experiments confirm that subjects invariably adopt the innovation that is both the leading and the popular one. However, if one innovation is popular but the other is leading, then the leading innovation is chosen more often; again, subjects tend to follow their private signals — a pattern that occurred in both experimental treatments having public observation.

We draw out some of the implications of our experimental results for policies aimed at accelerating the adoption of cost-effective innovations to save energy and preserve the environment. Indeed, many policy recommendations have been put forward with the aim of closing the energy-efficiency gap. Among the policy proposals that have been put forward, some employ the standard economic instruments — such as taxes and subsidies — aimed at erecting monetary incentives to spur firms and consumers to adopt cost-effective technologies. Other proposals, however, target the behavioral anomalies which are seen to be impediments to adoption. Some proposals of this type are known as social “nudges” (Olander and Thøgersen, 2014). Rather than adjusting financial rewards, these proposals attempt to tailor the information on which adoption decisions are made. A good example of a nudge in our setting is to provide the decision maker with feedback about the history of their personal payoffs and information, or about prior decisions made by their peers. Our results suggest that information that is statistically irrelevant can have significant effect on the speed and quality of adoption decisions, and that policy makers can exploit this fact with little additional administrative cost.

The next section positions our paper in the literature on theoretical and laboratory experiments involving technology adoption with social learning, especially as it applies to adoption of energy and environmental innovations. Section “Experimental Design and Procedures” then describes our laboratory protocol and the three experimental treatments. Section “Theory” analyzes equilibrium in the game that underlies each treatment. We report the experimental results in section “Experimental Results,” comparing the treatments against one another and against our theoretical predictions. A final section summarizes our conclusions and suggests implications of our results for policy interventions that may work to accelerate adoption of innovations that save energy and protect the environment.

Related Literature

Theoretical Background

We contribute to the large body of behavioral research on social learning games and experiments. As with many related papers, our methodology has its roots in the pioneering work of Banerjee (1992) and Bikchandani *et al.* (1992) (“BBHW” for short). In these papers, agents make one-time binary investment decisions in a pre-determined order, each one informed by a single private signal plus the observation of all prior choices. While this structure generates valuable insights into adoption behavior (e.g., herding), the assumption of exogenously determined timing of decisions fails to represent many real-world decision problems, e.g., whether to invest in solar photovoltaic (SPV) panels or to buy an electric vehicle.

Our experimental treatments preserve much of the structure of the BBHW game in that subjects can choose either a safe or a risky innovation, and they may be able to observe the actions of their peers. Now, however, we allow agents to choose when to make their adoption decisions. While subjects start out with the same opportunities to acquire information, the amount of private and social learning that is possible depends on the timing of their decisions.

Following BBHW, a series of papers appeared that relaxed the assumption that agents made their decisions in a pre-determined order. In Chamley and Gale (1994) and Chamley (2004) agents choose when to make a binary investment decision in a discrete-time model. They find that the delays in adoption are directly related to the length of a period and inversely related to the number of players.

Zhang (1997) also considers a binary investment game with endogenous timing but with continuous time. In equilibrium, agents initially delay their actions in order to learn from others, and then invest rapidly thereafter. More recently, Levin and Peck (2008) construct an endogenous timing game that adds a second signal about the common value of an investment’s return, plus a signal about the agent’s cost of investing. Murto and Välimäki (2011) conduct a similar study with an irreversible investment in an endogenous, discrete-timing game involving both private and observational information. They find that observational learning delays decisions in equilibrium, a conclusion consistent with the energy-efficiency gap and with our experimental results.

Energy and Environmental Technology Adoption Literature

Despite the critical energy and environmental issues involved, and despite evidence confirming an energy-efficiency gap, relatively little empirical research has been conducted on the timing of this kind of adoption.

Several observational studies have looked at the determinants of consumers' decisions to install SPV panels to supplement or replace electricity provided by the grid. Bollinger and Gillingham (2012) and Graziano and Gillingham (2015) documented the stimulative effect of neighbor adoption on the diffusion of SPV panels among California and Connecticut residents, respectively. For instance, Bollinger and Gillingham (2012) found that installation of each additional SPV increased the likelihood of SPV adoption in the respective zip codes by 0.78 percentage point. Learning about energy-efficient options has also been documented in laboratory settings. For example, Gaker *et al.* (2010) found that experimental subjects were more inclined to select conventional or hybrid vehicles when a higher fraction of their peers did likewise.

Similar trends are observed in the commercial context of environmental technology adoption. For instance, Covert (2015) studies how oil companies use different versions of hydraulic fracturing technology over time, using detailed data from the Bakken Shale in North Dakota. State disclosure laws made it possible for firms to analyze competitors' data as well as their own, but he finds that even these sophisticated and high-stakes decision-makers place higher weight on their own experience than on that of others. Overall they appear to learn somewhat slowly, and in particular they do not often experiment in the sense of choosing high-risk, high-return approaches.

A number of field experiments have explored the role of private and social information on residential electricity consumption. Fischer (2008), Ayres *et al.* (2013), Costa and Kahn (2013), and Jessoe and Rapson (2014) examine how information provision can induce conservation. Collectively these studies confirm the presence of behavioral anomalies among this important class of consumers, and also the effectiveness of information dissemination to counteract the effects of those anomalies. Comparatively little has been done in the way of laboratory experimentation examining these same decisions, however. Our paper attempts to fill this void with respect to the energy-efficiency gap.

Experimental Tests

Several papers have tested models of herding behavior in the laboratory. Among the first was Anderson and Holt (1997) who, like many who followed, attempted to replicate the decision problem described by BBHW. They find evidence of the predicted uniformity of behavior, including selection of the “wrong” investment, consistent with Bayesian updating. As will become clear, our experiments extend Anderson and Holt’s 1997 stochastic structure, allowing decision timing to be endogenous. Çelen and Kariv (2004) were the first to distinguish between herd behavior and information cascades in a laboratory setting.

A few early laboratory experiments did allow for endogenous timing but in a strategic setting that differed from BBHW, e.g., Baik *et al.* (1999). Sgroi (2003) was the first to adapt the BBHW structure to permit subjects to select the timing of their actions. The results of these experiments showed that, while herding behavior still occurred, subjects tended to delay their adoption in an apparent attempt to learn from their peers’ private signals.

More recently, Ivanov *et al.* (2013) implemented laboratory experiments of the two-period, two-signal game structure developed in Levin and Peck (2008). They are especially interested in the implications of group size as it alters the information content contained in the observation of peers. They find that subjects do not take full advantage of the information supplied by others’ actions. Unlike them, our treatments are restricted to a fixed group size (six) but allow for more than two periods (eight). We do find, like Ivanov *et al.* (2013) and others, evidence that subjects attach a disproportionately large weight to private information compared to public observation — a pattern that appears in many experiments using the BBHW model with exogenous timing.

Weizsäcker (2010) merges 13 experimental datasets, each of which implemented the BBHW framework, and finds that subjects deviate from choice that would be optimal leading up to a decision. In particular, he finds that in a majority of cases subjects follow their private signal when an alternative would earn them higher payoffs. We similarly assess subjects earnings, and find that they learn too little from observation of others.

Our study employs discrete time with a finite horizon, where a finite number of agents have a choice between two innovations and a third

option, the *status quo*, which delays the choice. The difference in payoffs between the safe innovation and the *status quo* option determines the cost of delay. The underlying tension, as in actual decisions such as pollution reduction, involves incremental improvements in standard technologies versus waiting longer to learn more about the potential costs and benefits of an attractive, but as of yet unproven, alternative green technology.

Experimental Design and Procedures

All of our experiments were conducted using the Experimental Social Science Laboratory, or “Xlab,” at the University of California, Berkeley. The Xlab primarily recruits undergraduate students as subjects, supplemented with graduate students and university staff members.

The experiments themselves were conducted on laptop personal computers using the Zurich Toolbox for Readymade Economic Experiments (“z-Tree”) program (Fischbacher, 2007). Workstations were separated from one another with physical partitions to prevent subjects from looking at one another’s screens and to discourage conversation. Apart from the introductory period when subjects could ask clarifying questions, no oral communication was allowed.

Subjects participated in just one experimental session and, hence, in just one of the three treatments. In two of the treatments, subjects were randomly assigned to groups of six. In those cases, members of a group observed actions of others through the z-Tree interface, remaining unaware of the identity of their peers. Groups were randomly reassigned for each of the 15 decision problems, making it difficult for subjects to learn about behavior specific to any individual.

Subjects were given monetary incentives for their participation in the experiments. Subjects earned “tokens,” which were exchanged for dollars at a rate announced at the beginning of each experimental session. When a session ended, tokens were added up for each subject and checks were issued to subjects in sealed envelopes. The average payout per subject per session was \$23.33, and the typical session lasted about one hour.

Our sessions included a total of 227 subjects in 13 experimental sessions, for an average of 17.5 subjects per session. At the beginning of

each session, subjects received written instructions and 10 minutes to read them, after which the instructions were read aloud by a researcher. Before beginning the experiment, subjects took a simple quiz that tested their understanding of the instructions and were not allowed to proceed until they answered each question correctly.⁸ Each session consisted of 16 repetitions of a decision problem. The first decision problem was unpaid practice and the results were discarded, leaving data for 15 decision problems per subject.

A copy of the instructions that subjects received for one of the experimental sessions is reproduced in Appendix A. Note that while we use the language of ‘technology adoption’ and ‘innovation’ in this paper to describe the decision problems facing subjects, neutral language was used in the verbal and written instructions during the experimental sessions.

Payoffs, Probabilities, and Information

In each decision problem, subjects choose between three technologies: *A*, *B*, and *C*. Subjects begin each decision problem with technology *A* by default, and remain with *A* until they make the irreversible decision to adopt “innovation” *B* or *C*. *A* is inferior in expected payoff to *B* and *C*, but remaining with *A* allows subjects to postpone their adoption decision at least one more period. Thus, *A* can be interpreted as an inferior *status quo* technology that an agent may maintain while she accumulates information on the relative merits of the two innovations.

Each decision problem consists of an eight-round “decision sequence.” Each technology has a per-round payoff that accrues to a subject who has chosen that technology. The *status quo* technology *A* pays 1 token per round and the innovation *B* pays 2 tokens per round. The payoff to innovation *C* is stochastic, paying 0 or 4 tokens per round, each with equal likelihood. The per-round value of *C*, once realized, remains constant for the entire decision sequence. Subjects can earn the most if they choose innovation *C* when it takes a high value (4), and innovation *B* when *C* takes a low value (0). The earlier in the decision sequence

⁸The quiz was introduced after the first three sessions. No difference in experimental outcomes was found before and after adding the quiz.

that a subject makes this choice, the greater the total number of tokens she earns.

Subjects receive an informative but noisy private signal about the realized value of innovation C before they make their decision for round 1. Each subject who chooses technology A in round 1, and thus declines to adopt innovation B or C , remains “active,” and will receive further information about innovation C in subsequent rounds. Thus, innovation B can be interpreted as a guaranteed improvement over *status quo* technology A , while C is a risky innovation. It may be superior to both A and B , but it may also be worthless.

Active subjects — those who have yet to adopt innovation B or C — begin each round of a decision problem by observing a private signal. Draws of the signal are independent and identically distributed across rounds. They are informative of the true value of innovation C . If the value of C is 0, the subject receives a signal of “ H ” with a probability of $\frac{1}{3}$ and “ L ” with a probability of $\frac{2}{3}$. If innovation C ’s value is 4, then each signal is “ H ” with probability $\frac{2}{3}$ and “ L ” with probability $\frac{1}{3}$. Applying Bayes Rule, the conditional probability that C pays 0 and 4 tokens, given that the “ L ” signal is observed is $\frac{2}{3}$ and $\frac{1}{3}$, respectively, absent any other information.

By design, subjects incur an opportunity cost of delay equal to one token per round since postponing adoption to the next period generates 1 token per round under *status quo* A compared to 2 tokens per round if they chose innovation B instead. The benefit of delay is the possibility to make a “better” choice between B and C after drawing an additional private signal and, depending on the treatment, observing another round of decisions by peers. The cost of delay is captured in the opportunity cost as opposed to imposing an arbitrary discount factor on subjects’ payoff streams as is done by other experimental studies. It is possible that there is a countervailing nonmonetary temptation to delay, simply to forestall boredom as raised in Lei *et al.* (2001). This is not specific to lab settings, however. Further, we would expect this effect to be approximately symmetric across our treatments.

The Three Treatments

Treatments differ in two dimensions: the opportunity to observe the actions of other subjects, and the correlation of payoffs across subjects

Table 1: Characteristics of the treatments.

Treatment	Payoffs	Observation of peers
<i>P</i>	Private	No
<i>PO</i>	Private	Yes
<i>CO</i>	Common	Yes

in a group. The observability of actions and commonality of payoffs were nested across treatments to identify the influence of each characteristic on adoption decisions. The three treatments are summarized in Table 1. In the baseline Treatment *P* subjects only observe their private signals of the payoff to *C* which is independently drawn for that subject. As there is no social interaction in this treatment, there can be no social learning.

In all three treatments, all subjects draw at random a private signal of the payoff to innovation *C*. A “P” indicates that only private signals inform the subjects. In some treatments, beginning in the second round and continuing through the eighth and final rounds, active subjects observe the choices made by other members of their group in the previous round prior to making their own decision. When observation of others’ actions is allowed, we label the treatment with an “O.” This feature allows for the possibility of social learning as subjects can learn about their peers’ private signals by observing the resulting actions.

The second characteristic of treatments is whether the payoffs to innovation *C* are independent across subjects or they are perfectly correlated among members of the same group. A “C” stands for common value (or perfect correlation) of these payoffs and we refer to such a treatment as “common value.” When the *C* payoffs are statistically independent across group members, we refer to this as the case of “private values” but do not use separate label for the treatments. The experimental instructions inform subjects whether they are participating in a private or common value treatment, and subjects must demonstrate their understanding of the experiment by achieving a perfect score on a short quiz before proceeding.

For instance, Treatment *PO* allows subjects to observe the decisions of each member of their six-person peer group, but these decisions have

no statistical bearing on their own decisions. Thus, any differences in behavior between Treatments *P* and *PO* cannot be explained by Bayesian updating. On the other hand, Treatment *CO* represents the case when there is an informational externality. In that treatment social learning is possible because subjects can update their beliefs about the payoff of innovation *C* by observing the choices of other members of the group.

Theory

This section characterizes optimal behavior for risk-neutral expected-utility maximizers in each treatment. Such subjects will select the innovation *and* the timing of the adoption that maximizes their expected payoff. Let *h* and *l* represent the number of private draws of *H* and *L*, respectively, seen by a subject through round *t*, and N_j^{t-1} the number of subjects that have adopted technology *j* as of round *t* - 1.⁹ $V_j(h, t, N_B^{t-1}, N_C^{t-1})$ denotes the per-round expected value of technology $j \in \{A, B, C\}$ earned by a subject in round *t* based on all available information.¹⁰

The *total* expected profit from technology *j* over the remainder of the decision problem if technology *j* is chosen in round *t* is $\pi_j(h, t) = V_j(h, t) \times (9 - t)$. Finally, let $X^*(h, t)$ be a subject's optimal choice in round *t* given *h*, that is, $j \in X^*(h, t)$ if $\pi_j \geq \pi_k$ for all $k \in \{A, B, C\}$.

For Treatments *P* and *PO*, the solution is an optimal decision rule that maps their information sets into a probability distribution over the three technologies. In Treatment *CO*, subjects' optimal actions may also depend on the adoption decisions of others, rendering subjects' optimal actions interdependent. As a consequence, we must look for equilibrium noncooperative strategies.

Treatments *P* and *PO*

As treatments *P* and *PO* are theoretically equivalent, we can characterize optimal decisions for both simultaneously. As decisions are

⁹Subjects observe the N_j^{t-1} other subjects had chosen technology $j \in \{A, B, C\}$ before making their own choice in round *t*.

¹⁰For example, $V_A(h, t, N_B^{t-1}, N_C^{t-1}) = 1$ and $V_B(h, t, N_B^{t-1}, N_C^{t-1}) = 2 \vee (h, t, N_B^{t-1}, N_C^{t-1})$.

independent, V_C depends only on h and l . Using Bayes' Rule, the likelihood that the payoff to innovation C equals 4 is: $P_4(h, t) = \frac{2^h}{2^h + 2^{t-h}} = \frac{2^{h-l}}{2^{h-l} + 1}$. Importantly, $P_4(h, t)$, and thus $V_C(h, t)$, depends only on the difference $(h - l)$. $\pi_C(h, t) = V_C(h, t) \times (9 - t)$ and $\pi_B(h, t) = 2 \times (9 - t)$. The computation of $\pi_A(h, t)$ is more complex, requiring the computation of $\pi_{X^*}(h, t + 1)$ and $\pi_{X^*}(h + 1, t + 1)$:

$$\begin{aligned} \pi_A(h, t) &= 1 + [P_H \times \pi_{X^*}(h + 1, t + 1) + (1 - P_H) \\ &\quad \times \pi_{X^*}(h, t + 1)] \times (8 - t), \end{aligned} \tag{1}$$

where $P_H(h, t)$ represents the probability that the next private signal will be H :

$$P_H(h, t) = \frac{2}{3} \times P_4 + \frac{1}{3} \times P_0. \tag{2}$$

As the opportunity to delay adoption complicates early decisions, we begin our analysis at the end with Round 8, where a profit-maximizer need not consider remaining with A . Subjects remaining in Round 8 should adopt C for $h > 4$, B for $h < 4$, and either for $h = 4$. Because delay is costly, this also reveals that profit-maximizing subjects should adopt C if $h \geq 4$ and B if $l \geq 4$ in *any* round. As a consequence, such subjects should adopt immediately for all (h, l) in Round 7.

$(h = 3, t = 6)$ is the only circumstance in the last three rounds where we must consider the possibility of retaining technology A . However, $\pi_B(3, 6) = \pi_C(3, 6) = 2 \times (9 - 6) = 6$, while $\pi_A = 1 + \frac{1}{2}\pi_C(4, 7) + \frac{1}{2} \times \pi_B(3, 7) = 1 + \frac{8-6}{2} [\frac{2}{3} \times 4 + 2] = \frac{17}{3} < 6$, so the additional information is not worth waiting for, and thus $X^*(3, 6) = \{B, C\}$, ruling out technology A when $h \geq 3$ or $l \geq 3$ for any t , etc. The end result of this backwards-induction process is displayed in Table 2.

Table 2 characterizes the entire decision space.¹¹ $C \in X^*(h, t)$ if $h \geq l$ and $B \in X^*(h, t)$ if $h < l$, except for $(h, t) = (1, 2)$ or $(2, 4)$, where $X^*(h, t) = A$. Subjects get the highest payoff from adopting the innovation suggested by their private information (C if $h > l$, B if $h < l$) in any round of the experiment. When private information suggests indifference ($h = l$) early in the experimental decision sequence (Rounds 2 and 4), subjects are best served by delaying, and adopting according to their next signal. By Round 6, however, not enough rounds remain in

¹¹See the Appendix for the calculations behind the predictions in Table 2.

Table 2: Optimal decisions for treatments P and PO .

t (Round)	$h : A \in X^*(h, t)$	$h : B \in X^*(h, t)$	$h : C \in X^*(h, t)$
8		0, 1, 2, 3, 4	4, 5, 6, 7, 8
7		0, 1, 2, 3	4, 5, 6, 7
6		0, 1, 2, 3	3, 4, 5, 6
5		0, 1, 2	3, 4, 5
4	2	0, 1	3, 4
3		0, 1	2, 3
2	1	0	2
1		0	1

which subjects could reap the benefits of the newly acquired information, and adopting either innovation B or C immediately maximizes expected payoffs when $h = l = 3$.

Importantly, backwards-induction yields a unique Bayesian Nash equilibrium (“BNE”) of ‘immediate adoption’ for Treatments P and PO , in which subjects adopt innovation B following a private signal of L and innovation C following a private signal of H in Round 1. The *ex ante* expected payoff of this strategy is $(0.5 \times 2 + 0.5 \times \frac{2}{3} \times 4) \times (9 - 1) = \frac{56}{3} \approx 18.67$. A complete characterization of the optimal decision rule calls for X^* to include B if $h \leq l$ and C if $h \geq l$, except $X^*(1, 2) = X^*(2, 4) = A$.

Treatment CO

The characterization of optimal behavior in Treatment CO is complicated by observational learning. Delaying adoption is potentially *more* attractive in this treatment because delay reveals not only private signals, but also the information of others. Subjects adopting in earlier rounds effectively confer an informational externality on later adopters.

We begin by testing whether immediate adoption can occur in BNE, as it does in Treatments P and PO . Following any such strategy again yields $\pi_C(1, 1) \approx 21.33$ and $\pi_B(0, 1) = 16$, while deviating and choosing A effectively reveals six additional private-signal equivalents. The payoff to this deviation is more complicated than Equation (1), as the six additional pieces of information yield seven information sets in the next

round, rather than two. Equation (3) calculates the distribution of H signals among the additional private-signal equivalents, D_t ,¹² under the maintained assumption that all subjects adopt in round t .

$$Pr(h_a = i) = \frac{D_t!}{i! \times (D_t - i)!} \times P_H^i \times (1 - P_H)^{(D_t - i)} \quad (3)$$

Equation (4) uses the probability distribution derived from Equation (3) to calculate the expected profit of choosing technology A, where the optimal choice in the next round, $X^*(t + 1) \in \{A, B, C\}$ depends on the number of effective H draws:

$$\pi_A(h, t) = 1 + \sum_{i=0}^{D_t} Pr(h_a = i) \times \pi_{X^*}(t + 1, h + i) \quad (4)$$

In the ‘immediate adoption’ strategy profile, deviating by choosing technology A in Round 1 yields one more private signal and the decision of five colleagues, or $D_1 = 6$. For example, following a private draw of L in round $t = 1$, a subject adopts technology B at $t = 2$ for $h_a \leq 3$, and technology C for $h > 3$:

$$\begin{aligned} \pi_A(0, 1) &= 1 + \left(2 \times \sum_{i=0}^3 Pr(h_a = i) + 4 \times \sum_{i=4}^6 Pr(h_a = i) \times \frac{2^i}{2^i + 2^{7-i}} \right) \\ &\quad \times 7 \approx 16.57, \\ \pi_B(0, 1) &= 2 \times 8 = 16 \end{aligned} \quad (5)$$

As $\pi_A(0, 1) > \pi_B(0, 1)$ according to this strategy profile, ‘honest’ adoption in the first round cannot occur in equilibrium in Treatment CO . Further, for the same strategy profile, $21.73 = \pi_A(1, 1) > \pi_C(1, 1) = 21.33$. Thus, in Treatment CO , there can be no equilibrium in which all subjects adopt with certainty in the first round. The information revealed in the decisions of others would be too valuable for subjects to profitably follow such a strategy. Note that this argument does indeed show that there is no BNE (not simply perfect Bayesian equilibrium, or ‘PBE’) in which all agents adopt immediately, because in that case a

¹² $D_t = 7 - N_B^{t-1} - N_B^{t-1}$, equivalent to the $6 - N_B^{t-1} - N_B^{t-1}$ colleagues that are yet to adopt, plus one for the additional private signal that will accompany the selection of A .

unilateral deviation would be profitable. In any symmetric equilibrium, therefore, all subjects must delay adoption with at least some positive probability.

Implications

We do not expect our experimental subjects to play exactly according to the formal equilibrium strategies. This is due to noise (and other errors), ‘behavioral’ factors (discussed further later), and modeling simplifications such as risk neutrality. Nevertheless, we can draw a few robust qualitative conclusions from the theory in order to help guide the interpretation of the results.

For Treatments *P* and *PO*, the first implication is that average behavior should be the same in both cases. Any divergence must be due to either misunderstanding the rules (which we tried to avoid to the extent possible); to some nonstandard component of utility; or something like ‘meta-learning’ about how to play the game. The second robust implication is that players should adopt one of the new innovations very quickly. Note for instance that any positive degree of risk-aversion will break indifference in expected payoff values: at $t = 2$, if $h = l = 1$, there will be a strict preference for *B* over *C* since it is less uncertain.

For Treatment *CO* the main implication is that, *relative* to Treatments *P* and *PO*, delay should occur and the average time of adoption will be later. This follows whatever the nature of the exact mixed equilibrium in the former case, and it is robust to the addition of risk aversion. Indeed, other than the threshold case mentioned in the previous paragraph, risk aversion will tend to delay adoption since by waiting more uncertainty is resolved. However, this effect is the same across treatments and does not change the relative prediction of later adoption. Finally, as subjects can always choose to simply ignore the observed behavior of others (and therefore optimally adopt immediately), payoffs should be weakly larger in Treatment *CO* as compared to the other two.

Experimental Results

Our laboratory experiments were designed to explore how subjects balance the competing goals of speed and accuracy, and how that trade-off depends on their private signals and observations of peers’

adoption decisions. This section assesses the data generated by the experiments using several empirical techniques.

Looking first at the choice between the two innovations, we say that a subject is *accurate* if she adopts the innovation that would deliver the highest *ex post* payoff in the decision sequence. Specifically, innovation *B* is the accurate choice in a decision sequence when $\pi_C = 0$, while innovation *C* is accurate when $\pi_C = 4$. Note that at any round of the decision sequence, the private signals drawn by a subject may nonetheless point to the innovation with the lower payoff. Turning next to adoption timing, the *speed* of decision making is measured by the number of rounds that a subject stays with *A* before choosing either *B* or *C*.

We begin by reporting simple summary statistics to compare the experimental treatments. We then employ the statistical techniques of *survival analysis* to analyze the timing of adoptions. Initially, we make non-parametric comparisons of adoption timing of pairs of treatments. Next, we estimate Cox proportional hazard models, which relate adoption timing to the history of a subject's private signals and, when observable, the prior adoptions made by her peers. The estimated coefficients provide a means to quantify the differences among the various treatments as well as to assess the influence of private and public information. To allow distinct patterns of adoption for the safe and risky innovations, we estimate a *competing risks* model and relate the timing of adoption of *B* and *C* to the time-varying covariates.

Finally, the complexity of the decision problems stemming from different sources of information led us to consider whether behavioral rules could better explain the experimental data than BNE strategies. We propose nine behavioral rules of thumb that map private and public information into adoption decisions. Each of the rules is a reasonable response to the available information, but they are not necessarily optimal. Based on subjects' adoption decisions, we attribute one of the nine behavioral rules to each subject. After calculating the average payoff corresponding to each rule, we rank order the rules in terms of average payoff.

Accuracy of Decisions

Table 3 provides an initial assessment of the accuracy of adoption decisions by comparing summary statistics across the three treatments.

Table 3: Summary statistics of decision accuracy.

Treatment	Obs.	Adopted B (%)	Adopted C (%)	Accurate adoption (%)	Avg. round of adoption
<i>P</i>	525	51.8	48.2	59.4	2.17 (0.20)
<i>PO</i>	900	56.0	44.0	64.7	2.06 (0.10)
<i>CO</i>	720	52.6	47.4	61.5	1.91 (0.10)

Robust standard errors in parentheses, clustered by subject.

Each decision sequence amounts to a single observation, and so there are 15 observations per subject.

Participants show a slight preference for the safe innovation *B*, as if subjects were on average averse to the risk of innovation *C*.¹³ Adoption is significantly less accurate in Treatment *P* ($p = 0.023$), suggesting that the ability to observe peers and the correlation of payoffs in Treatments *PO* and *CO*, respectively, may guide subjects toward the superior innovation. More specifically, examining the contingency tables relating subjects' choice between *B* and *C* conditional on the true value of *C*, we found they were particularly accurate in identifying case of $C = 0$ and choosing *B* in both of those two treatments.

As evidenced by the average number of rounds before an innovation was chosen, subjects adopted more quickly in the common value treatment, *i.e.*, *CO*.¹⁴ At first glance, it appears that the observation of peers results in faster and more accurate adoptions, whether values are perfectly correlated or statistically independent. We now turn to a more formal analysis that takes account of difference in adoption timing over all rounds of the decision sequence.

¹³Whether this is due to underlying risk aversion or to a change in expressed norms as a function of the varying contexts across treatments, as suggested by List (2007) and Bardsley (2008), is difficult to separate empirically. The policy implications are similar in either case.

¹⁴The difference is significant at the 1% level when the standard errors are not clustered. When they are clustered, the significance level is better than $p = 0.15$ for one-sided tests relative to Treatments *P* and *PO*.

Survival Analysis of Adoption Timing

We want to be able to compare statistically the time pattern of adoption in our experimental data, and for this purpose we employ the statistical technique of “survival analysis” or what is also called “duration analysis.” These techniques are specifically designed to model the timing of an “event,” which in our context is the adoption of an energy-environmental innovation. In economics survival analysis most often appears in the field of labor economics.¹⁵ It is perfectly suited, for example, to modeling the duration of unemployment spells and their determinants. Survival analysis was used to statistically model the timing of adoption of energy-efficient building codes by Nelson (2012) and deployment of solar panels by Islam and Meade (2013).

Since survival analysis was designed to perform statistical analysis on the *timing of an event*, it is well-suited to analyzing the adoption times generated by our experiments.¹⁶ In our case, the event is the adoption of an innovation, and so the dependent variable becomes the round in which subjects adopt either innovation *B* or *C*.¹⁷

We begin by plotting the empirical hazard functions observed for each treatment. Figure 1 overlays the Kaplan–Meier non-parametric hazard functions for all three treatments after they have been smoothed for easier comparison. A rising hazard function indicates that the propensity to adopt one of the two innovations increases from one round to the next, while a falling hazard indicates a decreasing propensity.¹⁸

It is clear that the baseline Treatment *P* results in uniformly slower adoption than the other two treatments. The deviation between Treatments *P* and *PO* is especially surprising. By design, observation of others’ actions in Treatment *PO* conveys no information about subjects’ expected payoffs; all relevant information is contained in their individual

¹⁵See, e.g., Heckman and Singer (1984).

¹⁶To our knowledge, this is the first application of survival analysis to the results of endogenous timing experiments.

¹⁷Event data are often censored because the experiment ends or subjects leave the study. With negligible exceptions (*i.e.*, 8 out of 2,145 decision sequences), our subjects made an adoption sometime during the eight-round decision sequence.

¹⁸In other words, the longer the subject delays making a decision, the less likely she will adopt in the next period. In labor economics such “duration dependence” arises when the longer a person is unemployed, the less likely they will re-enter the workforce in the near future.

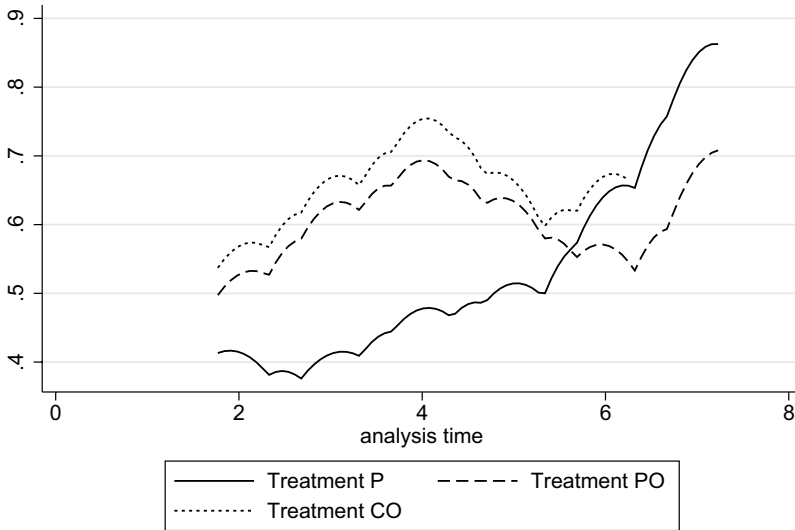


Figure 1: Smoothed hazard function by treatment.

private signals. The fact that its hazard rate increases steadily over rounds indicates that postponing adoption leads to increased propensity to adopt. It appears that the observation of actions of other members of the group prompts subjects to act when they would otherwise procrastinate.

Cox Proportional Hazard Models

Pair-wise comparisons of adoption timing are likely to be affected by the private signals and observation of peers’ decisions. To control for the two factors, we fit time-varying covariates to the observed hazard rates. Specifically, we estimate Cox proportional hazard models that specifies the hazard rate to be proportional to a baseline non-parametric hazard, where the proportionality factor is linear in the included covariates. Two time-varying covariates are of primary interest: (i) the number of *B*s and *C*s adopted by each subject’s peer group prior to the current round, N_B and N_C ; and (ii) the absolute difference in the number of *H*s and *L*s drawn by each subject including the current round, $|h - l|$. Since at this point we do not distinguish which of the two innovations

Table 4: Cox Proportional hazard models with covariates.

Model	1	2	3	4†	5	6
Treatment <i>PO</i>	0.132 (0.060)**		-0.332 (0.073)***			
Treatment <i>CO</i>	0.239 (0.061)***		-0190 (0.073)***			
N_B		0.135 (0.036)***	0.130 (0.036)***	0.129 (0.036)***		
N_C		0.124 (0.039)***	0.114 (0.039)***	0.112 (0.039)***		
$ h - l $		0.304 (0.034)***	0.312 (0.034)***	0.316 (0.034)***	0.305 (0.034)***	0.308 (0.034)***
$N_B + N_C$					0.130 (0.033)***	
$ N_B - N_C $						0.029 (0.030)
N	4,293	3,701	3,701	3,701	3,701	3,701
Pseudo R^2	0.001	0.004	0.005	0.005	0.004	0.004
Log likelihood	-14,236.4	-12,521.9	-12,511.1	-10,816.6	-12,521.9	-12,528.9
Log like at zero	-14,244.8	-12,575.3	-12,575.3	-10,871.7	-12,575.3	-12,575.3

Key: * = significant at 10%, ** = significant at 5%, *** = significant at 1%. † = Stratified by treatment.
Standard errors are clustered by subject.

is adopted, B or C , it should not matter the absolute number of H s and L s that are drawn.

Table 4 contains the results of estimating Cox proportional hazard models with various combinations of covariates on the entire sample of all three treatments. Dummy indicators of the treatments were sometimes included in which case Treatment P was taken as the excluded category.

The positive coefficients on the two treatment dummies in Model 1 confirm that the observation of prior adoptions accelerates adoption regardless of whether or not they are informative. The larger coefficient on Treatment CO dummy compared to that for Treatment PO is consistent with the comparison of hazard rates in Figure 1.

Models 2 through 6 include some combination of the private and public information variables. All estimated coefficients are positive and, with one exception, statistically significant. The sizes of the coefficients on the time-varying covariates are robust across specifications. We conclude that observation of peers' prior choices induces subjects to

make an adoption decision sooner (without regard to which of the two innovations was chosen). As a generalization, the results confirm that subjects tend to adopt more quickly as observations of B s and C s increase and as the absolute difference between draws of H s and L s increases. Note that the positive coefficients on N_B and N_C are about the same size within each model, and across the models. Also, the coefficients on the variable $|h - l|$ are more than twice as large as those on the variables N_B and N_C . This is our first evidence that subjects are relatively more responsive to their private signals when deciding when to adopt compared to the observation of peers' prior actions.

Table 4 also reports coefficients when the estimated hazard model was stratified by treatment: Model 4 allows the baseline hazard to vary across the three treatments. Compared to when a common baseline hazard is assumed, coefficients do not differ much.

Model 5 includes the *total* number of prior adoptions of both types, $N_B + N_C$, replacing the variations with the number of the two innovations separately, N_B and N_C . The coefficient on the sum is again positive and significant, and close in size to the coefficients on the separate variables.

Finally, we check whether subjects might be spurred on by an imbalance of the prior adoptions between to the two innovations. Model 6 includes the absolute difference in the adoptions of B s and C s as of the prior period, *i.e.*, $|N_B - N_C|$. The coefficient on this variable was not significant indicating that the cumulative number of adoptions, not the innovations chosen, determine the speed of adoption.

Competing Risks Analysis

Up to this point, our survival analysis did not make any distinction as to which of the two innovations subjects adopted. It is possible, however, that there could be differences in timing of adoption between the two innovations if only because the payoff of B is certain while the payoff of C is risky. We allow for such differences by estimating a “competing risks” model of adoption timing. A typical application of this model occurs in medical studies in which individuals can die from two or more causes, such as from a heart attack or from some type of cancer.¹⁹ In this medical application, as in our experiments, each of the two possible

¹⁹A typical example of competing risks found in labor economics occurs when workers leave unemployment for one of several reasons, as when they return to work

events is irreversible. In our case, once a subject adopts either B or C , she cannot make a change for the remainder of the decision sequence. Consequently, when one of the two events occurs, the other event can no longer be observed for that subject. Estimation of the timing of adoption of B taking the adoption of C as a competing risk, relates the covariates to this specific event rather than to the event of adoption of either innovation. This approach allows us to disentangle the causes for adoption of B and C .

Figure 2 plots the Cumulative Incidence Functions (“CIF”) for adoption of B and C separately, for each of the three treatments. The CIF gives the portion of all subjects that have adopted an innovation at some time up to and including a round of the decision sequence. Solid curves are used for CIFs for adoption of B while the dashed curves are used for C . Consistent with Figure 1, the timing of adoption is similar under Treatments PO and CO , but differs from the baseline Treatment P . Under Treatment P , adoption of both technologies is greater in the

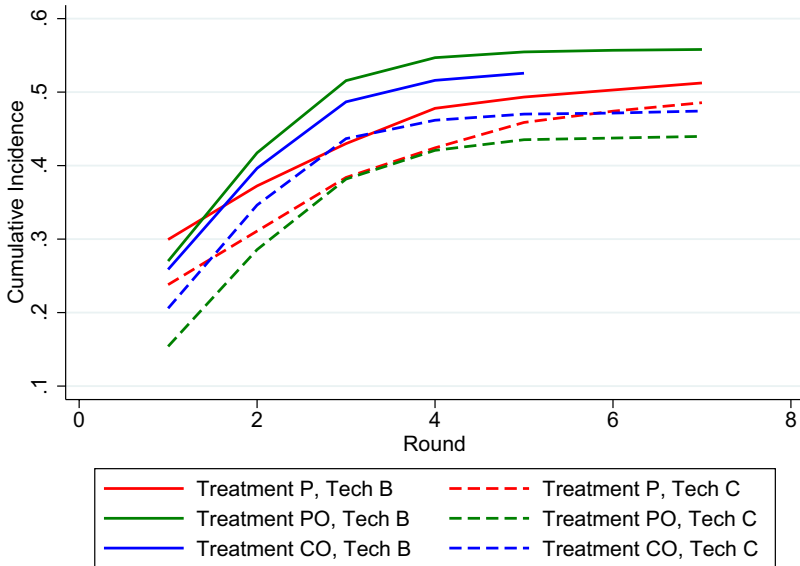


Figure 2: Cumulative incidence of B and C by treatment.

with their previous employer, take a job with a new employer or drop out of the workforce entirely.

first round, but the other two treatments catch up quickly. Despite equality in expected payoff, technology B is adopted more frequently than technology C for all three treatments, i.e., solid CIF is below the dashed CIF in Figure 2. Subjects apparently delay their adoption of C as they gather additional information of both kinds. Note, however, that the adoption of B would also be delayed as this information is gathered, and yet it occurs sooner as if the risk associated with adoption of C prolongs deliberation.

We next estimate a parametric model that has a similar specification to the Cox proportional hazard model in the presence of a competing risk. The results are presented in Table 5. Models were run for both B and for C with the other innovation acting as a competing risk. Two specifications are included in the table, one of which excludes Treatment P (because no recording of prior adoptions in that case) and one which includes fixed effects for the other two treatments. Both specifications include the cumulative number of prior adoptions of B and C , N_B and N_C , as well as the relative difference in private signals, $h - l$. Note that now we do not take the absolute value of the difference in signals as the

Table 5: Competing risks models of adoption of B and C .

Event of interest	Adoption of B		Adoption of C	
Competing risk	Adoption of C		Adoption of B	
N_B	0.488 (0.025)***	0.505 (0.024)***	0.350 (0.026)***	0.381 (0.025)***
N_C	0.280 (0.029)***	0.313 (0.029)***	0.615 (0.025)***	0.623 (0.025)***
$h - l$	-0.569 (0.020)***	-0.556 (0.020)	0.583 (0.023)***	0.581 (0.023)***
Treatment				
Fixed effects	No	Yes	No	Yes
N	5,685	6,209	5,685	6,209
Log likelihood	-10,812.89	-12,200.1	-10,088.01	-11,185.61

Key: * = significant at 10%, ** = significant at 5%, *** = significant at 1%. Standard errors are clustered by subject.

measure of private information since the number of each type of signal should matter.

The coefficients are very precisely estimated and have the expected signs. For instance, observation of prior adoptions of B s, N_B , leads to more rapid adoption of B . Similarly, observation of prior adoptions of C s, N_C , accelerate adoption of C . But note the cross-effects in Table 5: observation of prior C s (B s) also accelerates the adoption of B (C). The cross-effects, however, are smaller than the direct effects.

Also, as expected, the speed of adoption of B (C) is directly increasing in the draws of L s (H s). The parametric model allows us to compare the relative impacts of these factors on speed of adoption of the two innovations. In particular, the observation of prior adoptions of C has a greater impact on adoption of C (about 0.62), whereas the counterpart effect of prior adoptions of B on the adoption of B is smaller (about 0.50). The net number of signals has about the same impact on adoption of B and C , though opposite in sign (*i.e.*, +0.58 and -0.56, respectively).

The Speed-Accuracy Tradeoff

We have seen how the two types of information affect the speed and accuracy of subjects' adoption decisions in different ways. For instance, observation of peers' decisions and correlation of group payoffs accelerate adoption in Treatments PO and CO , respectively. This was quantified, among other ways, in the last column of Table 3. We might expect, therefore, that adoption accuracy under Treatment P would be superior because in that case subjects draw more private signals; in fact, subjects are less accurate in that treatment than under the other two as shown in the fifth column of Table 3. In this section we summarize using a single measure of how subjects trade off between speed and accuracy of their adoption decisions.

Table 6 is an initial attempt to capture both speed and accuracy of subjects' decisions in each of the three treatments. First, note the third column offers an alternative measure of accuracy as the average number of rounds that subjects earn the highest payoff possible — itself a measure of combined speed and accuracy of adoption decisions. Again, subjects perform worst in Treatment P spending fewer rounds receiving the high-payoff innovation ($p = 0.052$). The difference in this

Table 6: Tradeoff between speed and accuracy.

Treatment	Obs.	Avg no. of high-payoff Rounds	Average payoff by round of adoption				
			Round 1	Round 2	Round 3	Round 4	Round ≥ 5
<i>P</i>	525	4.06 (0.22)	17.59 (0.83)	17.95 (1.31)	15.76 (0.86)	15.39 (0.86)	10.63 (0.82)
<i>PO</i>	900	4.49 (0.14)	17.52 (0.59)	18.14 (0.58)	15.66 (0.69)	16.02 (0.84)	12.36 (1.37)
<i>CO</i>	720	4.27 (0.12)	16.96 (0.61)	18.43 (0.61)	16.49 (0.72)	17.10 (0.95)	12.53 (1.61)

Robust standard errors in parentheses, clustered by subject.

performance measure for Treatments *PO* and *CO* is not statistically significant. As before this is surprising to the extent that Treatment *PO* provides subjects with only irrelevant, potentially confusing, information relative to Treatment *P*.

The remaining portion of Table 6 groups subjects by the round in which they adopted, and reports the average payoff of each group of subjects. For all three treatments, the average payoff earned by subjects who adopted in Round 2 was higher than the average payoff to a subject that chose to adopt in a different round. Delay pays off: empirically, subjects do better when they do not adopt in Round 1 as prescribed by the Bayesian Nash Equilibrium. In fact, the difference in average profits between first- and second-round adopters is not necessarily statistically significant at the usual 5% level.²⁰ Nevertheless, subjects who adopt in Round 3 or later earn significantly less than those who adopt in Round 2.²¹

Apparently, it pays to delay long enough to observe the first round of peers' adoptions (and to draw a second private signal as well) but beyond that average payoffs tend to suffer. This regularity, which holds

²⁰Treatment *CO* comes closest to significant difference, with those adopting in the second round out-earning their first-round counterparts by 1.47 tokens ($p = 0.09$, robust).

²¹The majority of pairwise comparisons between Round 2 average payoffs and payoffs to adoptions in other rounds are statistically different at the 5% level.

across all treatments, compels us to examine adoption rules that, while not necessarily BNE strategies, generate high profits in our experimental settings.

Evaluating Behavioral Adoption Rules

The different types of information available to subjects in our experiments, and the complex incentives that affect their adoption decisions, make theoretical predictions difficult. As described in section “Theory,” in Treatments *P* and *PO*, the Bayesian Nash equilibrium has subjects adopt in the first round, while the analysis of Treatment *CO* is less conclusive. As optimal behavior in Treatment *CO* depends on the *actual* behavior of subjects’ colleagues, we turn to an empirical analysis to enhance our understanding of optimal behavior in Treatment *CO*, with the same calculations performed for Treatments *P* and *PO* included for comparison.

We propose nine plausible behavioral rules that subjects might follow, and calculate the hypothetical profits that subjects would have earned under each rule²² in each decision sequence. The nine rules exhaust all strategies for adopting innovation *B* or *C* in the first three rounds that are ‘rationalizable,’ in the sense that subjects do not adopt an innovation that is neither *Leading* (“*L*”) nor *Popular* (“*P*”), if one of them exists.²³ The strategy labels denote the rank-order of priority. Thus, *L1* is the strategy in which a subject adopts the leading innovation in the first round, while *L2–P2–L3* has the subject adopting the leading innovation in Round 2, breaking ties if neither is leading at Round 2 by choosing the popular innovation, and choosing the leading innovation in Round 3 if neither is popular in Round 2. The symbol “*R*” indicates that a subject randomizes between the adoption of *B* and *C* when their other adoption criterion fails to force a decision. Thus, the rule *L2–P2–R2* means that a subject adopts the leading innovation in Round 2, breaks ties by adopting the popular innovation, and picks one or the other if there is no popular or leading innovation.

²²In performing this exercise, we implicitly assume that other group members’ behavior did not depend on the subjects represented in each observation.

²³We define ‘Leading’ and ‘Popular’ technologies as in previous sections. There are fewer strategies listed for Rounds 1 and 3, as there is always a Leading innovation, so less tie-breaking is required.

Table 7: Average payoff under behavioral rules.

Treatment P		Treatment PO		Treatment CO	
L1	19.2 (0.49)	L1	19.0 (0.36)	L1	19.1 (0.41)
L2-L3	17.8 (0.45)	L2-P2-L3	17.9 (0.28)	L2-L3	18.3 (0.33)
L3	17.0 (0.36)	L2-L3	17.8 (0.29)	L2-P2-R2	18.0 (0.34)
		L2-P2-R2	17.6 (0.29)	L2-P2-L3	18.0 (0.34)
		L3	16.9 (0.27)	L3	17.4 (0.30)
		P2-L2-R2	15.6 (0.27)	P2-L2-P3-L3	16.8 (0.33)
		P2-L2-P3-L3	15.5 (0.26)	P2-L2-R2	16.6 (0.34)
		P2-P3-L3	15.2 (0.26)	P3-L3	16.5 (0.29)
		P3-L3	14.8 (0.24)	P2-P3-L3	16.5 (0.33)

Table 7 reports the average profits calculated for each strategy, in descending order of average payoffs for each treatment.²⁴ Clearly, honest adoption (*L1*) performs the best in Treatment *CO*, as it does in the other two. The calculated payoffs of Table 7 are largely consistent with the calculations of section “Theory.” In all treatments, it is best to follow private information, though following peer observation is punished less in Treatment *CO*, due to the benefits of observational learning.

Policy Options and Conclusions

We design an experiment that captures three elements inherent in the decision to adopt environmental innovations: (i) uncertainty about net

²⁴Only three of the nine behavioral rules are possible under Treatment *P*.

benefits of the new technologies, (ii) observation of private information and of others' actions, and (iii) freedom to choose the timing of the adoption. We find that subjects routinely depart from rationality, often adopt a new technology later than was optimal given available information. This pattern corroborates the real-world delay that lies at the heart of the "energy-efficiency gap" discussed in the Introduction, and leads us to explore the features of our experimental manipulations that affect the speed of adoption.

Analyzing the experimental data using parametric and non-parametric models, we conclude that, when choosing between the safe and risky innovation, subjects place greater weight on their private signals compared to observation of prior adoptions by their peers. Observation of prior decisions does, nevertheless, tend to accelerate adoption, all else equal. This pattern is consistent with field evidence regarding adoption of environmental innovations. As mentioned above, Bollinger and Gillingham (2012) and Graziano and Gillingham (2015) find that observation of installations of photovoltaic solar panels by households in the same geographic area tends to accelerate adoption by other nearby households.

One interpretation of this finding is that peers' actions convey their private information about the new technology. However, we also found in our experiments that adoption is hastened even when payoffs of adoptions are statistically independent, in which case peers' adoption decisions do not contain any useful information. In fact, prior adoptions not only accelerate adoption by laggards, but they do so regardless of what innovation is chosen by peers. Prior adoptions spur the laggards to act, but do not cause them to imitate the early movers. If, as we would suspect, subjects are slow to adopt due to some form of *status quo bias*, we are seeing that observation of actions of any sort by their peers overcome their inertia.

Reducing, or removing, this inertia is central to policies that aim to promote adoption of new technologies that conserve energy or protect the environment. Rather than inducing adoption through monetary incentives — such as through the use of taxes, subsidies or price controls — evidence points to informational strategies to promote these adoptions.²⁵ An immediate recommendation emerging from our results

²⁵See, e.g., Fischer (2008) and Olander and Thøgersen (2014).

is to focus on the volume of adoption, especially the total counts of adoptions by whomever the decision makers perceive as their peer group, e.g., households located in the same neighborhood, or other firms in the same industry.²⁶

These results may have practical implications for the real-world adoption of environmental innovations — for both decision makers who must not only select among competing technologies but also choose the timing of the adoption, and for policy makers who wish to influence those decisions. Specifically, our experiments affirm the delay that lies at the heart of the energy-efficiency gap discussed in the Introduction, and could enlighten policy makers that aim to hasten the adoption of a superior alternative.

The diffusion of energy conservation technologies has received considerable attention recently. Numerous field studies have been conducted that quantify how financial and informational programs affect residential energy usage. For instance, Delmas *et al.* (2013) conduct a meta-analysis of 156 field trials of electricity conservation programs for residential consumers. They found that programs that supplied consumers with information about usage of electricity by their peers resulted in an average reduction of 11.5% in electricity use, a significantly larger reduction than the average of 7.4% registered by all methods including the provision of information about consumers' prior personal usage or offering monetary incentives for usage reduction. Few of these studies, however, focus on the speed with which consumers respond to the programs.

Our results are consistent with the findings of these field experiments in that the timing of adoption of our subjects is influenced by the decisions of others. It could be that in both the laboratory and the field the typical decision maker is relatively certain as to which innovation to pursue but his uncertainty as to when to take action results in delay. An additional source of information — even when the feedback provides no useful information — can spur an agent to take action without necessarily steering them toward a particular option. Setting aside concerns about privacy, policy makers could accelerate adoption by collecting and disseminating individual adoption decisions

²⁶While we do not test how subjects might respond differently to others inside or outside of a reference group, field studies confirm this difference, as do laboratory experiments. See Viscusi *et al.* (2011).

to prospective adopters, or by offering subsidies to early adopters.²⁷ Our findings suggest that policy makers ought to favor the use of information dissemination relative to the recently popular use of “nudges” which leverage cognitive biases. The concept of choice architecture falls somewhere in between: the context of a decision, including both information and framing (such as the source of the information), can positively affect outcomes as we have observed here.

As has been often noted, e.g., Rogers (1962), lead adopters can be instrumental in promoting diffusion of new technology. Jaffe and Stavins (1994b) observe that “if the act of adopting a new technology is, itself, a source of useful information for others, then the act of adoption creates a positive externality by providing information to others for which the adopter is unlikely to be compensated. This (positive) externality is another form of market failure [slowing down adoption]” (p. 805). The policy question becomes how to internalize this positive externality that early adopters confer on procrastinators. One way to encourage early adoption is to offer a subsidy that is declining over time and that is independent of the adopted innovation — so as not to bias technology choice, but rather to encourage adopters to use their private signals. To an extent, this approach was adopted with the federal solar panel tax credit plan.²⁸ From the time it began in 2006 and running through 2019, residential adopters of solar panels could/can claim a 30% credit off their tax bill. The credit drops exogenously to 26% in 2020 and then to 22% in 2021, at which time the program ends. Adopters are given an incentive to move sooner rather than later, and this could be seen as a reward for motivating other households to adopt sooner than they otherwise would.²⁹

We need to add that, while early adoption of an innovation may advance the date when it begins to pay dividends, delaying the decision generates additional information on the relative merits of the competing innovations. Indeed, our experiments show that the highest average payoffs are earned by subjects who adopt an innovation in the second

²⁷The 2006 California Solar Initiative had this effect since it offered rebates that declined over time as cumulative installations reached certain thresholds. See Bollinger and Gillingham (2012).

²⁸See <https://energy.gov/savings/residential-renewable-energy-tax-credit>.

²⁹To the extent that costs of solar panel systems fall over time, the present value of the marginal incentive is steeper than would appear from the credit percentages.

round, and not in the first as is prescribed by the risk-neutral equilibrium. Further, subjects who select in the first or second round the innovation that is favored by private signal(s) earn the highest average payoffs, compared to subjects who follow other plausible behavioral rules of thumb which specify wait-and-see strategies. The importance of timing, as well as the tradeoffs involved for both market actors and policy makers, is particularly relevant for certain environmental innovations, e.g., with potentially limited infrastructure in place, such as the choice between a hybrid and a hydrogen fuel cell vehicle. Hence, programs that disseminate the adoption decisions of peers may not only encourage earlier movement away from inefficient legacy technologies but also offer the added bonus of improving the choices of later movers regarding innovative new technologies.

Appendix A — Instructions for Laboratory Subjects

Experimental Instructions: [treatment CO]

Welcome to this experiment on decision making and thank you for being here. You will be compensated for your participation in the experiment, though the exact amount you will receive will depend on the choices you and others make and on random chance (as explained below). Please pay careful attention to these instructions, as a significant amount of money is at stake.

Information about the choices that you make during the experiment will be kept strictly confidential. Your name will appear only on the payment-receipt form and will not be linked to any specific choices you make. You will not be asked to reveal your identity or the content of your decisions to anyone else (either the experimenter or other participants) at any time during or after the experiment. In order to maintain privacy and confidentiality, please do not speak to anyone during the experiment and please do not discuss your choices with anyone even after the conclusion of the experiment.

The experiment will consist of a series of 16 problems, the first of which will be a practice problem for which you will not be paid. Each problem is called a Decision Sequence and is made up of eight rounds. During each Decision Sequence, you will decide if and when to move between three alternative courses of action: options A, B, and C.

You will begin each Decision Sequence following the default option, A. During Round 1 of 8, you will be asked to choose one of three options: to remain with option A, to switch to B, or to switch to C. As long as you remain with A, you will be given the same three options during each subsequent round in the sequence. After you switch to B or C, you will no longer have the opportunity to move between options. You will remain with either B or C (depending on which one you selected) each subsequent round until the end of the sequence. For example, you may choose to remain in option A for the first three rounds, and then switch to option C in Round 4. You must then remain in option C for the rest of that Decision Sequence.

At the beginning of each sequence, you will be placed into a group with five other participants that are in the lab with you. Group assignments will be made randomly and will change at the end of each sequence. At the beginning of each round, you will be told how many of your group members were following each alternative (A, B, and C) as of the end of the previous round.

There is a payoff associated with following each option. For each round that you follow option A, you will receive 1 token. For each round that you follow B, you will receive 2 tokens. The payoff for C in each decision sequence is the same for all members of your group and will be randomly determined at the beginning of each decision sequence; it will remain the same until the end of that decision sequence. The payoff for C will either take a value of 0 tokens per round or 4 tokens per round, both being equally likely. You will not be told if your group's payoff for C is 0 or 4 in any given decision sequence, but at the beginning of each round you will be shown the result of a coin flip that will help you determine what the payoff for C is. If the payoff for C is 0, the coin flip will result in an H with $1/3$ probability and an L with $2/3$ probability. If the payoff for C is 4, the coin flip will result in an H with $2/3$ probability and an L with $1/3$ probability. The payoff for C will be the same for each member of your group in each decision sequence, but the coin flips themselves are generated randomly for each group member individually, and therefore may or may not be the same. Figure A1, below, illustrates the payoffs.

Your payoffs for the eight rounds in a sequence will be tallied at the end of the sequence. For example, if you decided to remain with option A for the first four rounds and then switched to option B, you

- Option A** = 1 token per round
- Option B** = 2 tokens per round
- Option C** = either 0 or 4 tokens per round. Each is equally likely.

Coin Flips:



Figure A1: Payoffs for options A, B, and C.

would receive $4 * 1 + 4 * 2 = 12$ tokens. Your tokens from all 15 Decision Sequences will be added at the end of the experiment and converted to U.S. dollars (the exchange rate is: 12 tokens = \$1), and you will be given a check for that amount.

The payoff for option C can be 4 or 0 tokens. For example, if you remain with option A for the first two rounds and switch to option C in the third round, your payoff will be $2 * 1 + 6 * 4 = 26$ if the payoff for option C turns out to be 4, and your payoff will be $2 * 1 + 6 * 0 = 2$ if the payoff for C turns out to be 0. Note that the true value of the payoff for C will remain the same (either 0 or 4) for all eight rounds of the Decision Sequence.

For each round that you remain with option A, you will view a flip of the coin that will help you determine the payoff for option C. For example, if you choose option A during rounds 1, 2 and 3, you will see four flips of the coin (at the beginning of Rounds 1–4).

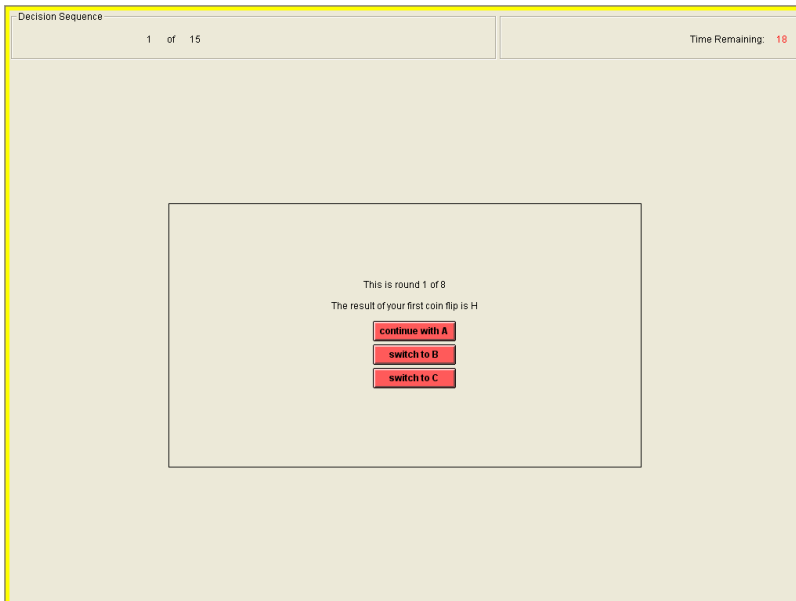
After all eight rounds of a Decision Sequence, you will be told how many tokens you accumulated in that sequence. Then a new sequence will begin and you will go through the same process again. Remember that payoff for C is independently chosen for each sequence, so the value of C’s payoff in previous sequences is not an indication of what it will be in future sequences. After all 15 Decision Sequences, your tokens from all sequences will be added together, converted to dollars, and given to you in the form of a check.

Computer Program Description

To make your decisions you will use the computer in front of you. Right

now, you can see an initial waiting screen. The program will be activated when the instructions are finished.

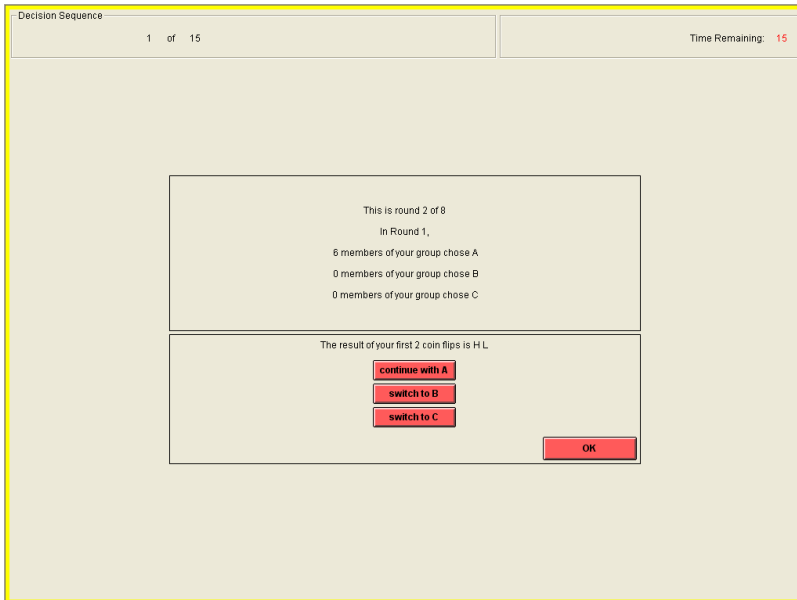
Once the program is activated, you will be given a short quiz and go through a practice decision sequence to make certain that you correctly understand how the experiment will work. After you have completed the quiz and the practice sequence, the experiment will begin. At that time, a new window will pop up and replace the initial waiting window. The new window will resemble the one shown below. The upper-left of the screen informs you that you are in the first of 15 Decision Sequences. The upper-right of the screen shows how much time you have remaining to make your decision. You will have a total of 20 seconds to make your decision, after which the computer will choose option A for you. The center of the screen will inform you that you are in Round 1 and tell you the result of a coin flip that will help you determine the payoff for option C. The screen below shows that the first flip of the coin resulted in an H. Finally, you are asked for your decision. You may choose option A, option B, or option C by placing the cursor on the corresponding button and left-clicking.



If you choose option B or C in Round 1, you remain with that option for the remainder of the rounds of the Decision Sequence. In

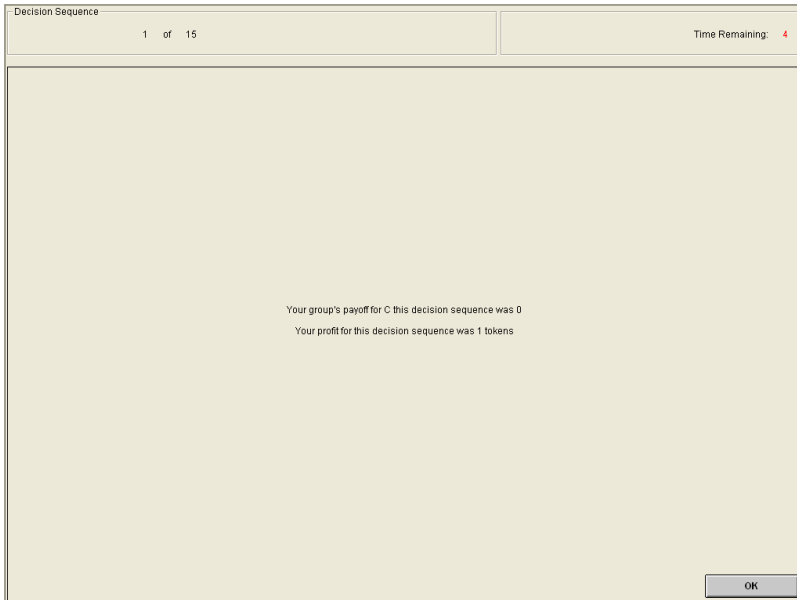
this case, you will again see the waiting screen that is in front of you now, which will remain until the other participants are finished with the first Decision Sequence.

If you choose option A in round 1, you will see a similar screen as the second round begins. Notice that the upper-left of the screen shows that you are still in the first decision sequence. As shown in the upper-right, you will again have 20 seconds to make your decision, after which the computer will select option A for you. The screen will inform you that you are in Round 2. It will also show the options chosen by other members of your group in the previous round. Below, the screen shows that six group members chose A, while no group members chose B or C. Just underneath, the screen shows the result of a new coin flip. The screen shown below indicates that the coin flips for rounds 1 and 2 resulted in an H and an L. Notice that you are shown the results of the coin flips for both round 1 and round 2.



If you choose option B or C in round 2, you remain in that option and are therefore shown the waiting screen. If you choose option A in round 2, you will be told the result of an additional coin flip and be asked to make another decision.

This process will repeat itself until you choose option B or option C, or until you complete round 8. After the Decision Sequence is over, you see the feedback screen. The feedback screen shows you the payoff for C, 0, and your profit for the round, 1 token.



You have a maximum of 20 seconds to observe the feedback screen from the first Decision Sequence. You may click the OK button in the lower right corner of the screen when you are ready to proceed to the next Decision Sequence. After 20 seconds, you will proceed automatically.

Then, the new Decision Sequence will start, and the computer will again randomly place you into a group with five other participants, randomly select a coin, show more information and ask for a new decision.

After the last round is finished, a final screen will pop up, informing you of your total earnings for this experiment.

Rules

Please do not talk with anyone during the experiment. We ask everyone to remain silent until the end of the last round.

Your participation in the experiment and any information about your earnings will be used solely for research purposes. Your name and association to your decisions will be kept strictly confidential. Your payment receipt and participation form will be the only places in which your name is recorded.

Questions

Any clarification questions should be asked at this time. Please raise your hand and wait for an instructor to come to your desk.

Appendix B — Survival Analysis

Non-parametric Comparison of Adoption Timing

Survival analysis summarizes the timing of an event using its hazard function, or equivalently, its corresponding survival function. There are several ways to test statistically the equality of two survival functions. A standard non-parametric test of equality of two survival functions is the “log-rank test.”³⁰ Table B1 reports the significance levels of the log-rank test for each pair of the three treatments. The figures in the table are the Chi-squared statistics of a non-parametric test of the equality of two estimated survival functions over all eight rounds of each decision sequence.

Consistent with a visual comparison of the hazard plots in Figure 1, the tests reject that survival under Treatment *P* is statistically the

Table B1: Log-rank tests of equality of survival functions.

	<i>P</i>	<i>PO</i>	<i>CO</i>
<i>P</i>			
<i>PO</i>	4.31**		
<i>CO</i>	13.94***	5.30**	

Significance: * = 10%, ** = 5%, *** = 1%.

³⁰This test is also known as the “Mantel–Haenszel test.” See Kalbfleisch and Prentice (1980).

same as survival under each of the other two treatments. The tests also reject that timing under treatments *PO* and *CO* are the same — affirming that the existence of common values speeds up the pace of adoption.

Cox Proportional Hazard Models of Adoption Timing

The log-rank test imposes no parametric restriction on the three survival functions. In this section, we again test for differences in timing under the three treatments but now using the popular Cox proportional hazard model. The Cox model assumes that the observed hazard is proportional to a base hazard rate. We take the base hazard rate to be the hazard under Treatment *P*, and estimate the proportionality of the hazards of the other two treatments relative to this baseline. The results of this estimation are found in Table B2. The coefficient β in the table determines the factor of proportionality, e^β , relative to the hazard for Treatment *P*.

Table B2: Cox proportional hazard models relative to the baseline treatment, *P*.

Treatment	Coef.	Std. Err.	z	$P > z $	Lower 95% CI	Upper 95% CI
<i>PO</i>	0.131	0.059	2.20	0.028	0.014	0.249
<i>CO</i>	0.239	0.061	3.91	0.000	0.119	0.359

The fact that both estimated coefficients are positive confirms that the adoption rate is faster under Treatments *PO* and *CO* than under Treatment *P*. The coefficients are all statistically significant at the usual 5% level. As observed earlier, adoption occurs much faster when outcomes are perfectly correlated compared to when they are statistically independent. Indeed, the coefficients on the treatments involving a common payoff are roughly *twice* as large as those with independent payoffs, indicating that common values accelerates adoption.

References

- Allcott, H. and M. Greenstone (2012), “Is There an Energy Efficiency Gap?”, *Journal of Economic Perspectives*, 26(1), 3–28.
- Allcott, H. and D. Taubinsky (2015), “Evaluating Behaviorally Motivated Policy: Experimental Evidence from the Lightbulb Market”, *American Economic Review*, 105(8), 2501–38.
- Anderson, L. R. and C. A. Holt (1997), “Information Cascades in the Laboratory”, *American Economic Review*, 87(5), 847–62.
- Ayres, I., S. Raseman, and A. Shih (2013), “Evidence From Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage”, *Journal of Law, Economics, and Organization*, 29(5), 992–1022.
- Baik, K. H., T. L. Cherry, S. Kroll, and J. F. Shogren (1999), “Endogenous Timing in a Gaming Tournament”, *Theory and Decision*, 47, 1–21.
- Banerjee, A. (1992), “A Simple Model of Herd Behavior”, *Quarterly Journal of Economics*, 107(3), 797–817.
- Bikchandani, S., D. Hirshleifer, and I. Welch (1992), “A Theory of Fads, Fashion, Custom and Cultural Change as Informational Cascade”, *Journal of Political Economy*, 100(5), 992–1026.
- Bollinger, B. and K. Gillingham (2012), “Peer Effects in the Diffusion of Solar Photovoltaic Panels”, *Marketing Science*, 31(6), 900–12.
- Çelen, B. and S. Kariv (2004), “Distinguishing Informational Cascades from Herd Behavior in the Laboratory”, *American Economic Review*, 94(3), 484–97.
- Chamley, C. (2004), “Delays and Equilibria with Large and Small Information in Social Learning”, *European Economics Review*, 48(3), 477–501.
- Chamley, C. and D. Gale (1994), “Information Revelation and Strategic Delay in a Model of Investment”, *Econometrica*, 62(5), 1065–85.
- Costa, D. L. and M. E. Kahn (2013), “Energy Conservation Nudges and Environmentalist Ideology: Evidence from a Randomized Residential Electricity Field Experiment”, *Journal of the European Economic Association*, 11(3), 680–702.
- Covert, T. R. (2015), “Experiential and Social Learning in Firms: The Case of Hydraulic Fracturing in the Bakken Shale”, Available at SSRN: <https://ssrn.com/abstract=2481321>.

- Delmas, M. A., M. Fischlein, and O. I. Asensio (2013), "Information Strategies and Energy Conservation Behavior: A Meta-analysis of Experimental Studies From 1975 to 2012", *Energy Policy*, 61, 729–39.
- Fischbacher, U. (2007), "z-Tree: Zurich Toolbox for Ready-made Economic Experiments", *Experimental Economics*, 10(2), 171–8.
- Fischer, C. (2008), "Feedback on Household Electricity Consumption: A Tool for Saving Energy?", *Energy Efficiency*, 1(1), 79–104.
- Gaker, D., Y. Zheng, and J. Walker (2010), "Experimental Economics in Transportation: Focus on Social Influences and Provision of Information", *Journal of the Transportation Research Board*, 2156, 47–55.
- Gerarden, T., R. G. Newell, and R. N. Stavins (2015), "Deconstructing the Energy-Efficiency Gap: Conceptual Frameworks and Evidence", *American Economic Review*, 105(5), 183–6.
- Graziano, M. and K. Gillingham (2015), "Spatial Patterns of Solar Photovoltaic System Adoption: The Influence of Neighbors and the Built Environment", *Journal of Economic Geography*, 15(4), 815–39.
- Heckman, J. J. and B. Singer (1984), "Econometric Duration Analysis", *Journal of Econometrics*, 24(1–2), 63–132.
- Islam, T. and N. Meade (2013), "The Impact of Attribute Preferences on Adoption Timing: The Case of Photo-Voltaic (PV) Solar Cells for Household Electricity Generation", *Energy Policy*, 55, 521–30.
- Ivanov, A., D. Levin, and J. Peck (2013), "Behavioral Biases in Endogenous-Timing Herding Games: An Experimental Study", *Journal of Economic Behavior and Organization*, 87 (March), 25–34.
- Jaffe, A. B. and R. N. Stavins (1994a), "The Energy Paradox and the Diffusion of Conservation Technology", *Resource and Energy Economics*, 16(2), 91–122.
- Jaffe, A. B. and R. N. Stavins (1994b), "The Energy-efficiency Gap: What Does It Mean?", *Energy Policy*, 22(10), 804–10.
- Jessoe, K. and D. Rapson (2014), "Knowledge is (Less) Power: Experimental Evidence From Residential Energy Use", *American Economic Review*, 104(4), 1417–38.
- Kalbfleisch, J. D. and R. L. Prentice (1980), *The Statistical Analysis of Failure Time Data*, Hoboken, NJ: John Wiley & Sons.
- Lei, V., C. N. Noussair, and C. R. Plott (2001), "Nonspeculative Bubbles in Experimental Asset Markets: Lack of Common Knowledge of Rationality vs. Actual Irrationality", *Econometrica*, 69(4), 831–59.

- Levin, D. and J. Peck (2008), “Investment Dynamics with Common and Private Values”, *Journal of Economic Theory*, 143(1), 114–39.
- Murto, P. and J. Välimäki (2011), “Learning and Information Aggregation in an Exit Game”, *Review of Economic Studies*, 78(4), 1426–61.
- Nelson, H. T. (2012), “Lost Opportunities: Modeling Commercial Building Energy Code Adoption in the United States”, *Energy Policy*, 49, 182–91.
- Olander, F. and J. Thøgersen (2014), “Informing versus Nudging in Environmental Policy”, *Journal of Consumer Policy*, 37(3), 341–56.
- Rogers, E. (1962), *The Diffusion of Innovations*, New York: Free Press.
- Sgroi, D. (2003), “The Right Choice at the Right Time: A Herding Experiment in Endogenous Time”, *Experimental Economics*, 6, 159–80.
- Stadelmann, M. (2017), “Mind the Gap? Critically Reviewing the Energy Efficiency Gap with Empirical Evidence”, *Energy Research & Social Science*, 27, 117–28.
- Van Soest, D. P. (2005), “The Impact of Environmental Policy Instruments on the Timing of Adoption of Energy-Saving Technologies”, *Resource and Energy Economics*, 27(3), 235–47.
- Viscusi, W. K., O. R. Phillips, and S. Kroll (2011), “Risky Investment Decisions: How are Individuals Influenced by Their Groups?”, *Journal of Risk and Uncertainty*, 43(2), 81–106.
- Weizsäcker, G. (2010), “Do We Follow Others When We Should? A Simple Test of Rational Expectations”, *American Economic Review*, 100(5), 2340–60.
- Zhang, J. (1997), “Strategic Delay and the Onset of Investment Cascades”, *Rand Journal of Economics*, 28(1), 188–205.