Contributions to Economic Analysis & Policy

Volume 5, *Issue* 2 2006 *Article* 4

FIELD EXPERIMENTS

Online Fund-Raising Mechanisms: A Field Experiment

Yan Chen* Xin Li†

Jeffrey K. MacKie-Mason‡

*University of Michigan, yanchen@umich.edu

†University of Michigan, xinl@umich.edu

‡University of Michigan, jmm@umich.edu

Online Fund-Raising Mechanisms: A Field Experiment*

Yan Chen, Xin Li, and Jeffrey K. MacKie-Mason

Abstract

We implemented one of the first web-based online field experiments of fund-raising. We embedded our experiment in the Internet Public Library to test four mechanisms: Voluntary Contribution (VCM), Premium, Seed Money and Matching. Although the gift size is not significantly different across mechanisms, the Seed and Matching mechanisms each generate significantly higher user click-through response rate than the Premium mechanism. Because this is one of the earliest embedded, web-based field experiments, we report our methodology findings in some detail. Cookies work better as participant assignment techniques than pop-up windows and elicitation of geographic information. Participant clickstream data that nominally demonstrate a desire to donate is a poor predictor of actual giving.

KEYWORDS: online fund-raising, public goods, field experiment

---

*We thank Erica Ackerman, Michael Cohen, Maurita Holland and Shelley MacMillan for helpful discussions, John List, the Guest Editor, and two anonymous referees for their constructive comments which significantly improved the paper. We thank Erica Ackerman, Kathy Koehler and John Northup for programming for the experiment, Sasha Achen and Benjamin Chiao for testing the programs, the entire staff of the Internet Public Library for their support, and Kelly Roan and the University of Michigan Development Office for logistic support. Chen gratefully acknowledges the financial support from the National Science Foundation through grant no. SES-0339587, and MacKie-Mason from no. CSS-9988715. Chen and MacKie-Mason: School of Information, University of Michigan, 1075 Beal Avenue, Ann Arbor, MI 48109-2112. Email: yanchen@umich.edu, jmm@umich.edu. Li: Department of Economics, University of Michigan, 611 Tappan Street, Ann Arbor, MI 48109-1220; Email: xinl@umich.edu.

Voluntary contributions to support activities with public goods characteristics is an important economic activity. Individuals in the U.S. gave $188 billion dollars in 2004 (Giving USA 2005). Scholars have studied the motivations for individuals to make voluntary contributions since the 1960s. More recently, there has been attention to the design of fund-raising campaigns in order to elicit higher contributions.[1]

The emergence of the Internet has introduced both new public goods that need funding, and new opportunities for raising funds. An enormous amount of information is freely provided over the Internet, and electronically-distributed information is a near-perfect example of a public good (nondiminishable and nonexcludable). Examples include reviews (of, e.g., books, movies, appliances, former romantic partners), data (e.g., economic, political, sociological), literature, nonfiction treatises, maps, directories and so forth. The creation, collection, organization, storage, indexing, presentation, revision and other functions of an information service all are costly, even when electronic reproduction and distribution cost approximately zero at the margin.

Simultaneously new opportunities for soliciting and collecting voluntary contributions are made available by the low-cost, wide-area, increasingly ubiquitous communications network provided by the Internet. Fundraisers can implement both targeted and diffuse direct e-mailing campaigns at much lower cost than traditional direct mail campaigns.[2] More novel is the ability to embed a contribution system directly in the information delivery mechanism. For example, when a user goes to a Web site to obtain a freely provided document, a button can be provided right next to the document labelled "Donate!", and the button can be programmed to present a relatively simple and immediate electronic funds transfer form, e.g., for a real-time verified credit card transaction.

During the early years of the public Internet (largely the 1990s) there was an explosion of voluntarily-provided information resources. The costs were born by altruistic individuals and organizations who seeded creation and initial operations. However, more and more providers of public information goods have discovered that they need to secure new sources of ongoing funding to support maintenance operations, as well as continuing investments in creation and organization of new material. Thus, interest in effective Internet-based fund-raising mechanisms has grown rapidly at the same time that the scholarly community has been focusing on the more general question of fund-raising mechanism design.

[1] Usually volunteer-supported public goods are *underprovided*, so at the margin increasing contributions is likely to increase efficiency.

[2] The Howard Dean presidential campaign drew much attention for the success of its online fundraising. During the first three months of campaigning Dean raised nearly half of his $7.5 million total through online donations, using a commercial fundraising system (Convio) for nonprofits that allows organizations to implement several of the mechanisms we describe in this paper (Weiss 2003).

In this article we report on a field experiment to test four different contribution mechanisms incorporated directly into a major Internet information service. The four mechanisms are the voluntary contribution mechanism, voluntary contribution with premiums, seed money and matching mechanisms (detailed below). The experiment was run on the Internet Public Library site.

The Internet Public Library (http://www.ipl.org) began as a graduate seminar project at the University of Michigan School of Information in winter 1995. It provides two major services: a subject-classified and annotated collection of materials and links on a wide range of topics, and question-answering reference service. In its ten years of operation, live reference experts have answered over 50,000 questions. In the most recent calendar year (2004) there were about 7.5 million unique visitors and 12.4 million separate visits.[3] Dozens of articles have been written about IPL, and awards received. Several observers conclude that the IPL is "perhaps the largest and most well known" of the free electronic reference and online library collection services (McCrea 2004).

The IPL has been supported by miscellaneous in-kind donations from technology companies, funding from the University of Michigan, and volunteer labor. No ongoing sources of regular funding have been established. Hegenbart (1998) concluded that without regularized funding, the IPL's "economic viability is in jeopardy." This real long-term funding need of the IPL motivates our study.

In our study, we designed a *natural field experiment* (Harrison and List 2004) to test the effectiveness of the four fund-raising mechanisms. We implemented the mechanism by modifying the IPL web site and instrumenting it to track user interactions. To our knowledge, this is the first embedded field experiment which studies fund-raising mechanisms on the Internet. A natural field experiment provides a bridge between a laboratory experiment and direct field observations. It allows us to study behavior in a more natural environment than the lab with participants who are the actual users of the site. Meanwhile, it gives the researcher more control than field observations as we can randomly assign users to different mechanisms and keep all aspects of the environment constant across mechanisms except the description of the mechanism itself. (See Harrison and List (2004) for a comprehensive discussion and classification of field experiments.) In recent years, there have been an increasing number of fund-raising field experiments. List and Lucking-Reiley (2002) study the effect of seed money and refunds in a university capital campaign and find that both mechanisms are effective in increasing contributions. Eckel and Grossman (2005) compare the effects of rebates and matching mechanisms in a public radio fund-raising campaign. They find that matching mechanisms result in larger total contributions compared to their

[3]Private communication with Prof. Maurita Holland, IPL Executive Director.

strategically equivalent rebate mechanisms. Rondeau and List (2005) report results from a fundraising campaign to finance an expansion of an environmental education program of the Sierra Club in Canada. They find that announcement of challenge gifts considerably increases campaign contributions, while matching gifts do not have a significant effect on the average and total donations. While all three studies use direct mail and manipulate the payoff structure faced by individual donors, Shang and Croson (2005) manipulate social information during an on-air fund-raising campaign for a public radio station. They find that the most influential social information is the information drawn from the 90^{th} to 95^{th} percentile of previous contributions. Our study differs from the above in both the context and the medium of implementation. As this is the first online fund-raising field experiment, we discuss our methodological findings in detail. We hope that both the success and pitfalls in our methods will guide future online field experiments.

The rest of the paper is organized as follows. In Section 1 we review the literature on the mechanisms we test. Then, in Section 2 we detail our experimental design, with particular emphasis on the problems associated with implementing the experiment online, and the technologies we used to address these problems. We present our results on the economic hypotheses in Section 3, then discuss our findings on the methodology for online experiments in Section 4. We summarize our conclusions in Section 5.

1 Fund-raising Mechanism Design.

Charitable fund-raising is a sizable industry, spending more than $2 billion per year, with a strong strategic focus (Andreoni 1998). Potential donors are presumed to respond to incentives, and to have concern for whether other potential donors will contribute enough to make a program viable. Thus, scholars have treated fund-raising as a problem for game theory and mechanism design. There is now a sizable theoretical and empirical literature.

We implemented four known mechanisms: a voluntary contribution mechanism (VCM), seed money (leadership giving), matching, and premium. In a VCM, donors contribute to a public good with no guarantee on the total amount that will be raised, and with no direct private return, although of course the donor can obtain private utility from the public good that is provided. The standard problem is an application of the theory of public goods (Samuelson 1954); for a recent review see Andreoni (2006 (forthcoming). The fundamental result is that from a social point of view, contributions will be inefficiently low.

Leadership giving or seed money is widely used by professional fundraisers. The charity secures initial donations, ideally of an unusually large amount, from prominent donors. Such gifts might be expected to encourage free-riding by

other potential donors, but seed money can help avoid an inefficient equilibrium if there is a nonconvexity in the production function for the public good: a minimum threshold that contributions must meet or surpass before any benefits of the public good can be consumed (Andreoni 1998). For example, a new university sports facility cannot be constructed unless at least $75 million is raised. In such a model there are generally two sets of Nash equilibria. The first set is inefficient, consisting of any combination of contributions which do not reach the provision point and no one can unilaterally increase her contribution to reach the provision point. The second set is efficient, consisting of any combination of contributions which add up to exactly the provision point and the fund-raising campaign is successful. The seed money is to encourage other donors to believe that the threshold will be met.[4] List and Lucking-Reiley (2002) recently tested the effect of seed money by manipulating an actual capital campaign at the University of Central Florida, and found that increasing the seed money as a percent of the goal increased the participation rate and had a more than linear effect on the total donations. Increasing the amount of seed money, however, has two opposite effects on public contributions. On the one hand, it might function as a credible signal of the quality of the charity (Vesterlund 2003); Potters et al. (2005) provide experimental evidence to support this hypothesis. Leadership giving might also increase total donations when individual utility functions exhibit warm glow or snob appeal effects (Romano and Yildirim 2001). On the other hand, increasing the seed money in a threshold public goods environment decreases the aggregate amount of money actually required from the general public. To partition these two effects, List and Rondeau (2003) conduct laboratory experiments and find that announcing leadership contribution has sufficient value to eliminate the negative effect created by the implicit reduction in threshold.

Another mechanism is matching: a benefactor commits to match donations at some fixed rate. The mechanism is sometimes used directly by a particular campaign (in a different context corporations sometimes offer to match employee donations to any recognized charity, rather than just one specific campaign). Matching subsidies are at least sometimes more costly to leadership donors than seed money, and in general they would optimally be offered only in conjunction with seed money (Andreoni 1998). The U.S. government uses a type of rebate mechanism for many types of giving through giving a charitable donation deduction on personal income taxes, though the experimental evidence from the laboratory suggests that people respond differently to tax deductions and matching funds (Eckel and Grossman 2003). In a field experiment which closely parallels their lab experiment, Eckel and Grossman (2005) find that matching mechanisms induce larger total contributions

[4]This model and the equilibria is closely related to the provision point mechanism (Bagnoli and Lipman 1989).

to the charity than their strategically equivalent rebate subsidies. The price elasticities are about one and one half to three times as large for the matching mechanism compared to the rebate subsidy. In a recent field study, Rondeau and List (2005) compare the matching mechanism with the seed money mechanism, and find little evidence that the former has a positive effect on charitable contributions. Seed money, on the other hand, significantly increases campaign contributions.

A third well known mechanism is to offer donors some sort of premium, which is an award or prize correlated with the gift size. For example, a charity might give donors a coffee mug for a modest donation and an umbrella for a larger gift. Performing arts organizations often give special treatment to different categories of donors (e.g., ticket priority and special dinners with visiting performers). In a related study, Andreoni and Petrie (2004) investigate the effects of publicity and categorization of gifts (without premium). They find that announcing donors by gift size category did not increase the total amount of giving in an experimental study, but did have a substantial effect on the distribution of gifts. Donors shifted the gift amounts toward the lower bound of a category, typically giving less if they otherwise would have been in the lower half of the category, and giving more to get into the next category if they were otherwise in the upper half of the category. In our experiment, we give gifts as premia, with cash values that are larger for larger categories of giving; such premia might be expected to have a greater incentive effect than merely announcing the names of donor by gift size class.

2 Experimental Design.

We designed a web-based field experiment using the Internet Public Library (IPL) as the platform to test the effectiveness of various fund-raising mechanisms proposed in the literature. We compare four mechanisms, the voluntary contribution mechanism (hereafter shortened as VCM), voluntary contribution with premiums (Premium), seed money mechanism (Seed) and matching mechanism (Matching). We use VCM as the control, and Premium, Seed and Matching as treatments. We compare the participation rates and the average gift size for each mechanism. This is the first online fund-raising experiment, and also one of a relatively small number of online field experiments embedded in production information systems. Therefore, we are also interested in users' behavioral responses to various design features (such as the use of pop-up and pop-under windows, and side-bar buttons), and measurement issues (such as the interpretation of clickstream data).[5]

[5] We attempt to consistently use the term "user" to refer to a person who came to use the IPL website, but may or may not have chosen to participate in some portion of the field experiment. We use "participant" to refer to those users who through some affirmative action choose to participate in at least part of the experiment, where the smallest degree of participation would be entering requested

We conducted two campaigns. The 2004 campaign started on 4 October 2004, and ended on 6 January 2005.[6] The 2005 campaign ran from 20 April through 6 July 2005. While the texts of the soliciting messages of the four mechanisms were the same across the two campaigns, we modified the techniques used for the random assignment of participants, and the design of the alerts that invited participation.

2.1 Campaign 2004

We had two goals for the assignment of participants to mechanisms. Our first goal was to balance the population characteristics assigned to each of the four mechanisms, especially the three treatment mechanisms.[7] The second goal was to ensure that each participant observes one and only one soliciting message throughout the entire campaign. While this latter goal can be easily accomplished in a conventional mail or phone campaign, it is much harder in an online environment because of limited mechanisms for identification and authentication of participants.

To achieve the first goal, we opened a pop-up window when the user first reached the IPL website, in which we asked for the first three digits of the user's postal code (see Appendix A for a screen shot), which we then used to demographically balance the treatments. To achieve the second goal, we decided to use cookies to create persistent identities for participants. We now explain these technologies and our use of them to implement our experimental design. In Section 4 we discuss the behavioral responses by users, and the lessons we learned for embedded online experimental design.

2.1.1 Postal Code Pop-Up Window

We first address our technique for balancing the size and demographics of the treatment populations through the use of postal code geographic identifiers and census demographic data by geography.

When users came to the IPL site for their first visit during the campaign, they were asked to enter the first three digits of their postal codes in a pop-up window.[8]

---

postal code information, or clicking on one of the campaign buttons or links to view the campaign solicitation messages.

[6] We had intended to end campaign on 30 December 2004. However, due to the holiday break, the campaign was not removed until 6 January 2005.

[7] Balancing the assignment of participants to treatments based on observable characteristics that are suspected of being correlated with the behavior under examination is the preferred technique in experimental design since the seminal work of R. A. Fisher (Fisher (1920), Cochran and Cox (1957)).

[8] A pop-up window is a new browser window that appears on top of the user's current window, and that can be created by the software code embedded in a website when triggered by a specified user

We conjectured that this information is relatively low-cost and non-invasive for the users. We specifically wrote the text in the pop-up window so that there is no indication it is connected with fund-raising, or an experiment (see screen shot in Appendix A). We checked the (partial) postal code information against a database of American and Canadian postal codes. Users who entered legitimate postal codes were randomly assigned to one of the four mechanisms. If the entry was not a legitimate postal code, or if the user ignored the pop-up window (entered nothing), she was allocated to the benchmark mechanism (VCM). Users from other parts of the world were also assigned directly to VCM. The IPL staff estimate that roughly 80 percent of its users come from the United States and 3-4 percent from Canada. This assignment method insures that the population characteristics assigned to each of the three treatment mechanisms is balanced, at the cost of a more varied set of characteristics in users assigned to VCM.

For participants from the United States, we adopt the system of three-digit codes that represent ZIP Code Tabulation Areas (ZCTAs) developed by the U.S. Census Bureau for tabulation summary statistics from Census 2000. We then designed an algorithm to categorize all three-digit codes of ZCTAs into four groups so that the variation in population size, urban and rural population, average household income and median household income were as small as possible across mechanisms.[9] Furthermore, as philanthropic attitudes and behaviors are likely to be correlated with socio-cultural characteristics, which, in turn, might be correlated with geographic areas, we use the definition on geographic regions adopted by the Bureau of Economic Analysis (BEA), which group the states into eight regions, i.e., Far West, Great Lakes, Mideast, New England, Plains, Rocky Mountain, Southeast, and Southwest. According to the BEA, the grouping was based on the homogeneity of states with respect to income characteristics, industrial composition of the employed labor force, and noneconomic factors, such as demographic, social, and cultural characteristics. We assigned the ZCTAs of a specific BEA region to the four mechanisms as equally as possible so that the population from any BEA region was not over-represented in any of the four fund-raising mechanisms.

Table 2.1.1 summarizes the population and income distribution by mechanism in the United States. As shown, all the population and income variables roughly follow a uniform distribution across mechanisms. The assignment of population across the BEA regions is also uniformly distributed across the four mech-

action. In our software, connecting to the IPL triggered code to evaluate whether a user should be shown the pop-up window, and if so, to create the window.

[9]The first three-digits of a zip code do not excessively localize populations (the first three digits might be shared by a city of more than 150,000), and in any case we assign many different three-digit aggregations to each mechanism, with the selection of which three-digit codes based on balancing the observable demographic averages to ensure that we are not selecting for homogeneous populations.

Table 1: Population and Household Income Distribution by Mechanism in the U.S.

Mechan.	Total Pop. (millions)	Urban Pop. (millions)	Rural Pop. (millions)	Agg. Household Income (billions)	Median Househ. Income (thous.)
VCM	71.4	57.5	13.9	148.4	39.2
Premium	71.2	56.2	15.1	152.0	39.8
Seed	71.2	56.2	15.0	150.5	40.1
Matching	71.4	56.1	15.3	150.1	39.9

anisms.[10]

We assigned Canadian participants in a similar manner. The form of Canadian postal codes is "ANA NAN", where A is an alphabetic letter and N is a numeric character.[11] Our software checks the three digits participants enter in the pop-up window with the FSA codes in the database to identify Canadian postal codes. We used the first letter to randomly group Canadian participants by province/territory, for which we obtain statistics on the distribution of population and income from Statistics Canada (2001). In Table 2.1.1 we present our aggregate assignment of Canadians by population and earnings across mechanisms.

Table 2: Population and Earnings Distribution by Mechanism in Canada

Mechanism	Postal Code Initial	Total Pop.	Rural/Total Pop.	Earnings Per Capita (in Canadian $)
VCM	A,T,Y	3,516,411	23%	18,449
Premium	C,E,G,H,J,S	9,081,204	24%	16,780
Seed	B,R,V,X	5,999,433	22%	17,848
Matching	K,L,M,N,P	11,410,046	15%	20,697

The proportional assignment of U.S. and Canadian participants to the four mechanisms was designed to ensure that IPL users assigned to each mechanism share the same income and rural/urban residential distribution as the entire population. To achieve this goal, we needed users to enter their postal code information truthfully.

[10] Tables are available from the authors upon request.

[11] The first character of a postal code represents a province or territory, or a major sector entirely within a province. The first three characters identify the Forward Sortation Areas (FSAs) that are associated with a postal facility from which mail delivery originates.

2.1.2 Cookies

A significant problem for our experimental design was the difficulty in identifying repeat visitors to the IPL so that we could ensure a participant received the same experimental treatment each time. This is a special concern for online experiments. Many web-based information services, especially those provided as public goods (IPL included), do not require users to identify or authenticate themselves. The IPL staff wanted us to not create an identification system that required users to create accounts and login for the experiment because they expected this would adversely affect usage and public satisfaction with the IPL service. Therefore we needed to devise another method to implement identification.

Identification requires that a user offer a unique identifier, such as a registered login name. Authentication is a process that verifies the proffered identity, to increase the confidence that the user proffering the identify is actually the owner of that identity. Authentication might also be used to ensure that a user can proffer only a single, unique identity. There typically will be problems for experimental design if multiple different participants offer the same identity, or if a single user returns to the experiment and offers a different identity than the first time.

Because we could not create a system to require active identification (e.g., a user account with login), we implemented a passive method for creating identities and authenticating them. Our method is reasonably reliable when users do not have a strong motivation to interfere with the system (e.g., there is no monetary gain from breaking the authentication system). This method uses cookies.

When a user visited the IPL for the first time during the campaign period, we not only solicited postal code information but also tried to write a cookie on the user's machine. We did this to create a persistent identity for participants, so that we could ensure that the same mechanism was presented to a participant each time he or she returned. A cookie is a small file that a website can request a browser to store on the user's local computer; the browser will comply unless the user has configured it to reject cookie creation requests. Our software generated a unique identifier to store in a user's cookie, and also stored this identifier in our database on the IPL server. Then, each time a user connected to the IPL, we checked to see if that user already had a cookie, and if so, to which treatment it had previously been assigned.[12]

[12]The cookie system provides a modest amount of authentication, as well: the cookie is stored on the user's account on the computer, and thus, is only accessible to our software when someone is logged into that user's account, so we effectively piggyback on the login system for the user's computer to provide authentication. Our technology does not authenticate that participants have and use only one unique identity; for example, we do not have a way to recognize a user at home and the same user at the office as the same person.

If a cookie was written successfully, we created an observation in our database, in which we stored an assigned record ID for the participant, the mechanism ID associated with the campaign message shown, a time stamp when the participant actively clicked to "Contribute now", and other information related to the participant's online activities at the IPL. If a participant returned to the IPL at a later time, we read the cookie on her machine, and identified her in the database. Return users were not shown the postal code window again, and would be shown exactly the same campaign message as on the first visit.

If we were unable to write a cookie on a user's machine, we did not create a record in the database, since we could not create a persistent identifier for the user. In addition, these users were not automatically shown the campaign solicitation message. If the user affirmatively clicked on one of the campaign buttons or links, she would be shown the VCM message. We are aware that assigning users who reject cookies to VCM might lead to a biased sample. For example, if users who reject cookies tend to be technically more savvy, which, in turn, might be correlated with certain demographic characteristics, the population characteristics of users in VCM might be different from the treatments. However, the alternative of random assignment of these users to all four mechanisms leads to a more severe problem. Since we cannot identify these users, they could be assigned to a different mechanism each time they visit. We view the latter problem as being more severe than the first. Therefore, we assign these users to VCM.

We used cookies to complement the postal code volunteered by participants to ensure that each user would be shown only one mechanism. Without cookies, we would have to rely on users to enter (the same) postal codes every time they came to the IPL. This method would greatly increase the risk that some users would be exposed to multiple mechanisms (due to intentional or unintentional entry of different or illegitimate codes). Participants identified by cookie would see the same mechanism, even if they did not provide the correct postal code information at the first visit.

2.1.3 Reaching the Solicitation Message

There were three ways that a participant could be presented with a campaign solicitation message: automatically in a pop-under window; by affirmatively clicking a "Friends of IPL" button displayed on the left side of the IPL homepage; or by affirmatively clicking a small horizontal link labeled "Friends of the IPL" and present at most of the IPL subject pages.[13] A pop-under window appears behind the

[13]The IPL has many subject pages, such as Arts & Humanities, Business, Computers, Education, Entertainment, Newspapers, KidSpace, and TeenSpace. Some users bookmark one or more of the subject pages, and enter the subject pages directly rather than through the IPL main page. We put a

user's current browser window, and thus is visible after the user exits the IPL and closes her main window. The use of a pop-under window to display the solicitation message automatically was suggested by the IPL staff. The use of a pop-under rather than pop-up window avoids interrupting a user's activities. The fund-raising specialist we consulted also argued that the best time to catch users' attention was after they closed the IPL pages, presumably happy with the information obtained from the IPL. In Appendix A we present screen shots of the "Friends" button, and the three screens of the fund-raising appeals.

Screen 1 offers a brief introduction, and then explains the particular fund-raising mechanism to which the participant is assigned. Participants were invited to make a pledge and informed that their gifts would be tax-deductible. Since one third of the IPL patrons are school children, a note states that one needs to be 18 years of age or older in order to make a contribution. Participants could contribute by clicking on the "Contribute Now" link in the lower right corner of screen 1. Alternatively, they could proceed to screen 2, which includes information on the deadline, suggested amount of gifts and a web link to information concerning the Michigan State tax credit.[14] Participants could then choose to contribute by clicking on the "Contribute now" link, or could go back to screen 1. The "Contribute now" link on both screen 1 and 2 will took participants to screen 3, the first page of the University Giving website, where detailed instructions were provided in a vertical frame of the screen, as shown in Appendix A. One could choose to contribute by writing a check or using a credit card online. After entering the University Giving page, we were unable to track participants further due to constraints imposed by the University Giving office. We did receive a weekly report from the University Giving office that showed for each donation received the amount contributed, the donor zip code information, and a time stamp indicating the second when the contribution was completed.

While participants can view the campaign messages for as many times as desired by clicking on one of the "Friends" buttons or links, our software would only display the message automatically via the pop-under window once per day, to avoid overload users. This was requested by IPL staff. In addition, we coded participants who had clicked on the "Contribute Now" button as contributors, and we displayed no further fund-raising appeals automatically.

---

horizontal link of "Friends of the IPL" on every subject page except KidSpace and TeenSpace.

[14]Residents of Michigan could elect to receive the full value of their contribution, up to $200 per taxpayer, as a tax credit, rather than deduct it as a charitable contribution; a deduction is worth less than a credit because the participant recovers only her tax rate times the amount of the contribution (or less if she has high enough adjusted gross income that the limits on itemized deductions are active).

2.1.4 Campaign Messages

The complete solicitation messages for all four mechanisms are in the Appendix B. Each of the four messages start with the same common introduction paragraph of the IPL, and end with the same paragraph, "Making a gift is easy, greatly needed and appreciated, tax-deductible, and directly supports the IPL service you use. Let us welcome you now as our newest IPL Friend!" Each message differed in the description of the particular mechanism.

In VCM, users are invited to make a contribution today, with the baseline description of "Our goal is to raise $20,000 by December 30, 2004, to cover a portion of the annual operational costs of the IPL."

In Premium, small gifts with the IPL logos (mouse pads, book lights and CD cases) were offered to donors according to the size of their gift. We displayed the following message: "Our goal is to raise $20,000 by December 30, 2004, to cover a portion of the annual operational costs of the IPL. All Friends contributing $25 or more will receive a special IPL thank you gift:

- IPL mouse pad ($25-$49)
- IPL book light flashlight ($50-$74)
- IPL CD case ($75 or above)"

Small images of the premium gifts are also included in screen 1 of the appeal.

In Seed, with the help of the Development Office at the School of Information, University of Michigan, we were able to use an anonymous donation of $10,000 as the seed for this campaign.[15] The corresponding message is thus, "Our goal is to raise $20,000 by December 30, 2004, to cover a portion of the annual operational costs of the IPL. We're already half way there with a gift of $10,000! Won't you help us with the final $10,000? Your gift in any amount will make a big difference."

Similarly, in Matching, the Development Office enabled us to offer one-for-one matching, to a total of $10,000. In the message we described this: "Our goal is to raise $20,000 by December 30, 2004, to cover a portion of the annual operational costs of the IPL. To help us get there, an anonymous donor has agreed to match contributions, $1 for $1, up to a total of $10,000. Double the impact of your gift and join the IPL Friends today!"

[15] Because this was a field experiment, in which users were making real donations without informed consent, all elements of the campaign had to be truthful, so we could not announce seed money unless the money was actually obtained as part of the campaign. Therefore, as part of our experimental design, we actually needed to obtain the $10,000 in seed money, and the $10,000 in matching money described below, in order to test these mechanisms.

The total amount of targeted gifts was equivalent cross four mechanisms. Therefore, in VCM and Premium, the goal was set at $20,000, respectively, whereas in Seed and Matching, we set a goal of an additional amount of $10,000, respectively.

In all mechanisms, although gifts of any amounts are welcome, we suggest that the participant consider "a gift of $25, $50, $75, $100, or $365." This technique is common practice in the fund-raising community.

Based on previous theoretical and experimental research, we expect that each of the three treatment mechanisms, Premium, Seed and Matching, will induce higher contribution amount and gift size than VCM. Because of the lack of robust empirical evidence on the ranking of the three treatment mechanisms, our null hypothesis is that they will perform equally well in terms of contribution rate and gift size.

2.1.5 Campaign Re-design

The 2004 campaign was launched on October 4, 2004. However, our original experimental design yielded a skewed distribution of online traffic, since the vast majority of the users did not enter a valid U.S. or Canadian postal code. As a result, they were directed to the VCM, which thus had a much higher number of participants than the other mechanisms. We subsequently revised the user assignment to alleviate this problem. The revision was put into production on November 10, 2004. Under the revised software, rather than sending all users without a valid U.S. or Canadian postal code to the VCM, these users were randomized among all four mechanisms. Users who entered valid U.S. or Canadian postal codes were still presented with the message based on the original design.

Even though the pop-up window was designed to appear unrelated to fund-raising or an experiment, and for the second half of campaign 2004 and all of campaign 2005 all users were treated the same, we do not rule out the possibility that there might be a sample selection bias from the different handling of those who complied with the pop-up window and the much larger number who did not.

2.2 Campaign 2005

Campaign 2005 was launched on April 20, 2005 and ended on June 30, 2005. We revised the assignment of users to mechanisms based on lessons learned from the 2004 campaign.

The major revision made was to eliminate all pop-up and pop-under windows, including the zip window and the pop-under window with fund-raising messages, as they were not effective in the 2004 campaign (see Section 4 for a discus-

sion of the problems we encountered with pop-up windows).

In the new design, a user would see a campaign message only when she voluntarily clicked on the campaign button on the IPL home page or the small horizontal campaign link located on each of the subject pages (again, except the KidSpace and TeenSpace). Users were assigned to each mechanism randomly and uniformly. Notice that we no longer could control the demographic balancing of our treatments, and thus we rely entirely on randomization for this balancing.

Cookies were again used to track users, so one user could see only one message consistently. Users who participated in the 2004 campaign were again identified through their cookies and shown the the same message they had seen in 2004. Users whose computers did not allow cookies were assigned to VCM.

The solicitation messages were largely the same, except for two minor updates. First, we changed campaign ending date from December 30, 2004 to June 30, 2005. Second, the introductory paragraph was revised slightly to incorporate the 10th anniversary celebration of the IPL.

On the IPL homepage, we use a new button that presented a logo with the text "Make an anniversary donation" (see Appendix A). The small campaign link on each of the IPL subject pages was made more prominent. Other design features of our experiment remained the same as for Campaign 2004.

3 Results.

We raised a total of $1,128 from 24 different individuals, $903 in the 2004 and $225 in 2005. We now analyze contribution rate, gift size, and click behavior. By way of comparison, in a recent seed money fund-raising experiment List and Lucking-Reiley (2002) raised $5509 from 183 donors through a direct mail campaign for Central Florida University, for an average donation of $30. We obtained a smaller sample of contributors, but an average donation of $47.

We present in Table 3 the total number of users who saw screen 1 of a solicitation message, the number of contributors, as well as the contribution rate for the 2004 and 2005 campaigns for each mechanism, as well as the overall contribution rate after combining the two campaigns (last column). The top panel includes all data from 2004 (2004All) and 2005, while the lower panel includes the 2004 data after the November 10 redesign (2004Red) and all of the 2005 data. As nearly 90% of the users ignored the postal code pop-up window in 2004, an overwhelming number of users were directed to VCM until we revised the design mid-way through the campaign. In comparison, in 2005, the number of users directed to each mechanism roughly follow a uniform distribution. Overall, the contribution rate was quite low, with only one out of every 6,252 users who saw a campaign

Table 3: Contribution Rate by Mechanism

Mechanism	#Users saw screen 1		# Contributors		Contribution Rate (%)		
Alll Data	2004All	2005	2004All	2005	2004All	2005	Combined
VCM	122,573	1,564	14	1	0.01	0.06	0.01
Premium	6,921	1,568	3	1	0.04	0.06	0.05
Seed	7,090	1,529	1	1	0.01	0.07	0.02
Matching	7,281	1,514	1	2	0.01	0.13	0.03
Total	143,865	6,175	19	5	0.01	0.08	0.02
Post Red.	2004Red	2005	2004Red	2005	2004Red	2005	Combined
VCM	2,269	1,564	4	1	0.18	0.06	0.13
Premium	2,304	1,568	1	1	0.04	0.06	0.08
Seed	2,322	1,529	1	1	0.04	0.07	0.05
Matching	2,423	1,514	0	2	0.00	0.13	0.05
Total	9,318	6,175	6	5	0.06	0.08	0.08

message making a contribution.[16] The contribution rate does appear to vary across mechanisms, however; we now turn to a statistical analysis of these data.

Table 4: P-values for Two-sample Fisher Exact Tests of Equality of Contribution Rates

Alternative Hypothesis	2004All	2004Red	2005	CombinedAll	CombinedRed
VCM $<$ Premium	0.06	0.82	0.75	0.03	0.78
VCM $<$ Seed	0.43	0.82	0.26	0.30	0.78
VCM $<$ Matching	0.42	0.95	0.49	0.11	0.78
Seed $\neq$ Premium	0.37	1.00	1.00	0.45	1.00
Matching $\neq$ Premium	0.36	0.49	0.62	0.72	1.00
Seed $\neq$ Matching	1.00	0.49	0.62	1.00	1.00

Note: The p-values in the first three rows are calculated for one-sided alternative hypotheses, while those in the last three rows are calculated for two-sided alternative hypotheses.

[16] A direct mail field experiment on nonprofit fundraising achieved about a 1% contribution rate (Katzev 1995). According to the Direct Marketing Association, average response rates to direct mail fundraising campaigns are about 4.5% (Zeller and Chernis 2004). List and Lucking-Reiley (2002) obtained about a 6% contribution rate. One problem with our contribution statistics is that many users of IPL are youth, and only users over the age of 18 were permitted to make donations. More generally, we have not found reliable estimates of typical donation rates for online nonprofit campaigns against which to compare our results.

Result 1 (Contribution Rate) *Using all contribution data, the Premium mechanism generates significantly higher contribution rate than VCM. None of the other pairwise comparisons between mechanisms is significant.*

Support 1 *In Table 3 we present the alternative hypotheses and the p-values for the two-sample Fisher exact tests for the 2004 (all data and post redesign), 2005 as well as the combined campaigns. The null hypothesis is that the contribution rates under each pair of mechanisms are equal. For the combined data set, contribution rate under VCM $<$ Premium at a p-value of* 0.03*. None of the other pairwise comparisons is significant at the 5-percent level.* ∎

As we discussed in Section 2, when we use all data, results might be subject to a sample selection bias. When restricted to post-redesign data, we can not reject the null hypotheses that contribution rate is equal between any pair of mechanisms. When each campaign is analyzed separately, the Premium mechanism generates a (weakly) significantly higher contribution rate than the VCM ($p = 0.06$) in 2004 using all data. The Fisher exact tests can not reject null hypotheses that any other pairs of mechanisms yield the same contribution rate. In campaign 2005, we can not reject the null hypotheses that contribution rate is equal between any pair of mechanisms, primarily because the sample of contributors is very small.

Table 5: Gift Size: Two-Sample Wilcoxon Rank Sum Test on the Gift Size Between Pairs of Mechanisms

All Data	# Contr.	Tot. Contr.	Mean Gift Size	H_1	p-value
VCM	15	\$805	\$53.67	VCM $<$ Premium	0.92
Premium	4	\$120	\$30.00	VCM $<$ Seed	0.88
Seed	2	\$50	\$25.00	VCM $<$ Matching	0.55
Matching	3	\$153	\$51.00	Premium $\neq$ Seed	1.00
				Matching $\neq$ Premium	0.59
Total	24	\$1,128	\$47.00	Matching $\neq$ Seed	0.55
Post Red.	# Contr.	Tot. Contr.	Mean Gift Size	H_1	p-value
VCM	5	\$250	\$50.00	VCM $<$ Premium	0.83
Premium	2	\$75	\$37.50	VCM $<$ Seed	0.33
Seed	2	\$50	\$25.00	VCM $<$ Matching	0.31
Matching	2	\$150	\$75.00	Premium $\neq$ Seed	0.32
				Matching $\neq$ Premium	0.22
Total	11	\$525	\$47.73	Matching $\neq$ Seed	0.10

We report the number of contributions, total contribution, and average gift size under each mechanism in Table 3. Again, the top panel presents all data, while the bottom panel presents data after the redesign. The last two columns present

the alternative hypotheses and the corresponding p-values from the two-sample Wilcoxon rank sum test on the gift size between pairs of mechanisms. We pool data from the two campaigns. Because of the skewed distribution of traffic in 2004, VCM generates the highest total dollar contribution. Partly due to the small sample size for each mechanism, the Wilcoxon rank sum tests can not reject the null hypotheses that any pair of mechanisms yield the same gift size conditional on giving.

Result 2 (Gift Size) *There is no significant difference between any pair of mechanisms in the gift sizes they generate.*

Support 2 *In Table 3 we present the alternative hypotheses and the corresponding p-values from the two-sample Wilcoxon rank sum test on the gift size between pairs of mechanisms. None of the pairwise comparisons is significant at the 5-percent level.* ∎

Result 2 may well be due to the fact that the number of contributions is very small under each mechanism. Both List and Lucking-Reiley (2002) and Rondeau and List (2005) found that the Seed mechanism generated significantly larger gift than VCM in a conventional mail campaign.

We now turn to user click behavior. Web communications are based on a client-server. The user is the client, and uses client software (a "browser") to send messages (in the HTTP protocol, or language) to a server, which is a software program usually running on a different computer that responds to the messages by returning information stored on that computer. For example, each time a user clicks on a link (or button or image, etc.) displayed by a browser, the browser sends a message to the Web server indicating that the user clicked that token and requesting that the server respond in the appropriate way. Thus, a "click" is a conscious behavior by a user to make a request or signal an action to a server. Further, the server observes the clicks, and can record the sequence of clicks and the semantic meaning associated with them (for example, "display additional information" or "start the donation process"). Clickstream data is a major new source of observable user behavior, and researchers (as well as marketing departments) actively collect and try to interpret user preferences and intentions from their click sequences.

As an example of the opportunities afforded by clickstream data, consider usage of scholarly (or other) publications: when journals are printed on paper and shelved in a library, it is extremely difficult for a publisher to determine which articles are in fact being viewed, and by whom. When articles are published electronically and accessed from a publisher's server, such as Elsevier's ScienceDirect (http://www.sciencedirect.com), the publisher can measure viewing with much less noise.

However, although clickstream data clearly reveals certain choices made by users, inferences about user preferences or intentions may be difficult. In the electronic publishing example, although clicking to display a journal article on one's screen is probably a fairly accurate measure of whether something about that article was *viewed*, it may be a much less accurate measure of whether, or how much of that article was *read*. Perhaps the title was catchy but the user realized from the abstract that the article was not of interest to her. We now discuss some of the user clickstream data we collected, and our findings about the caution that field experiments must employ when interpreting clickstream data.

In both campaigns, we were able to track user click data until the user departed from IPL by closing the browser window, going to a different website, or by clicking on the final step in our campaign system: the "Contribute now" button. Recall that the donation process is *not* complete when a user clicks "Contribute now"; rather, the user then is delivered to the University Giving page, where donations are actually processed by filling out a set of forms. In Table 3 we present the total number of users who saw screen 1 of a solicitation message, the number of users who clicked all the way through the "Contribute now" link, and the calculated click-through rates for the separate and pooled 2004 and 2005 campaigns. Both campaigns 2004 and 2005 yielded a consistent pattern with respect to the ranking of the click-through rates among the four mechanisms.

Table 6: Clicking "Contribute Now" Rate Across Mechanisms

Mechanism	#Users saw screen 1		# Clickers on "Contribute now"		Clicking Rate (%)		
All Data	2004All	2005	2004All	2005	2004All	2005	Comb.
VCM	122,573	1,564	250	57	0.20	3.64	0.25
Premium	6,921	1,568	7	33	0.10	2.10	0.47
Seed	7,090	1,529	26	51	0.37	3.34	0.89
Matching	7,281	1,514	28	51	0.38	3.37	0.90
Total	143,865	6,175	311	192	0.22	3.11	0.34
Post Red.	2004Red	2005	2004Red	2005	2004Red	2005	Comb.
VCM	2,269	1,564	4	57	0.18	3.64	1.59
Premium	2,304	1,568	1	33	0.04	2.10	0.88
Seed	2,322	1,529	4	51	0.17	3.34	1.43
Matching	2,423	1,514	6	51	0.25	3.37	1.45
Total	9,318	6,175	15	192	0.16	3.11	1.34

Result 3 (Click-through Rate) *Using all data, the Premium, Seed and Matching mechanisms each generate significantly higher rate of clicking through the "Con-*

tribute now" link than the VCM mechanism, while the Seed and Matching mechanisms each generate significantly higher rate of clicking through than the Premium mechanism. The latter holds when we restrict the sample to post redesign data.

Table 7: Clicking "Contribute Now": P-values for Two-sample Fisher Exact Tests

Alternative Hypothesis	2004All	2004Red	2005	CombinedAll	CombinedRed
VCM<Premium	0.97	0.82	0.99	0.00	1.00
VCM < Seed	0.01	0.62	0.64	0.00	0.69
VCM<Matching	0.00	0.42	0.62	0.00	0.66
Seed ≠ Premium	0.00	0.37	0.04	0.00	0.03
Matching ≠ Premium	0.00	0.13	0.04	0.00	0.02
Seed ≠ Matching	0.89	0.75	1.00	1.00	1.00

Support 3 *We present in Table 3 the alternative hypotheses and the p-values for the two-sample Fisher exact tests of equality of proportions for the 2004, 2005 as well as the combined campaigns. The null hypothesis is that the click-through rates under each pair of mechanisms are equal. For the combined data set, click-through rate under VCM < Premium, VCM < Seed, VCM < Matching, Premium ≠ Seed, and Premium ≠ Matching at a p-value of* 0.00 *respectively. None of the other pairwise comparisons is significant at the 5% level. Restricting the sample to post redesign data, we get Premium ≠ Seed at a p-value of* 0.03*, and Premium ≠ Matching at a p-value of* 0.02*.* ∎

Result 3 indicates that the click-through rate is significantly higher under the Seed and Matching mechanism than under the VCM and Premium mechanisms. For the combined data, the click-through rate under Premium is significantly higher than under VCM. Meanwhile, we cannot reject the null hypothesis that the click-through rate under Seed and Matching is the same. This result holds separately for the 2004 campaign, and generally holds separately for the 2005 campaign as well.

Comparing the click-through rates of 2004 and 2005, we note that, under each mechanism, the rates are much higher in 2005 ($p < 0.01$ for every pairwise comparison between 2004 and 2005). This might be due to the fact that in the 2005 campaign only those users who affirmatively clicked on the side bar "Give to IPL" button were shown the message; in 2004 all users were shown a pop-under window with the campaign message (if they did not have a pop-up blocker running). Therefore, 2005 participants who saw screen 1 were more active or curious than the 2004 users who might have seen screen 1 passively.

We are aware of no theoretical results that consider why one contribution mechanism might consistently draw more curiosity than another. To the extent that curiosity is a driving factor for our result, we find it somewhat surprising that participants were more inclined to click through for the Seed and Matching messages than for the Premium message. The Premium treatment is the only one that promised a direct private incentive for donating, and our prior was that, if anything, the chance to obtain some premium item would stimulate more curiosity. Since we cannot distinguish between the *actual* contribution rates or gift sizes for Premium, Seed and Matching, we do not offer concrete advice on this issue for fundraisers, but the possibility of a "curiosity" difference between mechanisms merits further research.

4 Discussion: Online Experiment Methodology.

We set out, in part, to develop methods for embedded online field experiments. By "embedded" we mean experiments that are run within production online services, not set up as stand-alone experiments that happen to use the Internet as a communications medium. There have been a number of stand-alone experiments that have used the Internet as a medium (Lucking-Reiley (1999) is one of the early, better known examples); there have been fewer attempts to embed experiments in production environments. In this section we summarize several lessons we learned about embedded online experimental design.

4.1 Subject Identification / Authentication

One important challenge for any online experiment is to ensure that each user only sees one mechanism. This is especially likely to be difficult for embedded experiments, since many information services do not use an identification / authentication system (such as login plus password). Indeed, in our case users do not need to log in to get access to the IPL. We used cookie technology to create unique, persistent identifiers to trace users, solely for the purpose of ensuring that on each return visit they see the same fund-raising mechanism.

Our method is not failsafe, however. We are aware of at least three failure modes. First, some users configure their browsers to block cookies. We dealt with this problem by assigning such users to a single mechanism (the control, VCM) which ensures consistency but is a source of potential unbalancing of our control sample.

Second, the cookie originally written on a participant's machine may be intentionally or inadvertently deleted by the participant before a return visit to IPL. We have no way to identify such instances: when we look for a cookie and don't

find one, we assume the user is making her first visit to IPL since the start of the campaign, and we assign her to a treatment — but quite possibly a different treatment than the last time she visited.

Yet a third failure mode is that the participant might visit the IPL from more than one client computer. Since cookies are stored on the local client computer, we would not find the cookie if the user made a subsequent visit from a different computer, and we would again treat the user as a first-time visitor and assign a — possibly different — treatment.

We considered an alternative to cookies which has some advantages, but we concluded that on balance it was inferior: identify users by the IP address of the computer they use to connect to IPL. The advantage is that the user cannot easily block the IP address from being read by our software. For machines with so-called "static" IP addresses, the user cannot easily delete or change the IP address before a subsequent visit. However, many machines are configured with "dynamic" IP addresses which can change relatively frequently. Other machines are located behind a firewall, and the IP address that they "publish" is actually the address of the firewall, and may represent a few to thousands of different machines. Likewise, a single machine may be used by multiple different users, who would then all have the same IP address.

Perhaps the most reliable method for identification and authentication at reasonable cost is to use a login plus password system. It is relatively simple for the experimenter to install such a system for an embedded experiment if the information service does not already use such a system. However, as we found in our IPL experiment, the service provider might be unwilling to allow the experimenter to implement a login system because it rather seriously changes the user experience of the service, and generally for the worse. Field experiment evidence suggests that such fears are well-founded: login plus password systems may impose substantial non-pecuniary user costs and have marked negative effects on usage; see the discussion in subsection 4.4 below.

A related design solution would be to choose as a host for the experiment an information service that already uses a login plus password system. However, this may not be feasible, or may rule out especially attractive opportunities, but the experimental results may not generalize well to the greater number of information services that do not require login authentication, especially given the evidence that the non-pecuniary user costs of login systems may be quite high.[17]

It may also be worth noting that no standard method of identification and

[17]Of course, login plus password systems are not failsafe either; there are several failure modes. Discussing them in detail is beyond the scope of this article. In short, they are similar to, but generally less severe than the failure modes for cookies.

authentication for experiments implement in other environments — e.g, mail, phone or laboratory campaigns — is proof against a participant repeating the experiment under a different identity. The problems seem to be more numerous and harder to solve for online experiments, but it is unreasonable to expect that any system will be perfect.

4.2 Balanced Assignment to Control and Treatment Groups

It is often important to gather information about experimental subjects in order to assign them to control and treatment groups in a way that balances the distribution of various characteristics, in order to control for spurious correlations and to render the results generalizable. Balanced assignment of subjects to groups can be more challenging for experiments conducted on the Internet than those conducted in the brick-and-mortar world. As we discussed above, the experimenter may be able to obtain little or no information identifying the individual subjects. Service providers hosting embedded experiments may not be willing to let experimenters mandate standard demographic pre-tests due to the adverse impact on user experience.

As described in Section 2.1.1 above, we used the web technology for generating pop-up windows to implement a simple method for balancing subject assignment. Our method imposed minimal cost on potential participants (they were asked merely to enter three digits and make one click), and was not privacy-invasive (no individual identifying information can be gleaned from only the first three digits of a postal code).

There are three main failure modes for our method. First, users could simply ignore the popup window, closing it without entering any information. Second, the pop-up window could be blocked altogether (so it was not even displayed) using pop-up blocking software that at that time had become a well-known feature for some browsers, and add-in available for others. Finally, the participant could enter three digits of a valid postal code, but they might not be the participant's true postal code, resulting in (possible) mis-assignment.

Out of the $155,591$ users tracked by our program in the 2004 campaign (when we used the pop-up method), $138,694$ users (89.14%) did not enter anything in the postal code window. The lack of entry could be due to either the use of a pop-up blocker, or user unwillingness to enter the digits. We are unable to distinguish between these two failure modes in our data.

In the 2004 campaign 683 users (0.44%) entered invalid U.S. or Canadian postal code digits. This number probably does not indicate very many (intentional or unintentional) errors because we recorded a valid postal code from another country as invalid, and IPL receives a sizable number of visitors from outside of North

America (indeed, two of the 24 contributors were from Germany and one from Egypt.).

Only 16,214 users (10.42%) entered a valid U.S. or Canadian postal code. Of these, some unknown fraction, while valid, may not have been the participant's true postal code. For users from the U.S. who actually made a donation, we are able to compare the mechanism under which they made the contribution with the mechanism that should have been assigned to them had they entered their true postal code information, since the true code is generally recorded in the contribution dataset.[18] Surprisingly, among the sixteen users from the U.S. who contributed in 2004, ten of them came from a different mechanism from the one that should have been assigned had they entered the true information in the postal code pop-up window. We conjecture that one reason for the very high rate of non-compliance with our request for the first three digits of postal codes may have been that users perceived there to be some non-negligible risk that this information could be used to identify them or otherwise invade their privacy.

We conclude that our pop-up window method for obtaining the geographic information needed to implement demographic balancing proved ineffective. In the 2005 campaign we relied solely on randomization to balance our samples.

4.3 Interpreting Clickstream Behavior

Since the emergence of the World Wide Web as a widespread phenomenon in 1994, many researchers have mined the server logs from web sites for "clickstream" data, and then tried to draw inferences from the clicks users make (that is, the affirmative choices to navigate to different locations, or to initiate some other action, such as processing a form input). In economic jargon, since clicking is an affirmative action, it reveals preference, and it might be possible to draw inferences from participants' revealed preferences. However, the experimenter must interpret *for what* the user has revealed a preference. Our data tell a cautionary tale on the interpretation of clickstream data.

One of the critical steps in our contribution solicitation was a link that an interested participant would reach that was labeled "Contribute now". Generally, only by clicking this link would a participant reach the University Giving online system through which she could process a donation.[19] What our data reveal,however, is

[18]Donors must provide a valid address to authenticate their credit card, and are asked to provide an address for a follow-up if they are making a check donation. Though the user could like in the latter case, this seems unlikely since the user has already decided to make a donation, and generally will give up address information by sending a check in any case.

[19]There are other ways to reach the University Giving system, but none were identified in the IPL website, so a user would have to have other knowledge about how to get around the "Contribute

that clicking "Contribute now" more often than not did not mean the participant would make a donation.

In particular, we observed dramatically higher rates of clicking "Contribute now" than the rate of actually contributing across all mechanisms. In campaign 2004, only 6.1% of users who clicked through the "Contribute now" link *actually* contributed. The rate in campaign 2005 was 2.6%. Across both campaigns the percentage was 4.8%. The differences are statistically significant for each mechanism separately by campaign, and pooled across campaigns ($p < 0.01$ for both one-tailed and two-tailed Fisher exact tests for all mechanisms by campaign and pooled campaign, except in Premium 2004All where the one-tailed p-value is 0.17 and the two-tailed is 0.34, and all mechanisms in 2004Red).

That is, whatever preference the participant was revealing by clicking "Contribute now", it apparently was not often a preference actually to donate. We conclude that great caution is required in interpreting clickstream data.

4.4 Interface Design

We noted above that only 5.2% of participants who affirmatively clicked on the "Contribute now" in fact completed a contribution. One possible explanation is that the design of the University Giving system interface, for example, the number of screens, length of forms, etc., imposes non-negligible cost that discourages participants from completing a donation that they did intend to make. We call this an "interface cost" effect.

Our pretesting indicated that, after users clicked "Contribute now" and entered the University Giving page, it took five to six minutes on average for someone who was fairly familiar with the giving system to complete the entire contribution process. In a field experiment on pricing online access to scholarly journals, Gazzale and MacKie-Mason (Forthcoming 2006) found that non-pecuniary user costs (such as the inconvenience of obtaining a free password and using it) decreased purchases of $7 articles by as much as 90%.

An alternative hypothesis is that many participants simply wanted to satisfy their curiosity about what would happen when they clicked "Contribute now", but they never intended to actually contribute: we call this the curiosity effect.

In our present experiment, since we were unable to instrument the participant's interaction with the University Giving system, we have no observational evidence to measure the relative importance of the curiosity and interface cost effects. However, taken together with Gazzale and MacKie-Mason (Forthcoming 2006), we

---

now" link. In any case, if any participant did find different means to get to the University Giving system, the resulting donation would not be recorded as associated with our campaign.

view the result as a strong hint that fundraisers should devote considerable attention to non-pecuniary users costs when designing their giving systems.

5 Conclusion.

We set out with a dual agenda: to test various fund-raising mechanisms for public goods, and to develop and evaluate methods for conducting online field experiments embedded in production services.

We are constrained by the small voluntary donation sample from obtaining statistically significant evidence on all of the mechanism design hypotheses for which the experiment was designed. Our design was somewhat ambitious: we implemented a control (VCM) and three treatments (Seed, Matching, Premium), and thus needed a substantial number of observations in each cell. One suggestion to researchers is to embed an experiment in a setting with a sufficient previously demonstrated level of activity to support the design. For example, we implemented a fund-raising campaign in a service with no track record for online fund-raising. We conjectured, incorrectly, that given the strong reputation of the IPL and the very high daily usage rate, we would obtain a reasonably large sample.[20]

Nonetheless, we did obtain significant evidence on several interesting questions. First, when we use all the data, we find that the Premium mechanism increases the fraction of participants who make a donation. This result, though, might be subject to a sample selection bias. Second, the Seed and Matching mechanisms each generated a higher rate of interest as measured by click-through to the "Contribute now" page than the Premium mechanism. Our results on gift size were insignificant due to the small sample sizes: we could not reject the null hypothesis that the gift size was the same for every pairwise mechanism comparison at even the 10% level of significance.

On our second goal of developing and evaluating methodology we learned several valuable lessons (some from successes, some from failures). First, participant identification and authentication, to ensure repeat visitors are presented with only one treatment, is a thorny problem. We considered several methods and implemented a relatively straightforward cookie tracking system that appeared to work reasonably well, though it is known to have some weaknesses that cannot be overcome. The most effective system with reasonable costs is login plus password, but this is unlikely to be acceptable for embedded experiments unless the hosting information service already uses such a system.

Second, our efforts to implement balanced assignment to control and treat-

[20] In fact, the IPL did try to implement one previous online fund-raising campaign, which also did not obtain many donors, but that was more than five years ago, when the service was not as well known and without the design support of a professional fund-raiser.

ment groups were stymied. The common approach of requiring subjects to complete a demographic survey before assignment will not be feasible in many production environments.[21] Our effort to obtain coarse granularity geographic information (through partial postal codes) was rejected by over 80% of participants, possibly due to (unwarranted) privacy concerns, so we had to fall back on pure randomization.

We also discovered strong evidence that care must be used when interpreting the preferences revealed by clickstream behavior: only about 5% of participants who clicked through "Contribute now" in fact made a contribution.

Our evidence on the final methodological concern is consistent with evidence from other sources: the non-pecuniary user costs of poor interface design may substantially affect participant behavior. In our case, the complexity of the donation forms may partly explain the rather low donation rate, particularly when compared to the much higher rate of "Contribute now" click-through. Planning for and allaying participant concerns about privacy invasion, even when unwarranted, may be important to obtain participation and protocol compliance.

Our summary of the methodological lessons is simple: whether the features of the online infrastructure are well suitable for the experimenter's design needs should enter as a nontrivial factor in the pre-experiment decision function. Particularly for field experiments embedded in production services, it may be wise to seek a service host with infrastructure that is already adapted to the experimenter's needs, since most production services will object to introducing intrusive new infrastructure to instrument and control an experiment.

In short, the Internet offers wonderful opportunities to aggregate participants and to implement computer-based experiments embedded in field production services, but the available technologies also confront the experimenter with imperfectly resolvable challenges. We hope that lessons learned from this project will help future embedded field experiments on the Internet.

[21] One exception we have seen arises when experimental subjects are employees of the hosting organization, and the hosts are willing to require the employees to provide demographic information (Bulkley and Van Alstyne 2005).

Appendix A. Screen Shots

Figure 1: Campaign 2004: Postal Code Pop-Up Window and "Friends of the IPL" Button on IPL Homepage

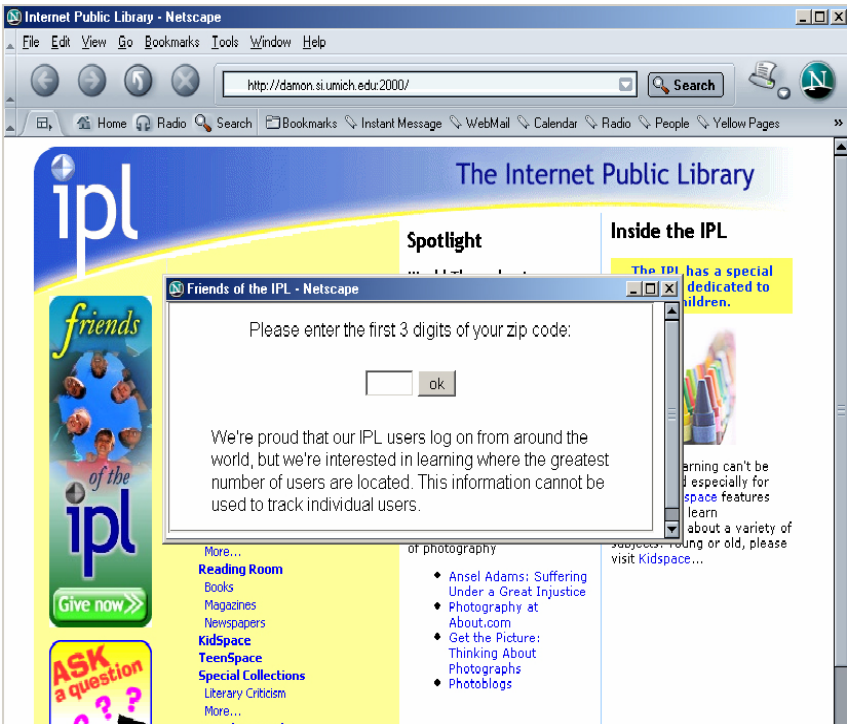
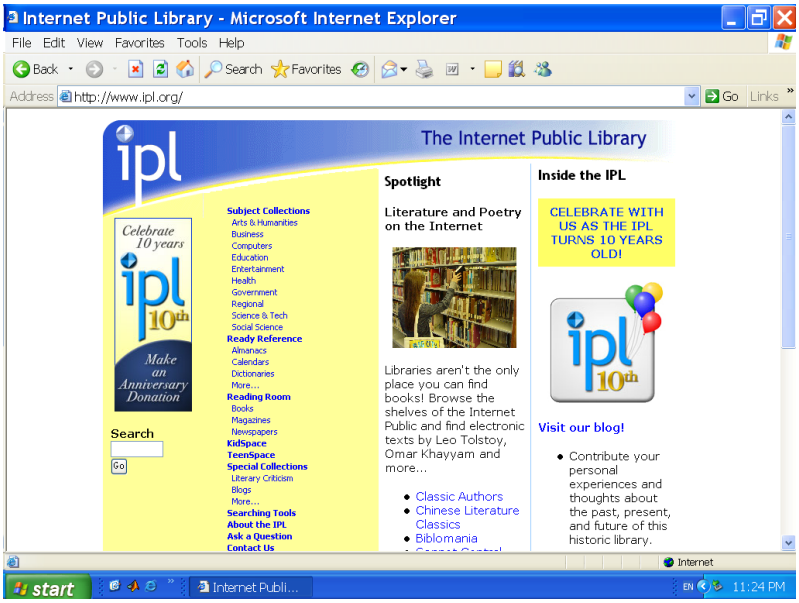


Figure 2: Campaign 2005: "Make an Anniversary Donation" Button on IPL Homepage and Horizontal Link on a IPL Subject Page



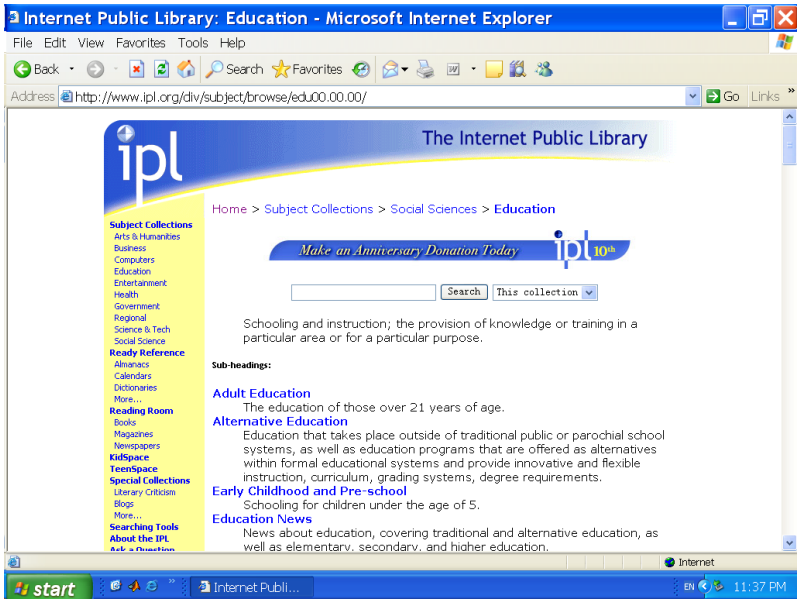
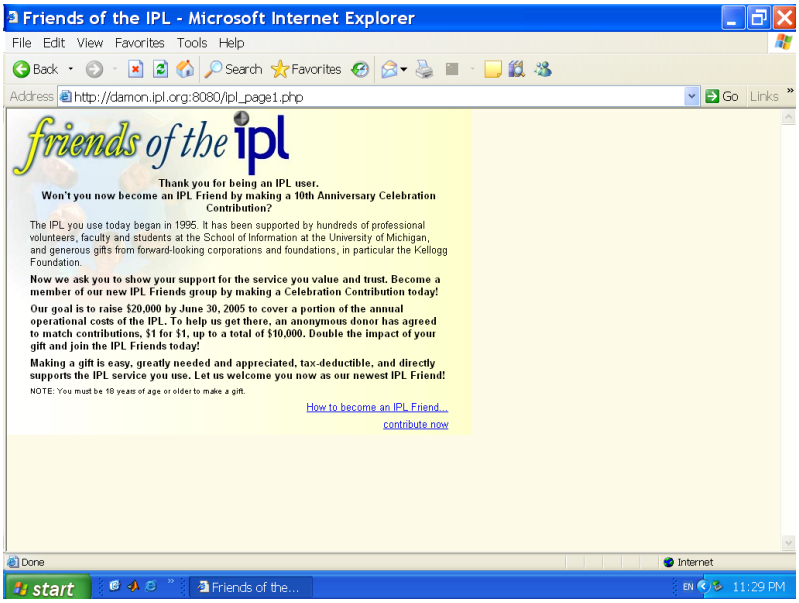


Figure 3: Campaign 2005: Solicitation Message Screens 1 and 2



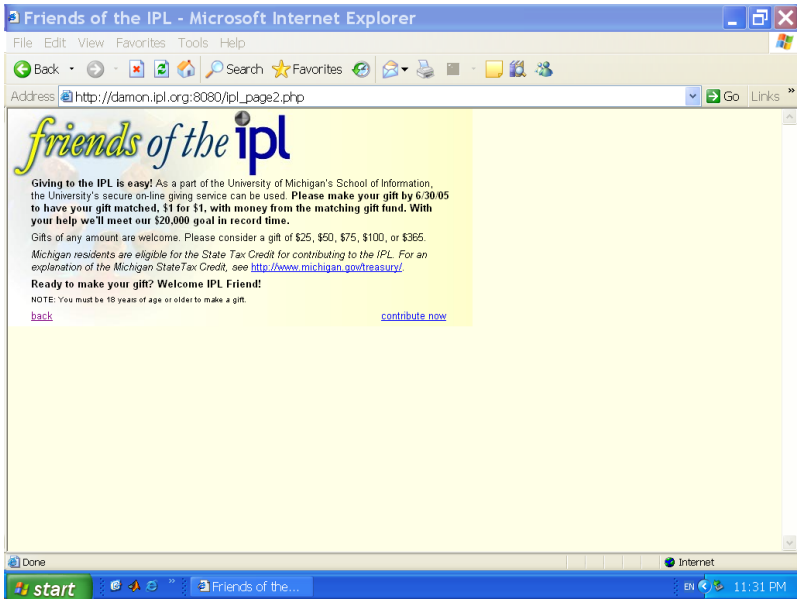
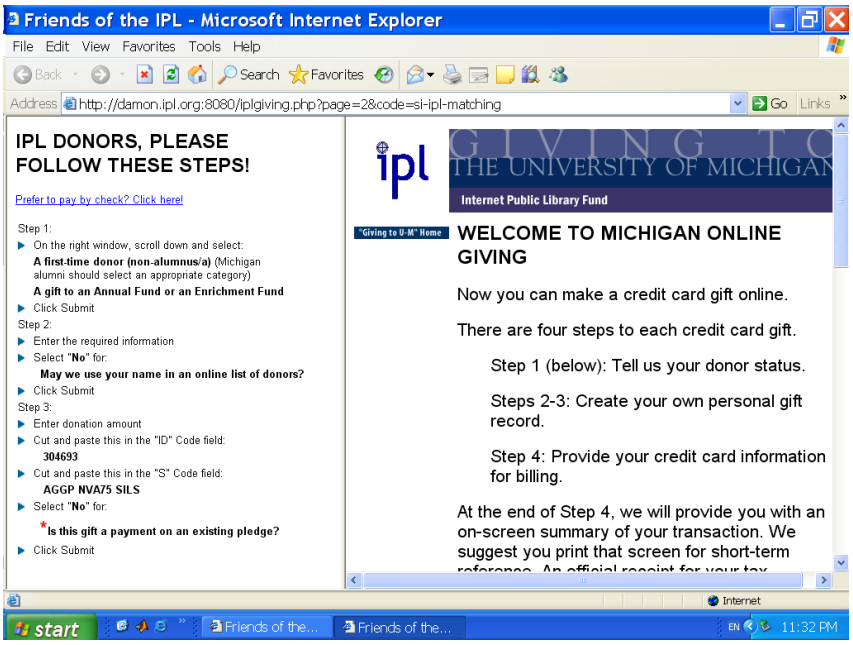


Figure 4: Campaign 2005: Solicitation Message Screen 3 - University Giving Page



Appendix B. Solicitation Messages: Campaign 2004

We include the solicitation messages for Campaign 2004. Those for Campaign 2005 are identical except for the ending dates, and the header, which says "Won't you now become an IPL Friend by making a 10th Anniversary Celebration Contribution?"

[VCM-screen 1]

Thank you for being an IPL user.
Won't you now become an IPL Friend?

The IPL you use today began in 1995. It has been supported by hundreds of professional volunteers, faculty and students at the School of Information at the University of Michigan, and generous gifts from forward-looking corporations and foundations, in particular the Kellogg Foundation.

Now we ask you to show your support for the service you value and trust. Become a member of our new IPL Friends group by making a contribution today!

Our goal is to raise $20,000 by December 30, 2004, to cover a portion of the annual operational costs of the IPL. Making a gift is easy, greatly needed and appreciated, tax-deductible, and directly supports the IPL service you use. Let us welcome you now as our newest IPL Friend!

NOTE: You must be 18 years of age or older to make a gift.

How to become an IPL Friend
Contribute now

[VCM - screen 2]

Giving to the IPL is easy! As a part of the University of Michigan's School of Information, the University's secure on-line giving service can be used. **Please make your gift by 12/30/04 to receive full tax-deductible benefits for 2004 ··· and to help us reach our goal of $20,000 for the IPL!**

Gifts of any amount are welcome. Please consider a gift of $25, $50, $75, $100, or $365.

Michigan residents are eligible for the State Tax Credit for contributing to the IPL. For an explanation of the Michigan State Tax Credit, see http://www.michigan.gov/treasury/.

Ready to make your gift? Welcome IPL Friend!

NOTE: You must be 18 years of age or older to make a gift.

back Contribute now

[Premium-screen 1]

Thank you for being an IPL user.
Won't you now become an IPL Friend?

The IPL you use today began in 1995. It has been supported by hundreds of professional volunteers, faculty and students at the School of Information at the University of Michigan, and generous gifts from forward-looking corporations and foundations, in particular the Kellogg Foundation.

Now we ask you to show your support for the service you value and trust. Become a member of our new IPL Friends group by making a contribution today!

Our goal is to raise $20,000 by December 30, 2004, to cover a portion of the annual operational costs of the IPL. All Friends contributing $25 or more will receive a special IPL thank you gift:

- **IPL mouse pad ($25-$49)**
- **IPL book light flashlight ($50-$74)**
- **IPL CD case ($75 or above)**

Making a gift is easy, greatly needed and appreciated, tax-deductible, and directly supports the IPL service you use. Let us welcome you now as our newest IPL Friend!

NOTE: You must be 18 years of age or older to make a gift.

How to become an IPL Friend
Contribute now

[Premium - screen 2]

Giving to the IPL is easy! As a part of the University of Michigan's School of Information, the University's secure on-line giving service can be used. **Please make your gift by 12/30/04 to receive full tax-deductible benefits for 2004 ··· and to help us reach our goal of $20,000 for the IPL!**

Gifts of any amount are welcome. Please consider a gift of $25, $50, $75, $100, or $365.

Michigan residents are eligible for the State Tax Credit for contributing to the IPL. For an explanation of the Michigan State Tax Credit, see
http://www.michigan.gov/treasury/.

Ready to make your gift? Welcome IPL Friend!

NOTE: You must be 18 years of age or older to make a gift.

back Contribute now

[Seed-screen 1]

Thank you for being an IPL user.
Won't you now become an IPL Friend?

The IPL you use today began in 1995. It has been supported by hundreds of professional volunteers, faculty and students at the School of Information at the University of Michigan, and generous gifts from forward-looking corporations and foundations, in particular the Kellogg Foundation.

Now we ask you to show your support for the service you value and trust. Become a member of our new IPL Friends group by making a contribution today!

Our goal is to raise $20,000 by December 30, 2004, to cover a portion of the annual operational costs of the IPL. We're already half way there with a gift of $10,000! Won't you help us with the final $10,000? Your gift in any amount will make a big difference.
Making a gift is easy, greatly needed and appreciated, tax-deductible, and directly supports the IPL service you use. Let us welcome you now as our newest IPL Friend!

NOTE: You must be 18 years of age or older to make a gift. How to become an IPL Friend

Contribute now

[Seed - screen 2]

Giving to the IPL is easy! As a part of the University of Michigan's School of Information, the University's secure on-line giving service can be used. **Please make your gift by 12/30/04 to receive full tax-deductible benefits for 2004 ··· and to help us raise the remaining $10,000 of our $20,000 goal.**

Gifts of any amount are welcome. Please consider a gift of $25, $50, $75, $100, or $365.

Michigan residents are eligible for the State Tax Credit for contributing to the IPL. For an explanation of the Michigan State Tax Credit, see
http://www.michigan.gov/treasury/.

Ready to make your gift? Welcome IPL Friend!

NOTE: You must be 18 years of age or older to make a gift.

back Contribute now

[Matching-screen 1]

Thank you for being an IPL user.
Won't you now become an IPL Friend?

The IPL you use today began in 1995. It has been supported by hundreds of professional volunteers, faculty and students at the School of Information at the University of Michigan, and generous gifts from forward-looking corporations and foundations, in particular the Kellogg Foundation.

Now we ask you to show your support for the service you value and trust. Become a member of our new IPL Friends group by making a contribution today!

Our goal is to raise $20,000 by December 30, 2004, to cover a portion of the annual operational costs of the IPL. To help us get there, an anonymous donor has agreed to match contributions, $1 for $1, up to a total of $10,000. Double the impact of your gift and join the IPL Friends today!

Making a gift is easy, greatly needed and appreciated, tax-deductible, and directly supports the IPL service you use. Let us welcome you now as our newest IPL Friend!

NOTE: You must be 18 years of age or older to make a gift.

How to become an IPL Friend
Contribute now

[Matching - screen 2]

Giving to the IPL is easy! As a part of the University of Michigan's School of Information, the University's secure on-line giving service can be used. **Please make your gift by 12/30/04 to receive full tax-deductible benefits for 2004 ··· and to have your gift matched, $1 for $1, with money from the matching gift fund. With your help we'll meet our $20,000 goal in record time.**

Gifts of any amount are welcome. Please consider a gift of $25, $50, $75, $100, or $365.

Michigan residents are eligible for the State Tax Credit for contributing to the IPL. For an explanation of the Michigan State Tax Credit, see
http://www.michigan.gov/treasury/.

Ready to make your gift? Welcome IPL Friend!

NOTE: You must be 18 years of age or older to make a gift.

back Contribute now

References

James Andreoni. The economics of philanthropy. In Serge-Christophe Kolm and Jean Mercier-Ythier, editors, *Handbook on the Economics of Giving, Reciprocity and Altruism*, volume 2. Elsevier/North-Holland, Amsterdam, 2006 (forthcoming).

— —. Toward a theory of charitable fund-raising. *Journal of Political Economy*, 106(6):1186–1213, 1998.

— — and Ragan Petrie. Public goods experiments without confidentiality: A glimpse into fund-raising. *Journal of Public Economics*, 88(7-8):1605–1623, 2004.

Mark E. Bagnoli and Bart L. Lipman. Provision of public goods: Fully implementing the core through private contributions. *Review of Economic Studies*, 56 (4):583–601, October 1989.

Nat Bulkley and Marshall Van Alstyne. Climbing the occupational ladder: An empirical analysis of strategies and efficiencies in social networks. *Proceedings of the Academy of Management*, 2005.

William G. Cochran and Gertrude M. Cox. *Experimental Designs*. John Wiley and Sons, New York, 1957.

Catherine C. Eckel and Philip J. Grossman. Rebate versus matching: Does how we subsidize charitable giving matter? *Journal of Public Economics*, 87(3–4): 681–701, 2003.

— — and — —. Subsidizing charitable contributions: A field test comparing matching and rebate subsidies. Department of Economics Working Paper, Virginia Polytechnic Institute and State University, 2005.

R. A. Fisher. *The design of experiments*. Hafner Publishing Co., New York, 7th edition, 1960 edition, 1920.

Robert S. Gazzale and Jeffrey K. MacKie-Mason. User cost, usage and library purchasing of electronically-accessed journals. In Jeffrey K. MacKie-Mason and Wendy Lougee, editors, *Byting the Bullet: Economics and the Usage of Digital Libraries*, chapter 6. MIT Press, Cambridge, Mass., Forthcoming 2006.

Giving USA. The annual report on philanthropy for the year 2004, 2005.

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

Barbara Hegenbart. The economics of the internet public library. *Library Hi-Tech*, 16(2):69–83, 1998.

Richard D. Katzev. Applying social psychology to charitable donations: Three experiments on nonprofit fundraising. Technical report, The Aspen Institute, 1995.

John A. List and David Lucking-Reiley. The effects of seed money and refunds on charitable giving: Experimental evidence from a university capital campaign. *Journal of Political Economy*, 110(1):215–233, 2002.

— — and Daniel Rondeau. The impact of challenge gifts on charitable giving: an experimental investigation. *Economics Letters*, 79(2):153–159, 2003.

David Lucking-Reiley. Using field experiments to test equivalence between auction formats: Magic on the internet. *American Economic Review*, 89(5):1063–1080, 1999.

Richard T. McCrea. Evaluation of two library-based and one expert reference service on the web. *Library Review*, 53(1):11–16, 2004.

Jan Potters, Martin Sefton, and Lise Vesterlund. After you-endogenous sequencing in voluntary contribution games. *Journal of Public Economics*, 89(8):1399–1419, August 2005.

Richard E. Romano and Huseyin Yildirim. Why charities announce donations: A positive perspective. *Journal of Public Economics*, 81(3):423–447, 2001.

Daniel Rondeau and John A. List. Exploring the demand side of charitable fundraising: Evidence from field experiments. Department of Economics Working Paper, University of Maryland, 2005.

Paul A. Samuelson. The pure theory of public expenditure. *Review of Economics and Statistics*, 36(4):387–389, 1954.

Jen Shang and Rachel Croson. Field experiments in charitable contributions: The impact of social influence on the voluntary provision of public goods. Wharton School Working Paper, University of Pennsylvania, 2005.

Lise Vesterlund. The informational value of sequential fundraising. *Journal of Public Economics*, 87(3-4):627–657, 2003.

Todd R. Weiss. Dean's online campaign dubbed noteworthy. *Computerworld*, July 2003.

Ann Zeller and Anna Chernis. The DMA 2004 response rate report. Technical report, Direct Marketing Association, New York, NY, 2004.