The Virtual Laboratory Infrastructure for Online Economic Experiments

Ben Greiner* H.–Arno Jacobsen† Carsten Schmidt‡

September 24, 2002

Abstract

The goal of this paper is to provide an overview on the *Virtual Laboratory* infrastructure for online economic experiments. We summarize our experience gained from performing several economic experiments on the Internet. The experiments we have run range from electronic markets to individual decision making. From there we synthesize and evaluate a set of methodological issues in performing economic experiments on the Internet. As a result for further exploration we sketch the design of an infrastructure that allows the automated execution of Internet experiments including marketing of experiments, control of application and participation, payment system integration, and evaluation of results. The infrastructure also aims at providing a generic interface for third parties to register and run experiments. We have prototyped the components of this infrastructure and ultimately aim at providing an Internet experiment service for the experimental economic research community.

Keywords: Internet experiments, Internet services, experimental economics, methodology.
JEL-Classification: C90

*Max Planck Institute for Research into Economic Systems, Strategic Interaction Group, Jena, Germany, Email: greiner@mpiew-jena.mpg.de

†Department of Electrical and Computer Engineering, University of Toronto, Email: jacobsen@eecg.toronto.edu

‡Corresponding author: Max Planck Institute for Research into Economic Systems, Strategic Interaction Group, Kahlaische Str. 10, D-07745 Jena, Germany, Tel: +49(0)3641 686 637, Fax: +49(0)3641 686 623, Email: cschmidt@mpiew-jena.mpg.de

1 Introduction

In experimental economics researchers start to leave the controlled environment of their laboratory and conduct experiments in the field. Many reasons for this trend can be identified: To run cross cultural studies one has to address subjects in different countries or to get a more diverse subject pool, compared to the usual students in the laboratory, researcher contact newspapers to run experiments (Bosch-Domenech, Nagel, Garcia-Montalvo, & Satorra, forthcoming; Güth, Schmidt, & Sutter, forthcoming). A main motivation behind this kind of research is to test for differences between standard results established in the laboratory and human behavior in the field.

The Internet provides a natural testbed to conduct experiments. Early psychological online studies proved that it was not only possible to conduct research online, but that it was also feasible to collect large samples of quality data in a short period of time (Birnbaum, 2000). Inspired by these early advances in psychological online research experimental economists started to run experiments via the Internet (Budimir, & Rieck, 1998; Lucking-Reiley, 1999; Anderhub, Müller, & Schmidt, 2001; Shavit, Sonsino, & Benzion, 2001).

Internet experiments have various advantages over computer–based laboratory experiments:

1. higher participation rates
2. feasibility to conduct experiments with long duration (e.g., days)
3. access to a more diverse subject pool (demographically, culturally)
4. higher ecological validity (artificial laboratory vs. familiar environment)
5. avoid experimenter effects
6. feasibility to conduct experiments without an expensive laboratory setup
7. automation of many experimenter tasks
8. Internet experiments run in the laboratory, but laboratory experiments do not run via the Internet, at least most of the times

The main drawback of the open environment of the Internet is that the experimenter looses some control. One aim of this paper is to provide IT-based solutions to some of the following problems

1. less control of subjects (double participation, group decision, drop-out)

2. less control of environment (use of aids, quality of network connection)
3. immediate payment (currently not feasible)

Early papers investigated whether Internet experiments are an alternative to standard laboratory experiments. A natural way to do this is to run the same experiment via the Internet and in the laboratory to compare the economic decision variables. Whether there are differences between the results of an experiment in the Internet and the laboratory, and what properties of the underlying method are causing them, seems to be an interesting research topic (Reips, 1997). While in psychological research correlational research already showed stimulating results (Krantz, Ballard, & Scher, 1997; Reips, 1997), to the best of our knowledge, Anderhub et al. (Anderhub et al., 2001) was the first study in experimental economics to address the validity of Internet experiments in this form. The experiment focused on individual decision making in multi-period savings decisions. To support the argument that data collected from Internet subjects can be useful and valid, an Internet sample has been compared with a laboratory sample in terms of several attributes. The Internet and laboratory experiment data was not significantly different when average economic behavior was concerned. Yet, in general the Internet setting provided higher variances in most of the compared variables and significant lower participation times.

In the study of Anderhub et al. (Anderhub et al., 2001) an unresolved issues remained. A significantly larger number of participants on the Internet did not spend all their money in the last period, which is a clearcut irrationality. This difference in the collected data could result either from the different subject pool or the different experimental environment. Yet, it was unclear which of these sources, if any, will have an influence.

The results of Internet studies will most likely be challenged because of their environment which appeared to be less controllable than in the laboratory, and less on the grounds of a more diverse subject pool. Therefore, Shavit et al. (Shavit et al., 2001) excluded the influence of a different subject pool by drawing subjects only from a student population. In their study individual decision making in risky situations is evaluated. In particular students were asked to bid buying prices for five simple lotteries. They find bids for lotteries on the web significantly higher compared to the laboratory and standard deviations higher on the Internet as well. They conclude that subjects' risk aversion might be lower on the Internet and find again that the Internet medium might increase noise in experimental data.

Similar in the potential to attract a large body of participants are newspaper experiments, that have mostly been conducted in the context of the beauty contest game (Bosch-Domenech et al., forthcoming). Newspaper experiments relate to online experiments like pen and paper experiments relate to their computerized counterparts in the laboratory: The computerized experiment allows to verify the subject's inputs, whereas the pen and paper version allows for input mistakes. Our experiences with a newspaper experiment on ultimatum bargaining (Güth et al., forthcoming), that offered the participants the choice between mailing a newspaper form or to fill out a computerized form on the Internet, revealed a rate of 18.4% incomplete filled out pen and paper forms, whereas all Internet forms where complete.[1]

Besides individual decision experiments technically more demanding strategic interaction studies have been conducted via the Internet. Early approaches used email as a means of message transportation, this includes a bargaining experiment (Baier, Bolle, Buschbaum, & Swiniarska, 1997) and an auction experiment (Lucking-Reiley, 1999). More recently we have run induced value auction experiments on our site where subjects participated in several auctions via the Internet during one week time. The main coordination problem of interactive experiments seems to be the simultaneous start of two (or more) subjects. The asynchronous nature of the Internet requires to apply organizational solutions, like meeting times, to address this. A different approach are one-shot interaction experiments where the individual result of the game is hold back until the latest decision maker in the sequence has made her decision. This approach was taken by a cascade experiment hosted on our infrastructure (Drehmann, Oechssler, & Roider, 2002).

Internet experiments can be conducted in a more or less controlled environment. Some experiments are conducted close to the standards of laboratory experiments (Anderhub et al., 2001; Shavit et al., 2001) with values induced to participants and computerized user interface, some are conducted less controlled with home grown preferences and email communication (Lucking-Reiley, 1999). It seems that many variations between both extremes are possible. The goal of this paper is to provide a methodological and technical guide for experimental economists who want to conduct an online experiment. This paper is not about on how to implement an online experiment, other authors have dealt with this topic (Kirchkamp, 2000), rather on how to design an economic Internet experiment in the light of maximizing experimental control.

Other economic experiment infrastructures on the Internet exist. The Veconlab at the University of Virginia provides a service to run economic experiments for interactive learning. The site offers about 30 different experiments that can be parameterized by the teacher in order to run an economic experiment for teaching purposes (Holt, 2002).[2] Further, the Iowa electronic markets[3] (Forsythe, Nelson, Neumann, & Wright, 1992; Forsythe, Rietz, & Ross, 1999) and the AuctionBot[4] (Wurman, Wellmann, & Walsh, 1998), both of which provide an infrastructure for a very specialized set of economic experiments, i.e., trading based on political and economic events and auctions, respectively. An early attempt to allow to provide a more general experiment infrastructure is the Vlab at the

[1]For the Internet version a drop-out rate of 8% could be observed, whereas the drop-out rate via letter/fax is not under the experimenters control.

[2]`http://www.people.virginia.edu/~cah2k/programs.html`

[3]`http://www.biz.uiowa.edu/iem`

[4]`http://tac.eecs.umich.edu/auction`

University of California, Berkeley, that seems to be currently unmaintained.[5] There are numerous psychological online experiment sites on the Internet. As a good starting point the interested reader might visit the (virtual) experimental psychology laboratory.[6]

The contributions of the paper are as follows. We analyze the pros and cons of different experiment media with a focus on Internet experiments. Therefore, we discuss related work with particular focus on methodological findings of psychological online research. We identify problem areas for running Internet experiments, with experimental control being the key issue. Finally, we develop IT-based solutions to increase experimental control for online experiments and discuss our infrastructure to conduct controlled online economic experiments.

Section 2 discusses methodological issues of Internet experiments to identify the problem areas to gain experimental control. In Section 3 IT-based solutions that provide maximum control for the experimenter are presented. Finally, Section 4 sketches the design rational of the *Internet experiment infrastructure*, we are developing, and Section 5 draws conclusions.

2 Methodological issues

This section discusses methodological issues of Internet experiments. Therefore, we review methodological findings of online research in experimental psychology in the light of experimental economics. We provide organizational solutions for some identified problems in this section and outline the issues that have to be solved by the IT infrastructure for a discussion in the next section.

2.1 Experimental control and payments

The control of Internet experiments can be distinguished in three types of questions (Reips, 1997): (1) preventing subjects from cheating, (2) controlling variables in the sense of experimenting in a controllable laboratory, and (3) avoiding confusion. Whereas the latter question is not different from laboratory experiments, the first two questions are of main concern, both to participants and researchers. Table 1 gives an overview on the items a controlled economic experiment via the Internet can satisfy with respect to the control of subjects and the environment.

[5] `http://elsa.berkeley.edu/vlab`

[6] `http://www.psychologie.unizh.ch/genpsy/Ulf/Lab/WebExpPsyLab.html`

Requirements	Internet computerized	Lab computerized	Newspaper pen and paper	Lab pen and paper
Prevent subjects from cheating				
A subject should not play twice	yes	yes	no	yes
Decisions are not made by a group but by individuals	no	yes	no	yes
A subject should not contact former subjects who did the experiment before	partly	partly	partly	partly
Controlling variables in the sense of experimenting in a controllable environment				
Only controllable help devices should be available	partly	yes	no	yes
Control of subject interaction with GUI of experiment	partly	yes	–	–
Control drop-out of participants	yes	yes	no	yes
Participants should take the recommended time to solve the problem	yes	yes	no	yes
Control quality of the network connection	partly	yes	–	–
Payment of the subjects right after the experiment	no	yes	no	yes

Table 1: Satisfiable requirements for controlled experiments

A major concern related to the Internet technology is to ensure that subjects do not play twice. Internet protocols and Internet services do not provide the feature of unique identification of subjects, so far. We propose the following. For a controlled economic environment the use of a reward medium is a self evident percept (Friedman, & Sunder, 1994). Nearly all electronic payment systems have the built-in feature of identification of payer and payee, a feature that Internet protocols do not provide. By using the identification mechanism of the payment system to identify subjects, the problem of double participation can be solved in an elegant manner. So far, experiments have been conducted without check for double participation on the one hand (Budimir, & Rieck, 1998) or with manual selection of the experimenter on the basis of information of the underlying Internet protocol on the other hand (Anderhub et al., 2001). Using the payment mechanism to identify participants implies to collect payment information from every single participant. Here, the experimenter has to make a trade off between maximizing control and maximizing the number of observations.

An organizational approach to the issue of double participation is to build up a participant database and to invite subjects drawn from this database identified by a code in the invitation email. Although this method cannot fully exclude double participation a carefully maintained database with subjects who participate repeatedly in experiments will include mostly unique references to potential participants.

More difficult to control is that decisions are made by each subject individually and not by groups. In contrast to laboratory environments it is impossible to control if there is more than one person involved in the decision process. Similar to the laboratory environment the problem persists that subjects contact former subjects who did the experiment before.

Here, a short period of time the experiment is available on the Internet provides some help to prevent subject communication. In addition, the recruitment of a heterogeneous subject pool might be advantageous.

Controlling the environment of the subject when conducting an online experiment is our second major concern. The experimenter cannot control whether aids were used to solve the task. Here, the information technology provides the option to offer additional help devices to the subjects, for example a calculator, that can be controlled by the experimenter. A problem related to the WWW is the control of the subject interaction with the graphical user interface (GUI) of the experiment. The usage of the -BACK- and -FORWARD- buttons of current browsers is out of the experimenters control. Here, IT-based solutions have to be provided to prevent mis-usage of the GUI.

The motivation of subjects seems to be of considerable importance, because subjects might terminate participation at any time of the experiment. Situations, where subjects think they have to explain the interruption to the experimenter are unlikely to happen. The probability of drop-out in the Internet experiment seems higher than in the laboratory experiment (Reips, 1997). Therefore, the experimenter should have control over the rate of general drop-out. Especially a selective drop-out should be traced, where subjects leave the experiment with different frequencies depending on experimental conditions (Reips, 1997). Similar to this, subjects should take the recommended time to solve the problem. The experimenter can collect decision times by splitting decisions between several pages and collecting start and end time of the experiment.

Computerized experiments are dependent on its tools, computer and network. The actual quality of the Internet connection presents an important parameter of this tools, that is out of the experimenters control. A broken pencil in a paper-and-pencil experiment is an easy-to-solve problem, a broken Internet connection will result in a passive drop-out and an abrupt end of the subject's participation. Therefore, the experimenter should observe the quality of the Internet connection for the individual.

When receiving payments right after the experiment, subjects will be more motivated and the trust in the experimenter will increase.[7] Right now payments are mostly conducted outside the Internet, for example by transfers to a bank account (Budimir, & Rieck, 1998; Anderhub et al., 2001). This method of payment does not permit a fast transfer of the money and may thus be negative for the reputation of the experimenter.

One additional possibility for the payment procedure is the cooperation with a big on-line shop and the use of coupons. But providing people with anonymous coupons would vitiate the use of the identification mechanism by payment mentioned above, and the requirement of a unique account at the on-line shop would, in addition to that most

[7]Friedman argues, that to increase motivation the experimenter should pay subjects predominantly in cash and right after the experiment (Friedman, & Sunder, 1994). Some electronic payment systems come close to provide most of the features cash owns (MacKie-Mason, & White, 1997).

on-line shops don't care about double accounts, strengthen subject's self selection, which is the topic of the next but one subsection.

2.2 Representation and subject selection

The population of Internet users is still growing rapidly. Several surveys show that all parameters of the population converge to the regular population (Kehoe, & Pitkow, 1995). In the near future the group of Internet users will be as representative for the regular population as the group of TV-consumers. When a growing number of households get connected to the medium Internet, it will provide easy access to a representative population. As a consequence, reaching this population will be as convenient as the recruitment of students on campus for laboratory experiments, today.

Yet, with respect to some demographic variables, the participants of Internet experiments differ significantly from the participants of experiments that were conducted in the laboratory, and from the general population. Smith and Leigh report that their samples from Internet and class were similar in five[8] out of seven demographic variables but different in age and sex composition (Smith, & Leigh, 1997). Psychologist are cautious to generalize about the general population from the drawn samples, yet, Pasveer and Ellard found the student sample of their study to be comparable with the more heterogeneous Internet sample with respect to basic properties (Pasveer, & Ellard, 1998).

The new method of Internet experiments allows for research in areas where established methods failed (Reips, 1997). Studies of specific target populations are one example. So far, it was practically impossible to find subjects with properties that were of interest to the researcher (Reips, 1997). Using the Internet as a tool, the researcher has the ability to recruit specific target populations and to get access to subject pools with limited time budgets like managers, especially cross-cultural studies seem feasible (Hewson, Laurent, & Vogel, 1996).

Another aspect seems to be the "self-selection" of participants in Internet experiments. Reips mentions the example of surveys published on the Web, that are concerned with explosive topics (Reips, 1997). A "self-selection" of participants politically interested will take place (Reips, 1997). One solution to avoid this kind of self selection is by labeling experiments at an experimental site with nicknames or numbers/characters. This way the problem of self-selection is reduced to the participants of an experimental site.

A further topic is the pre-selection of participants. For example, Budimir and Rieck report of a portfolio selection experiment, where the data sets were selected on conditions of the responses to a pre-experimental questionnaire (Budimir, & Rieck, 1998). In this case, everybody could participate during the experiment, but the data sets of subjects

[8]Sexual orientation, marital status, ethnicity, education, and religiosity.

that failed the pre-experimental test where not incorporated in the evaluation (Budimir, & Rieck, 1998). An issue that was discussed when starting our experiments is the preselection of subjects by the choice of technical standards. To use state of the art technology available only to parts of the Internet community, a new Browser plug-in might serve as an example, is biasing and reducing the population of the potential participants. This can be useful when targeting a group of early adaptors, but harmful when conclusions should be drawn in a more general context. The experimenter should be aware of the problem of subject selection and report possible instances for the individual experiment.

2.3 Experimenter effects

A methodological danger of experiments is that subjects try to follow the presumed hypothesis (Reips, 1997). Subjects can be influenced by the demanded characteristics of the experimental situation (Hewson et al., 1996). For example, subjects may feel a need to please the experimenter or a pressure to conform to social norms. The anonymity of Internet seems to be suited to reduce such kind of effects.

In addition, the Internet can help to reduce effects of the attributes and behavior of experimenters, so-called experimenter effects, these having been acknowledged as factors in producing experimental demand characteristics (Hewson et al., 1996). Research carried out over the Internet is advantageous in this sense, because it is a more impersonal form of communication than face-to-face interaction, particularly when anonymity is preserved and contact is minimized (Hewson et al., 1996).

2.4 Ecological validity

The ecological validity[9] of laboratory experiments is “naturally” low (Reips, 1997). One reason might be that the subjects have to get into the new, unknown situation of the laboratory. Web experiments have an advantage in creating a situation where the experiment is brought to the subject, and the subject does not have to come to the experiment (Reips, 1997). A subject does not have to leave his familiar environment, when conducting the experiment on his computer at home or at work, for example. The only prerequisite for the subject’s participation is a computer connected to the Internet. In addition, Internet experiments do not have to rely on the opening times of laboratory experiments (Reips, 1997) or restrictions on the availability of computers during a session. The experiment is available 24 hours a day and 7 days a week.

[9]The term ecological validity refers to the extent to which behavior indicative of cognitive functioning sampled in one environment can be taken as characteristic of an individual’s cognitive processes in a range of other environments. Consequently, it is a central concern of cognitive scientists who seek to generalize their findings to questions about ”how the mind works” on the basis of behavior exhibited in specially designed experimental or diagnostic settings.

2.5 Ethics

Internet experiments have to rely on the same ethical standards as their counterparts in the laboratory. Usually, psychological experiments conducted in the United States adapt the American Psychological Association guidelines (American Psychological Association, 1992) to ensure voluntary informed consent[10] and confidentiality of the participants. Yet, special care is needed in studying on-line behavior, since the technology affords precise recording of many factors that would be inaccessible otherwise (Schiano, 1997). Schiano, in a study of an on-line community, allows for the explicit right to re-examine and delete any recorded content (Schiano, 1997).

In the case of the Internet, the recruitment of subjects is performed differently than in a laboratory environment. Mostly, subjects will be attracted by using the Internet itself. In general, participants do not interact with the experimenter. Therefore it seems necessary to devise methods to ensure that subjects receive fair treatment (Smith, & Leigh, 1997). Smith and Leigh develop some ethical guidelines suitable for conducting research on wide-area networks (Smith, & Leigh, 1997). They pay particular attention to five aspects: "subject recruitment, informed consent, the protection of a subject's right to withdraw participation at any time, protecting against subject fraud while maintaining subject anonymity, and data security" (Smith, & Leigh, 1997).

Subject recruitment will be mostly done over the Internet and needs to conform with the norms of the Internet. Posting to newsgroups is a sensitive area. Readers tend to be quite skeptical of requests that bear no relevance to the topic and dislike advertisements of any kind because they are often of dubious quality. The researcher should be convinced that the posting will be of interest to the addressed community. However, there exists a large body of newsgroups and mailing lists that have a broad enough scope to advertise online experiments.

2.6 Summary

Can Internet experiments satisfy requirements of experimental economics? Compared to laboratory experiments, the Internet experiment does not provide the experimenter with the same control. Additional noise is added: one source is the use of public networks, another is the lack of control of the subjects' environment, may that be at home or at work. When conducting online experiments the researcher has to observe some issues very closely: problematic aspects are the processes of subject selection, self–selection, and the drop–out of participants.

[10]In a laboratory environment subjects physically sign an informed consent form to indicate their willingness to participate. For an Internet experiment researchers present for example the form prior to the log on procedure (Smith, & Leigh, 1997).

Bringing the experiment to the subject provides the advantage of creating a situation where the participant faces a more "natural" environment. Typical problems of laboratory experiments, like experimenter effects, are easier to handle in an Internet setting, where the experimenter does not have to be present during the experiment.

The Internet experiment seems to be less influenced by systematic errors, however, more random noise is added. Therefore, results produced by the new medium should be applicable in a more general way than laboratory results. This implies, that economic experiments could provide more parallelism to the field by using the benefits of the Internet. When conducting an Internet experiment the researcher has to observe some issues very closely: problematic aspects are the processes of subject selection, self–selection, and the drop–out of participants.

3 IT-based solutions to problems identified

This section discusses techniques to solve the issues outlined in the methodological discussion in Section 2 (see Table 1 for an overview). In a first step techniques to increase subject control are discussed and in a second step the focus is on the control over the environment.

3.1 Techniques to increase subject control

The main objective with regards to subject control on the Internet is to uniquely identify subjects. Many online systems identify its participants via their e-mail addresses. Since it is not a problem to posses several different e-mail addresses, this is not a feasible approach for Internet experiments with the objective to avoid double participation by the same subject. A technique based on a tax payer identification number or a social security number might not be effective, as there is no way to check whether a given number is actually correct. Moreover, demanding very personal information from subjects may lead to self selection. For similar reasons do experiments conducted with a lottery-based reward ask for payment information only after the subject has been identified to win. In this case double-participation can only be determined at this (too) late stage.

To reduce the potential that one subject performs the same experiment under several different identities, payment information is used in addition to the e-mail address to uniquely identify each subject. While this reduces the number of potential double participants, it does not fully prevent such abuse. So far we use a traditional payment system, transfers to checking accounts, to identify subjects.

Currently we evaluate the integration of electronic payment systems in order to provide payments right after the experiment. For the above mentioned reasons, it is particularly important that the payment-system uniquely identifies the recipient of a transaction for Internet experiments. In addition, the electronic payment mechanisms must operate timely and credit the necessary funds in a predictable manner. Still electronic payment systems have been developed for different target applications in mind. Several systems provide extensive anonymity to the participants of a transaction and have properties similar to cash, e.g. DigiCash. These kind of systems do not provide the authentication needed for online economic experiments. Other systems, most notably credit card transactions via Internet, mainly specialized on customers paying their purchases. With regards to online experiments the option of peer-to-peer payments seem to be important, especially transfers from the experimenter to the subjects are necessary. A good starting point on the evaluation of electronic payment systems can be found in (MacKie-Mason, & White, 1997; Schmidt, & Müller, 1999).

A Web based technique to augment subject identification is the use of “cookies”. The cookie is an identifier that is stored on the users’ client browser. Subjects are able to prevent their browser to accept cookies, they can remove them, or simply use a different web browser/computer. Therefore, a cookie is not a unique identifier of a subject. However, most Internet users are not familiar with these issues, so that this method can at least increase control.

The issue of whether or not a specific subject has fully independently performed an experiment cannot be avoided, since the experimenter has no control over the environment in which the subject is performing. The issue of one subject contacting another former subject to intentionally or accidentally pass on influential information can hardly be avoided for any experiment, unless the experiment is only run exactly once and the online availability of the experiment is very short. The short availability of a future experiment can be made public in advance. Moreover, the chances of subjects actually knowing each other in a very diverse subject pool can be reduced by very selectively choosing among the user base of registered users, e.g., only notify one user per identifiable domain (company, department etc.) and present individuals with individualized URLs.

To cope with the issue of no-shows and subject’s exact appearance, we are currently working on a reputation mechanism. In our system the subjects are rated with regards to their in time participation in former experiments. The experimenter can invite subjects on the basis of an index that distinguishes between positive and negative ratings.

3.2 Techniques to increase control over environment

There are several objectives with regards to control over the subjects’ environment during the experiment. One demand of economic experiments is that only controllable help devices should be available. This includes the use of aids and tools, especially calculators. In case of Internet experiments only the use of provided aids can be controlled. Therefore, it seems important to integrate easy to use tools, like calculators, in the experiment to provide more control over the individual usage of helpers.

A further topic is to avoid and record drop-outs during the experiment. Most important, we think, is to distinguish between active and passive drop out. A passive drop out occurs due to a broken network connection or a crashed client computer. An active drop out occurs due to a user giving up. For either case it is easy to record the fact that for a specific user ID an experiment was not completed. Different options are available to handle this case. An experimentee may be given additional chances to complete the experiment on future logons, or maybe denied participation in the future.

It is more difficult to distinguish between both kinds of drop outs. To avoid active drop-out one needs to identify experimental stages with high drop-out rates and experiment with means to keep users' attention. The development and evaluation of such techniques is deferred to future work.

Participants should take the recommended time to solve the problem. We use several techniques to record time. Most important the time is recorded at the beginning and the end of the experiment. In case economic decision variables are recorded in the database the time will be added as well. The interaction with the user-interface can be monitored by a standard log file of a web-server in case of an HTML based user interface. The only problem is to identify individual sessions when two different user use the same IP-address and browser.[11] Therefore, we use an extended `access_log` file which also includes the session-ID, the login of the user, and the experiment identifier implemented as a server side-include. The logfile is shown in Figure 1.

```
pD9E2695B.dip.t-dialin.net - - [21/Mar/2002:10:25:32 00100] "GET
/254551037374317/Welcome.mhtml HTTP/1.0" 200 5695

pD9E2695B.dip.t-dialin.net - - [21/Mar/2002:10:25:32 00100] "GET
/254551037374317/Welcome.mhtml HTTP/1.0" 200 5695 254551037374317
testuser AC
```

Figure 1: Standard access_log top, extended access_log bottom

This extended logfile allows to compare actual and expected time to work over the whole and specific parts of the experiment of an individual subject identified by a username and a cookie based session-ID. And it provides a means of control of subject interaction with the GUI, i.e., which pages are requested.

A key design question for the implementation of the online experiment is the distribution of functionality to the server, to the client, or to both. It goes without saying that

[11] This might be the case when two participants are connected to the experiment via the same proxy server.

the implementation metaphor used must support this decision, i.e., pure HTML must be supported by server–side processing (e.g., JSP, PHP or MetaHtml), Java-Script can realize client–side processing, and Java can realize a hybrid approach. However, the more experiment functionality is "out sourced" from the server to the client, the higher the risk of potential fraudulent user interactions. The infrastructure provider has little control over the client side processing to prevent illegitimate use. On the other hand, data input integrity checks on the client side may substantially benefit the overall processing, since less browser–server interactions result, due to pre–validated input data (i.e., input errors are signaled to the user right away without a server connect).

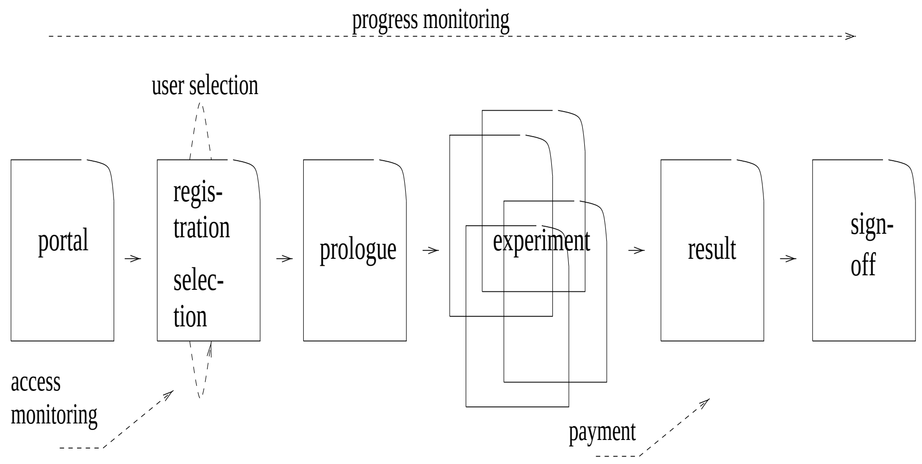

Figure 2: Experimental stages

Our experience in deploying economic experiments via the Internet has shown that some participants do — intentionally or unintentionally — try to "break" the system, e.g., through false input and aimless browser button use. This is a severe problem, since the experiment implementer has very little control over the way a participant uses her browser.

The solution we used for the individual decision experiment (Anderhub et al., 2001) is a client side approach by JAVA-applets and database connectivity classes (pre–JDBC) for the communication with the database. When using the client side approach, keeping state on the client is an easy task. In our case the database connectivity was implemented on the client, therefore some subjects could not participate because of policy restriction (e.g., proxy server and firewalls) of their local site connecting to the Internet. This problem can be eliminated by implementing a middle–tier, that handles the connection to the database, and communicates with the applet by standard http–protocol.

The server side approach to implement experiments uses (plain) HTML over the standard http–protocol; yet, it is more complex to keep track of the users' state. The technique we use in our infrastructure is based on a finite state machine representation of the experimental stages. A stage constitutes a unit of interaction between the experimenter and the participant. This may include experiment instructions, decision forms, and questionnaires. Commonly an experimental stage corresponds to one page delivered to the user's browser (see Figure 2).

In this scheme one central file implements the finite state machine. It assembles the page transferred to the client browser dynamically according to current state and user action. This has two major advantages. One, the server maintains all state information and releases the right information (according to the experiment design) to the user. Two, a single URL is associated with this central control file, i.e., a user cannot jeopardize the operation of the experiment through intentional manipulation of the current URL (e.g., by guessing URLs).

Experimental stages are represented by states, and all possible user actions represent state transitions. A user action, for instance, is the pressing of a form–submit button, but also the forward or backward browsing through the experiment instructions. We implement all actions via form buttons, except for actions performed by the user with her browser (e.g., a page re–load). These latter actions constitute an issue, since at the server–side it may not always be possible to "catch" these actions (e.g., a back–button action might access the client browser page–cache only). We influence this by setting cache invalidate flags, so that a page that should not be re-loadable from the cache will be fetched from the server by the browser upon the occurrence of the re-load action. However, this requires the correct implementation of this protocol by the browser used.

To reduce the subject–browser interaction to the allowed actions, one can use the following approach. First, the experiment will open in a separate window by using JavaScript. Thus, JavaScript has to be enabled in order to participate. The experiment window does only contain the delivered HTML page, all buttons, address windows, title and status bars are disabled. Second, in the experiment window the user is prevented from accessing the context-menu via the right mouse button by using the now surely enabled JavaScript again. At least, this is possible for the most used browsers Internet Explorer and Netscape Navigator. Thus, the only interaction a user can do with the experiment is either to use the provided buttons and links on the experiment page or to close the window and drop out. Although this method is not 100% safe, it increases the subject control. Sample code is provided in Figure 3.

4 Infrastructure architecture and components

In this section we provide an overview of the design of the experiment infrastructure we have built to perform economic experiments on the World Wide Web. Ultimately, we aim at offering an Internet service for use by the research community to perform and to participate in online economic experiments. This service offers a set of functions

```
 <SCRIPT LANGUAGE="JavaScript"> <!--
  function exppopup() {
           fenster=open("experiment.html", "Experiment",
           "width=550,height=500,location=no,toolbar=no,menubar=no,
status=no,directories=no,scrollbars=yes,resizable=no")
        }
 // -->
 </SCRIPT>

 <SCRIPT LANGUAGE="JavaScript"> <!--
 //Disable right click script III -
 //By Renigade (renigade@mediaone.net)
    var message="";
    function rightclickIE() {
            if (document.all) {
                    (message);return false;
                    }
            }
    function rightclickNS(e) {
    if (document.layers||(document.getElementById&&!document.all)) {
                    if (e.which==2||e.which==3) {
                            (message);return false;
                            }
                    }
            }
    if (document.layers) {
            document.captureEvents(Event.MOUSEDOWN);
            document.onmousedown=rightclickNS;
            }
    else {
            document.onmouseup=rightclickNS;
            document.oncontextmenu=rightclickIE;
            }
    document.oncontextmenu=new Function("return false")
 // -->
 </SCRIPT>
```

Figure 3: Top script *Open experiment window*, bottom script *Disable right click*

commonly needed by the experimenter, such as accounting, user authorization and registration. Furthermore, our approach is to provide an environment that automates many tasks that have to be performed by the experimenter, such as participant selection and payment.

The infrastructure aims at providing maximal control over the experiment and its subjects to cope with the issues outlined in the previous sections. Figure 4 depicts the individual components of the environment described in more detail below.

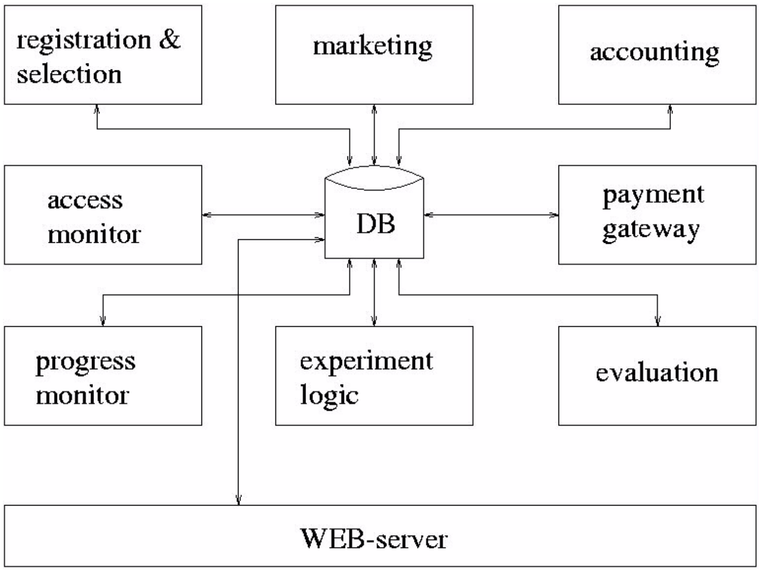

Figure 4: Architecture of the experiment infrastructure.

The **registration** and **selection component** performs user registration and makes user selection decisions. Registration is a simple dialog asking the user to enter name, email, banking account number, and other information. To ensure a certain degree of integrity of the entered data we interact with the user by sending her an access code by email. The email is generic not revealing the experiment site. A user must enter her identification and access code to actually sign–on for the experiment. With this procedure we want to ensure the presence of a valid communication channel with the subject.

User selection is done while registration is in progress. It is based on a set of rules granting or denying access for a registering user. Part of the rules derive from the particular needs of the experiment provider, who wants to address mostly students, or only females, for instance. Other rules are generic and directly address the experiment integrity, i.e., ensure that a novel subject is performing the experiment, for example. Note, that both kind of rules can only approximate the imposed constraints, and true integrity cannot be guaranteed in both cases. We think it is therefore best not to reveal the involved rules at this point.

The **access monitor** records access time, IP–address of client, system and browser type of client, and other information. The collected information is used by the selection sub–

system, for instance, to derive access decision. The information is also used by the marketing component, to measure marketing efficiency and direct further advertising efforts.

The **progress monitor** records similar information during the entire experiment and ensures that parameters returned to the client are appropriately set (e.g., cache–reload attributes). We use a set of techniques to monitor the progression of the subject in the experiment. An analysis of this information may reveal that a subject went multiple times over the experimental instructions while being asked for decision values.[12] This information may be helpful in evaluating the individual results.

The **accounting monitor** component manages the financial assets for the experiment and communicates about payments with the **payment gateway**, the interface to the banking system connected. Our design aims at maximizing security and control over the financial assets available. At each moment in time the component knows exactly how much money is still available. If a critical limit is reached or an unusual high amount is being transferred an administrator is immediately notified and the experiment is halted. Default thresholds are defined and may be configured for each new experiment.

The **payment gateway** is the interface to the banking mechanism used. We aim at supporting several mechanisms: manual banking, online banking, and electronic payment systems, as they become available. The ultimate goal is to provide a fully automated payment system integration.

The key problem we are facing is the incapability to perform peer–to–peer payments via credit cards, the primary means of payment on the Internet to date. Unlike most electronic commerce transactions, the economic experiments we are targeting do require that the Internet service (the experimenter) pays the customer (the subject). Only recently banks seem to offer an online API (in Germany the standard HCBI is emerging) to effectuate customer transactions automatically. Until now, online banks we know offer their services through an HTML–based form, or Java–applet targeted at the *human user*. The lack of a standardized online banking API renders program–driven payment transactions very difficult.

The **experiment logic interface component** constitutes a set of interfaces that permits to plug in experiment implementations. The interface is open, and any implementation compliant to the interfaces may be *plugged in*. Techniques for the management of methods in the Internet environment, which have been developed within the context of the MMM (Jacobsen, Günther, & Riessen, 2000) project, are used to realize this component.

The **evaluation component** serves as direct interface to a statistics packages and to perform result evaluation "on–the–fly". This can be useful for standardized questionnaires

[12]In case the experimental design foresees such actions.

provided by the experimental service, for example the 16PA (Brandstaetter, 1988). For specific experimental data the experimenter must identify how results are aggregated and evaluated, further processing of the results may then be carried out automatically. This evaluation is only a rough estimate and preliminary step, since outliers in the data are difficult to recognize automatically. We are currently working out details of this component, such as online vs. off-line processing and integration in the overall experiment infrastructure.

The **marketing** and recruitment component serves to advertise the experiment before and possibly during the experiment. It draws upon a large database of email lists, individual email addresses, newsgroups, and free Internet–ad space providers. The data inherent to this component is highly domain dependent, and will have to be carefully collected for alternate use. The component automates the sending of email announcements to lists, the posting in news–groups, and the advertising of the experiment on free Internet–ad sites. The component also automates the return traffic processing as much as possible, e.g., management of bounced emails. The experimenter states in a graphical query what kind of subjects she wants to address: students or general public, specific sex, and/or geographical origin. At the end the experimenter states how many subjects should participate. Finally, the query draws the specified number of subjects randomly out of the eligible ones. In the future we aim at further developing the functionality of this component by also incorporating paid–ad providers and means to analyze feedback.

The **sign-off** component is a very simple component that manages mailing and interest lists. It prompts the user and, if she is interested, signs her up for different mailing lists concerning distribution of research reports about the experiment and further experimental economics research.

The implementation of the experiment infrastructure is based on Meta–HTML (Fox, 1998), a server–side include programming environment that enables to establish and maintain session state, to manipulate databases out of HTML–documents, and to author dynamic HTML–pages, among others. The components are built around a database that maintains all experiment and participant data.[13] The Virtual Laboratory is online[14] and currently draws on a mailing list of more than 1000 former participants.

5 Conclusion

Internet experiments have become a popular tool for several research disciplines, such as experimental economics and experimental psychology. We have outlined several methodological constraints that govern Internet experiments as opposed to computer–based lab-

[13] The infrastructure is build with open source software components.

[14] `http://experiment.mpiew-jena.mpg.de`

oratory experiments. One of the major restrictions being the lack of control over the participant. We have motivated the design of an experiment infrastructure that aims at providing an improved degree of control and an automated management of many experiment tasks to the experimenter. The infrastructure we are developing constitutes a generic system with functional entities used in most e–commerce systems. These components comprise access monitoring, progress monitoring, marketing, user authorization and registration, and payment system integration. Furthermore these components may be used for online polling and market surveys, alike. In the future we aim at offering these infrastructure services to the research community to perform online experiments.

Acknowledgments

The financial support from the German Research Society (DFG grant nos. SFB 373/A3 and GRK 316) and the Max Planck Society is gratefully acknowledged. We would like to thank Werner Güth and the participants of the Berlin Internet Economics Workshop 1999 for valuable comments on a previous manuscript. We owe a great debt to coauthors and colleagues who helped us learn and improve Internet experiments, especially Vital Anderhub, Rudolf Müller, and Martin Strobel.

References

American Psychological Association (1992). Ethical principles of psychologists and code of conduct. *American Psychologist*, **47**, 1597–1611.

Anderhub, V., Müller, R., & Schmidt, C. (2001). Design and evaluation of an economic experiment via the Internet. *Journal of Economic Behavior and Organization*, **46**, 227–247.

Baier, A., Bolle, F., Buschbaum, L., & Swiniarska, M. (1997). Please believe me! - Bargaining under incomplete information with communication. Discussion Paper 98. Fakultät für Wirtschaftswissenschaften Europa-Universität Viadrina, Frankfurt(Oder).

Birnbaum, M. H. (2000). Decision making in the lab and on the web. *in* M. H. Birnbaum, ed., 'Psychological Experiments on the Internet'. Academic Press. pp. 3–34.

Bosch-Domenech, A., Nagel, R., Garcia-Montalvo, J., & Satorra, A. (forthcoming). One, two, (three), infinity: Newspaper and lab beauty-contest experiments. *American Economic Review*, forthcoming.

Brandstaetter, H. (1988). Sechzehn Persönlichkeits-Adjektivskalen (16 PA) als Forschungsinstrument anstelle des 16 PF. *Zeitschrift für experimentelle und angewandte Psychologie*, **35(3)**, 370–391.

Budimir, M., & Rieck, C. (1998). A portfolio selection experiment in the Internet. Discussion paper. Institut für Kapitalmarktforschung, Johann Wolfgang Goethe-Universität, Frankfurt am Main.

Drehmann, M., Oechssler, J., & Roider, A. (2002). Cascades in the Internet. mimeo.

Forsythe, R., Nelson, F., Neumann, G., & Wright, J. (1992). Anatomy of an Experimental Political Stock Market. *American Economic Review*, **82**, 1142–1161.

Forsythe, R., Rietz, T., & Ross, T. (1999). Wishes, Expectations and Actions: A Survey on Price Formation in Election Stock Markets. *Journal of Economic Behavior and Organization*, **39**, 83–110.

Fox, B. J. (1998). *The Meta-HTML Reference Manual*. Universal Access Inc. [web page] URL `http://www.metahtml.org`.

Friedman, D., & Sunder, S. (1994). *Experimental Methods - A primer for Economists*. Camebridge University Press.

Güth, W., Schmidt, C., & Sutter, M. (forthcoming). Fairness in the Mail and Opportunism in the Internet - A Newspaper Experiment on Ultimatum Bargaining. *German Economic Review*, forthcoming.

Hewson, C. M., Laurent, D., & Vogel, C. M. (1996). Proper methodologies for psychological and sociological studies conducted via the Internet. *Behavior Research Methods, Instruments, & Computers*, **28(2)**, 186–191.

Holt, C. (2002). Webgames and Strategic Behavior: Recipes for Interactive Learning. mimeo.

Jacobsen, H.-A., Günther, O., & Riessen, G. (2000). Component Leasing on the World Wide Web. *Netnomics*, **2**(2), 191–219.

Kehoe, C. M., & Pitkow, J. E. (1995). Surveying the territory: GVUs fife WWW user surveys. *The World Wide Web Journal*, **1**(3). [web page] URL `http://www.cc.gatech.edu/gvu/user_surveys/papers/w3j.html`.

Kirchkamp, O. (2000). WWW Experiments for Economists, a Technical Introduction. Discussion Paper 00-09. Sonderforschungsbereich 504, Universität Mannheim. [web page] URL `http://www.sfb504.uni-mannheim.de/~oliver/webexp.html`.

Krantz, J. H., Ballard, J., & Scher, J. (1997). Comparing the results of laboratory and World-Wide Web samples on the determinants of female attractiveness. *Behavior Research Methods, Instruments, & Computers*, **29(2)**, 264–269.

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

MacKie-Mason, J., & White, K. (1997). Evaluating and selecting digital payment mechanisms. *in* G. L. Rosston, & D. Waterman, eds, 'Interconnection and the Internet. Selected Papers from the 1996 Telecommunications Policy Research Conference'. Mahwah, NJ: Lawrence Erlbaum [pdf-document] URL `http://www-personal.umich.edu/~jmm/papers/digidoll.pdf`.

Pasveer, K. A., & Ellard, J. H. (1998). The making of a personality inventory: Help from the WWW. *Behavior Research Methods, Instruments, & Computers*, **30(2)**, 309–313.

Reips, U.-D. (1997). Das psychologische Experimentieren im Internet. *in* B. Batinic, ed., 'Internet für Psychologen'. Hogrefe-Verlag. pp. 245–265.

Schiano, D. J. (1997). Convergent methodologies in cyber-psychology: A case study. *Behavior Research Methods, Instruments, & Computers*, **29(2)**, 270–273.

Schmidt, C., & Müller, R. (1999). A framework for micropayment evaluation. *Netnomics*, **1**(2).

Shavit, T., Sonsino, D., & Benzion, U. (2001). A comparative study of lotteries-evaluation in class and on the web. *Journal of Economic Psychology*, **22**, 483–491.

Smith, M. A., & Leigh, B. (1997). Virtual subjects: Using the Internet as an alternative source of subjects and resource environment. *Behavior Research Methods, Instruments, & Computers*, **29(4)**, 496–505.

Wurman, P. R., Wellmann, M. P., & Walsh, W. E. (1998). The Michigan Internet AuctionBot: A configurable auction server for human and software agents. *in* 'Proceedings of the Second International Conference on Autonomous Agents'. Minneapolis, MN, USA.