

# An Advanced Multi-Disciplinary Facility for Measurement and Experimentation in the Social Sciences (MESS)

## Principal investigators:

**Dr. J.W.M. (Marcel) Das** (contact person)  
CentERdata, Tilburg University  
Warandelaan 2  
P.O. Box 90153  
5000 LE Tilburg  
The Netherlands  
phone +31 13 466 8226/8325  
fax +31 13 466 2764  
e-mail das@uvt.nl

**Prof. dr. ir. A. (Arie) Kapteyn**, RAND, Labor and Population Program, and CentER, Tilburg University. E-mail: kapteyn@rand.org

**Prof. dr. A.H.O. (Arthur) van Soest**, Tilburg University, Faculty of Economics and Business Administration, Department of Econometrics and Operations Research, and RAND, Labor and Population Program. E-mail: avas@uvt.nl

## Co-applicants:

**Prof. dr. J.G. (Jelke) Bethlehem**, University of Amsterdam, Faculty of Economics and Econometrics, Department of Quantitative Economics, and Statistics Netherlands

**Prof. dr. ir. B.G.C. (Benedict) Dellaert**, Maastricht University, Faculty of Economics and Business Administration, Department of Marketing

**Prof. dr. J.A.P. (Jacques) Hagenaars**, Tilburg University, Faculty of Social and Behavioral Sciences, Department of Methodology and Statistics

**Prof. dr. M. (Matthijs) Kalmijn**, Tilburg University, Faculty of Social and Behavioral Sciences, Department of Social Cultural Sciences

**Prof. dr. P. (Peter) Kooreman**, University of Groningen, Faculty of Economics, Department of General Economics

**Prof. dr. J.P. (Johan) Mackenbach**, Erasmus University Medical Center, Department of Public Health

**Prof. dr. W.E. (Willem) Saris**, University of Amsterdam, Faculty of Social and Behavioral Sciences, Department of Methods and Techniques for Political Science, and ESADE Business School, Barcelona, Spain

## International Board of Advisors:

**Prof. dr. M.P. (Mick) Couper**, University of Michigan, Department of Sociology

**Prof. dr. D.A. (Don) Dillman**, Washington State University, Departments of Sociology and Community and Rural Sociology

**Prof. dr. D. (Daniel) Kahneman**, Princeton University, Department of Psychology (Nobel Laureate in Economics, 2002)

**Prof. dr. D.L. (Daniel) McFadden**, University of California at Berkeley, Department of Economics (Nobel Laureate in Economics, 2000)

**Prof. dr. N. (Norbert) Schwarz**, University of Michigan, Department of Psychology

**Dr. mr. A.H.E.M. (Nout) Wellink**, President, Dutch Central Bank

## Table of Contents

Executive summary.....	1
A. Outline of proposal and specific aims.....	2
A.1. Build a representative panel of households.....	2
A.2. Design a cost effective recruiting, compensation, and loyalty system for panel members.....	2
A.3. Extend and develop the interviewing software.....	2
A.4. Build a system for incorporating special groups.....	3
A.5. Develop a system for prioritizing access by researchers.....	3
A.6. Incorporate new measuring devices and approaches.....	3
A.7. Integrate with administrative data.....	3
A.8. Integrate the questionnaires of a number of important social and economic surveys into the panel.....	3
B. Background and significance.....	4
B.1. Internet interviewing.....	5
B.2. Sampling and selective response.....	5
B.3. Reliability and mode effects.....	7
B.4. Presentation of tasks, visual displays, preloading and other forms of feedback.....	8
B.5. Experience sampling and the Day Reconstruction Method.....	8
B.6. Retrospective measurement of life histories.....	9
B.7. Biomarkers and other non-interview information.....	11
B.8. Response rates and incentives.....	11
B.9. Advantages and disadvantages of a panel.....	12
B.9.1. Advantages.....	13
B.9.2. Disadvantages.....	14
C. Previous work by the principal investigators.....	14
D. Description of the facility.....	17
D.1. The Internet Panel.....	17
D.2. Focus groups, control groups, and special samples.....	18
D.3. Quality control procedures.....	19
D.4. Links with administrative data.....	19
D.5. Use of non-interview measurements and special experimental set-ups.....	20
D.6. Software development.....	20
D.7. Data dissemination.....	21
D.8. Access procedures and user feedback.....	23
D.9. A parallel facility in the U.S.....	24
D.10. Examples of data collection and experiments.....	24
E. Budget and organization.....	25
E.1. Budget.....	25
E.2. Organizational structure.....	25
F. Comparison with “BIG” criteria.....	27
G. Principal investigators, co-applicants and consultants.....	28
H. References.....	32

## Executive summary

We propose to build an advanced data collection environment for the social sciences, which creates maximal opportunities for innovation, which is fast, cost effective, and easy to use for everyone in the scientific community. This facility will boost research in the social sciences in the Netherlands and abroad, making it possible to gain insight in many aspects of people's lives, such as socio-economic status, preferences, attitudes, social networks, physical and mental health, well-being, and expectations for the future. The core of this "laboratory" is a representative panel of households who have agreed to be available for regular interviews over the Internet. The Internet panel is representative in the sense that recruiting of respondents is based on a probability sample drawn from registers. Internet access is not a prerequisite for participation in the panel. If a respondent does not have Internet access at the time of recruitment into the panel, he or she is provided with a WebTV. In addition to the Internet panel, special groups will be sampled and interviewed in different ways, e.g. immigrants or the oldest old. Particular attention will be given to procedures for maximizing initial unit response and minimization of panel attrition.

The multidisciplinary and longitudinal nature of the survey and the links to administrative data make it possible to substantially improve research into many social phenomena; for instance the interaction between health and socio-economic status, or poverty and social exclusion, or career choices and well-being, among many others.

The facility has four main components:

- It will integrate a number of major permanent data collection efforts, of relevance to different disciplines in the social sciences.
- It will allow for experiments, both in survey methodology and in testing and quantifying of theories in the social sciences.
- It will incorporate and pioneer new forms of data collection and the collection of new types of data, including, but not limited to, PDAs ("personal digital assistants"), self administered measurement devices for the collection of biomarkers, web cams, experience sampling, Day Reconstruction Methods, intensive methods to increase both unit and item response, preloading, and data quality checks.
- It will link survey data to administrative data.

The proposed facility is planned for a period of ten years and entails both an optimal infrastructure for empirical research in the social sciences and the financial resources to carry out this research. The proposal builds on the applicants' extensive experience with measurement and experimentation in the social sciences and aims at a globally unique new infrastructure. Although not part of the present proposal we anticipate that a similar facility will be created in the U.S. (funded by U.S. funding agencies). We expect that more countries will follow. This will add a capability for carrying out comparative research across countries.

The proposal builds on previous efforts in The Netherlands (particularly CentERdata) and in the U.S. (including RAND, Berkeley, and University of Michigan) and on software jointly developed by RAND and CentERdata.

The new facility potentially revolutionizes empirical social sciences: (1) It takes maximal advantage of newly available technology and procedures; (2) Access to the facility is simple and open to every academic researcher; (3) Special emphasis is placed on the quality of measurement; (4) By combining the content of a number of usually separate but overlapping surveys it is highly cost effective; (5) Data become available for analysis very quickly, thus greatly improving the efficiency of scientific research and its societal relevance.

## **A. Outline of proposal and specific aims**

In this section we briefly describe the specific aims of the proposed facility. In Section B, we provide an extensive discussion of background and significance, to position the role of the proposed infrastructure in today's social sciences. Section C briefly discusses the relevant experience of the principal investigators. Section D describes the proposed facility in more detail. Section E summarizes the requested budget and organization. Section F discusses how the proposal relates to the "BIG" criteria. Section G presents brief biographical information for the principal investigators, co-applicants, and advisors. Finally, Section H contains a bibliography.

The goal of this project is to build an infrastructure for data collection that will boost research in the social sciences in The Netherlands and abroad, making it possible to gain insight in many aspects of people's lives, such as socio-economic status, preferences, attitudes, social networks, physical and mental health, well-being, and expectations for the future, among many others. The multidisciplinary and longitudinal nature of the survey data to be collected and the links to administrative data make it possible to substantially improve research into many important scientific issues, such as the causal effects of health on socio-economic status and of socio-economic status on health, of economic poverty on social exclusion, of career choices on well-being, etc. The specific aims below summarize the innovative features of the infrastructure that will be built and maintained.

### **A.1. Build a representative panel of households**

The project will recruit and maintain a panel of households with or without current Internet access. Households without current Internet access will be provided with a WebTV and if needed a TV set and a telephone line<sup>1</sup>. The full panel is expected to comprise about 5,000 households (about 7,500 individuals). All households will be given a fast speed Internet connection if not yet available.

### **A.2. Design a cost effective recruiting, compensation, and loyalty system for panel members**

We will design a compensation schedule for respondents that rewards loyalty. In addition we will develop other loyalty programs and respondent relations programs to motivate respondents and to minimize attrition. Special emphasis will be placed on recruiting methods (including phone, letters, personal visits, etc.), response rates, representativity, multi-mode approaches in cases where Internet interviewing poses difficulties, etc.

### **A.3. Extend and develop the interviewing software**

We will develop additional software features to facilitate easier use of the panel. In particular we will develop a data export module that updates the resulting dataset and a codebook with frequencies in real time and provides interactive documentation for each questionnaire item. This module also will have analytical functionality to allow researchers to produce immediate statistical results as data are being collected. This capability will include graphical and tabular presentation format options. Various types of visual displays and other help modules will be developed to aid in facilitating the presentation of questions to respondents. A preloading module will be built so that responses to earlier surveys can be used in the current questionnaire. In the somewhat longer run we envisage a user-friendly interface that will allow researchers to program interviews directly. As part of this interface we also will develop a library of questions.

---

<sup>1</sup> A Web TV, or Settop Box, or Internet player is a device that can be hooked up to a TV set and a telephone line to allow for Internet browsing, email, etc.

#### **A.4. Build a system for incorporating special groups**

Although the Internet has a number of very desirable characteristics for surveys and social experimentation (see e.g. Section B), there is a need for supplementary modes and sampling of special groups. Supplementary modes will be used to study mode effects where that is suspected to be important and to study special groups that may have particular difficulties with the technology. We also anticipate using focus groups for optimizing questionnaire design. Sometimes the focus groups may be on the Internet and sometimes they may be face-to-face. Another example of a special group is where one wants to conduct experiments over the Internet, but first wants to test a design in a traditional laboratory setting. The flexible multi-mode nature of our software facilitates the integration of the special groups in one data collection and data production system.

#### **A.5. Develop a system for prioritizing access by researchers**

Use of the facility will be free to academic researchers. Hence a system is needed to determine priority of access. We will propose a fast and “light” procedure to regulate access to the facility. Any academic researcher can apply for interview time. A board of overseers (for instance consisting of one of the principal investigators and the co-applicants) will consider the applications and make decisions on accepting the proposals as quickly as possible (usually within two weeks).

#### **A.6. Incorporate new measuring devices and approaches**

Data can be collected in a variety of ways and in various contexts. Next to traditional questionnaire settings, we accommodate the use of visual displays, immediate feedback, preloading of data, the use of PDAs, self administered measurement of biomarkers (saliva, heart rate, blood pressure, blood swaths, etc.), experience sampling, Day Reconstruction Methods (DRM, see Section B.5), etc.

#### **A.7. Integrate with administrative data**

Building on a current experimental agreement of Netspar (Network for Studies on Pensions, Aging and Retirement) with Statistics Netherlands we anticipate being able to link survey information of the respondents to administrative information. This greatly improves opportunities for checking data quality, reducing response burden (no need to collect information that is already available in the administrative data), analysis of selectivity of response, etc. In addition, datasets can be enriched with contextual data or information that is hard to obtain from respondents (e.g. information about exact pension entitlements). Finally, integration with administrative data offers interesting possibilities for more effective reweighing to correct for non-response.

#### **A.8. Integrate the questionnaires of a number of important social and economic surveys into the panel**

We propose to fill part of the available interview time with questions taken from existing or recently terminated surveys. Which surveys these would be will be decided after extensive consultation of researchers. Examples include the Socio-Economic Panel (SEP, conducted by Statistics Netherlands until 2002), a General Social Survey (GSS, with content similar to the European Social Survey), the Netherlands Kinship Panel Study (NKPS), the DNB Household Survey (DHS), the Survey of Health, Ageing and Retirement in Europe (SHARE), the European Values Study (EVS), and the National Election Survey (Nationaal Kiezersonderzoek). These surveys all have their own focus, but they also have substantial overlap. By combining the questionnaires of these panels (and using administrative data so that one can skip most questions about household composition, income, labor market, etc.) we can (1) compare the results obtained with the Internet panel with results obtained in these more traditional surveys; (2) exploit relations between domains that are separate in the

surveys from which we take the questionnaire, but are integrated in the Internet survey. If the comparison with the existing surveys shows sufficient promise, one would expect that in the future new research items could be added to the Internet panel in a highly cost effective way, rather than requiring new (rather expensive) surveys. More generally, this would suggest an entirely new approach to data collection, whereby rather than several separate surveys, one collects a substantially larger amount of information from the same people, thereby achieving economies of scale and a richer environment for analysis.

We expect to create a unique laboratory for the social sciences with unprecedented possibilities for measurement and experimentation. The facility should be available for researchers in The Netherlands and abroad. All data collected are available for the research community, free of charge and without delay. We will coordinate with existing infrastructures for optimal data dissemination, in particular DANS (Data Archiving and Networked Services).

## **B. Background and significance**

Several technological developments are opening up unique new possibilities for empirical research in the social sciences. These developments include the fast diffusion of Internet access, various new communication and measurement devices like cell phones, PDAs, GPS, self administered measurement of biomarkers. These new tools allow for much more accurate and cost effective measurement and experimentation than was possible in the past. To mention a few examples, it is now cost effective to cross validate self reports of food or alcohol intake with blood samples; one can use experience sampling by programming PDAs to give signals at random intervals at which subjects are asked to perform certain tasks or to provide information and opinions; one can replicate classical laboratory experiments within Internet panels; one can set up chat boxes on the Internet and monitor conversations (of course with prior consent by the respondents); one can break up interviews in short modules and let respondents decide when to respond and thus reduce respondent burden; one can randomize questionnaire content across respondents; one can use visual displays, videos, and preloads of data in interviewing; one can apply different modes for different respondents and integrate the results due to powerful software. Last but not least, linking survey data and administrative data is becoming increasingly common.

These are all just examples of how technology can revolutionize empirical research in the social sciences. In this application we propose a facility for research in the social sciences with six distinguishing characteristics:

1. It takes maximal advantage of newly available technology and procedures;
2. It is flexible and forward looking in the sense that it can easily accommodate new technologies and new approaches when these emerge;
3. Access to the facility is simple and open to every academic researcher, both in The Netherlands and abroad;
4. Special emphasis is placed on the quality of measurement, including reduction of non-response; accounting for and reduction of selectivity of samples; cross validation between measures and between survey data and administrative data;
5. It is efficient in the sense that a number of separate but overlapping surveys can be replaced by one;
6. Data become available for analysis much more quickly. This greatly increases the efficiency of scientific research, and fits better with the dynamics of society than the traditional and more conventional methods of data collection.

Below we discuss several aspects of these new technologies. We start with a discussion of the use of Internet interviewing and Internet panels, as this will be the backbone of the structure we are proposing to build. Most of the literature we discuss stems from the U.S., reflecting the dominance of the U.S. in this research field.

## **B.1. Internet interviewing**

Internet interviewing and experimentation create various new possibilities for empirical research in the social sciences. In principle one can measure new or complex concepts (e.g. expectations, probabilities, reservation prices) that are hard to measure with other interview modes and one can design better measurement methods for existing “standard” concepts (e.g. income, wealth). Moreover, all this can be achieved in much shorter time frames than is customary in more traditional survey research. The technology furthermore allows experimentation, follow-up data collection, and feedback from respondents, etc. This alone changes the opportunities for empirical research in the social sciences dramatically, but in addition Internet interviewing creates new possibilities for quality enhancement and quality control. Furthermore, in comparison to other ways of collecting survey data, Internet interviewing is very cost-effective, which in itself expands possibilities for empirical research substantially.

Internet interviewing bears several similarities to other kinds of interviewing (paper and pencil, computer-assisted personal interviewing (CAPI), and computer-assisted telephone interviewing (CATI)). In many ways it can be seen as a combination (or an extension) of these conventional interview modes. However, data collection via the Internet offers a number of advantages over traditional methods. First, it is much less expensive. Second, the Internet offers the possibility of graphical or animated presentation, e.g. display of probabilities through pie charts or exploding scales. Third, for some subpopulations, the response rate to an Internet survey is higher than to a traditional survey: very busy or active people may be more willing to take an interview at the time and place of their choosing. Fourth, the possibility to divide interviews in short sections and spread these over long time periods reduces response burden and actually allows for the collection of much more information from a respondent than would otherwise be possible. Our proposed structure heavily draws on this particular aspect.

In the following subsections, we further discuss several issues pertaining to Internet interviewing, including sampling and selective response, reliability and mode effects, presentation of tasks, the use of visual displays, preloading, and other forms of feedback, and advantages and disadvantages of surveying a panel.

## **B.2. Sampling and selective response**

Any interview mode affects the probabilities of including respondents in a sample. The most obvious example is that Internet interviewing will not work with respondents who do not have access to the Internet. Similarly, CATI is facing increasing difficulties as it becomes harder to reach respondents directly, partly because of the increased use of voice mail and cell phones (e.g. Oldendick and Link, 1994; Link and Oldendick, 1999; Berrrens, Bohara, Jenkins-Smith, Silva and Weimer, 2001) and partly because of the vast increase of unlisted telephone numbers (e.g. Piekarski, Kaplan and Prestegaard, 1999). Other modes have their own drawbacks that are well known.

Today there are many Internet-based samples used to conduct surveys of various kinds. Typically, no attempt is made to make these samples representative of anything other than the population of active Internet users. For example, prospective respondents may be recruited by e-mail or by placing banners on frequently visited websites.

There are obvious problems with such samples (cf. Couper, 2000). Not only are the respondents a selective sample of the population at large, they are the most savvy computer users and therefore may be expected to be much quicker at understanding and answering

survey questions than others. Because they may respond differently, one needs to find a way to validly generalize from such a sample to a broader population of interest.

There are two obvious routes to achieving a sample that is potentially representative of anything other than the population of active Internet users. If for example the target population is all households in The Netherlands, one may draw randomly from this population and provide Internet access to all households in the sample that do not yet have access. By and large that is the approach taken by CentERdata's CentERpanel (<http://www.centerdata.nl/>) and by Knowledge Networks (<http://www.knowledgenetworks.com/>) in the U.S.

CentERdata only provides WebTVs to households without current Internet access. Initially Knowledge Networks provided all participating households with a WebTV, thereby guaranteeing that all respondents experience each task in exactly the same way, while avoiding problems with incompatible software or browser settings. The obvious downside of this approach is that it is costly and potentially cumbersome for respondents who already have their own Internet access, since these respondents have to use two different systems. Perhaps for this reason, respondents with their own Internet access were considerably more likely to drop out of the panel (Dennis and Li, 2003).<sup>2</sup>

Rather than providing all households in a sample with Internet access, an alternative strategy is to employ several interview modes simultaneously: the Internet for those with Internet access and some other mode (e.g. CATI or CAPI) for the others. This approach is being followed in an ongoing project led by one of the Principal Investigators of this proposal (Arie Kapteyn, "Internet Interviewing and the HRS"; also see Section C) jointly between RAND and the University of Michigan. In this project the main focus is to learn about methodological issues of interviewing over the Internet vis-à-vis telephone interviewing. These issues include selectivity, mode effects, preloading, and elicitation of time preference parameters and subjective probabilities. Some of these issues will be discussed below. In comparison to Internet interviewing, the use of mixed modes has three major drawbacks: (1) turn around times between questionnaire design and production of the final data are considerably longer than for Internet alone; (2) it is much more costly and organizationally more involved, partly because one has to train interviewers for each interview; (3) to the extent there are mode effects, making results obtained via different modes comparable may be problematic if not impossible.

Even when providing all respondents in a randomly drawn sample with Internet access, there remain three potential issues of selectivity: (1) willingness to participate in a survey may differ systematically across individuals with different personal characteristics; (2) conditional on Internet access there may still be systematic differences in response rates; (3) the use of the Internet itself may influence response rates in the sense that respondents may like or dislike the Internet as a medium for doing surveys.

Regarding the first aspect, we pay separate attention to methods for improving response rates in Section B.8. Regarding the second aspect, there are well-known differences in Internet use between different groups in the population, e.g. with respect to age. Concerning response rates conditional on Internet access, it is somewhat encouraging that Couper, Kapteyn, Schonlau and Winter (2004) find that conditional on Internet use both the stated willingness to participate in an Internet survey and actual participation are only weakly linked to individual characteristics. These findings do not tell us what the response behavior will be

---

<sup>2</sup> There are of course other conceivable reasons, e.g. the fact that for people without own Internet access, dropping out of the panel also implies cutting off Internet access.

of respondents who would be given a WebTV. As mentioned, Dennis and Li (2003) find that attrition from their panel is largest among respondents who already have Internet access. So this suggests that response rates among non-current Internet users will be at least comparable to those of current Internet users. An analysis of the response rates in the CentERpanel shows that response rates are significantly higher among WebTV respondents.

Regarding the third aspect, a study by Crawford, McCabe, Couper and Boyd (2002) is suggestive. For a population of students at the University of Michigan they conduct a survey via two different modes: Internet and mail. It turns out that in this environment where Internet penetration is essentially 100% the response rate for the Internet mode is 20% higher than for the mail mode.

### **B.3. Reliability and mode effects**

Since most current data collection takes place by modes other than the Internet,<sup>3</sup> it is of importance to address the mode effect of using the Internet. Earlier studies of mode effects between CATI and CAPI interviewing for the HRS have not revealed any significant effect. This, however, should not lead us to conclude the same for Internet interviews. The web is a fundamentally different medium, combining self-administration with a visual medium and interaction with a computer (cf. Groves, 1989; Tourangeau, Rips and Rasinski, 2000).

Recent studies comparing Internet interviewing and telephone interviewing suggest that Internet interviewing generates higher quality information. For instance Chang and Krosnick (2003a) compare surveys done by telephone and over the Internet by two different companies (Knowledge Networks and Harris International). They conclude that Internet interviewing leads to higher predictive and concurrent validity, less random measurement error, less satisficing (i.e. a tendency to give the same answer to each question) and less tendency to give socially desirable answers. Conceivably these differences could be due to selectivity of the Internet samples or the result of experience (the Internet samples are panels, so that respondents have previous experience with interviews). In a follow-up study, Chang and Krosnick (2003b) compared both modes in a laboratory setting where subjects were randomly assigned to either self-administered interviews with a computer or to oral interviewing over an intercom system. They replicate their earlier findings. They also find significant recency effects (i.e. a tendency to pick the first item from a list of possible answers) in the intercom mode and not in the computer mode. Finally, completion time in the computer mode is significantly less than in the intercom mode (17.3 minutes versus 26.6 minutes).

As part of the aforementioned project "Internet Interviewing and the HRS" a number of studies have been completed on mode effects across CATI (the primary HRS interview mode) and Internet. Hurd and Kapteyn (2003) compare elicited subjective probabilities in the 2002 wave of the HRS with the same quantities measured among HRS respondents who participated in an Internet interview in 2003. Since this allows the authors to compare responses by the same respondents using two different modes, selectivity issues can be largely ignored. Essentially the authors find that at least on the basis of descriptive statistics there is no appreciable difference in results between the two modes. Hill, Perry and Willis (2004) estimate fairly complicated models of subjective survival chances and find that the subjective probabilities elicited over the Internet are less noisy than when elicited over the phone. Kapteyn and Van Soest (2004) compare the measurement of ownership and value of a number of asset categories across four surveys: 2002 HRS, 2003 HRS Internet, RAND Internet panel (a panel of respondents over 40 with Internet access; described in more detail

---

<sup>3</sup> We ignore the many convenience samples drawn on the Internet, because of their unknown sampling properties.

in Section D), and the RAND CATI panel (a panel of respondents over 40, interviewed by phone). The asset categories are home, checking and savings, and stocks. They find that after controlling for demographics, the Internet surveys tend to yield both higher ownership rates and higher amounts given ownership, in particular for stocks. A comparison of response rates between the 2002 HRS wave and the same respondents in the 2003 Internet interview shows substantially higher response in the Internet interview.

#### **B.4. Presentation of tasks, visual displays, preloading and other forms of feedback**

Given our approach of building our facility around the Internet, the findings summarized above are generally encouraging. Internet interviewing (or more generally Computer Assisted Self-Administered Interviewing (CASI)) provides in addition several possibilities to improve the presentation of tasks to respondents. For instance, Dornitz and Manski (1996) use training screens, help screens, error checks and review-and-revise screens. Such tools can be used both when eliciting beliefs and preferences, but also when giving feedback to individuals about their decisions.

#### **B.5. Experience sampling and the Day Reconstruction Method**

An overview of the method of experience sampling is provided at <http://psychiatry.uchc.edu/people/postdocs/conner/esm.php>. Scollon, Kim-Prieto and Diener (2003) provide a discussion of strengths and weaknesses. Essentially, experience sampling collects data on individuals in the context of daily life by prompting them to report their feelings, activities, physical measurements (e.g. blood pressure), etc. at prescribed time points. The preferred technology is one where a respondent carries a device like a PDA or perhaps a cell phone that prompts the individual to record the information required at that point.

A somewhat related, but less technology intensive approach is the so-called "Day Reconstruction Method" (DRM; Kahneman, Krueger, Schkade, Schwarz and Stone, 2004). The DRM asks respondents to systematically reconstruct activities and experiences of the preceding day, by using time diaries. "DRM respondents revive memories of the previous day by constructing a diary consisting of a sequence of episodes. Next, they describe each episode by answering questions about its circumstances and about the feelings that they experienced, as in experience sampling" (Kahneman et al., 2004, p. 3).

DRM was applied in a study of 909 employed women who had worked the previous day. Completion times for the self-administered instrument ranged from 45 to 75 minutes. The average number of episodes was 14.1, and average episode duration was 61 minutes. Respondents answered structured questions about each episode: *When* it occurred (start and end times); *What* they were doing (by checking one or more of 16 activities); *Where* they were; *With Whom* they were interacting; and, *How they felt*, using positive affect descriptors such as "warm/friendly, enjoying myself" and negative affect descriptors such as "frustrated/annoyed, depressed/blue, hassled/ pushed around, angry/hostile, worried/anxious, criticized/put down". The affect scales ranged from 0 = "not at all" to 6 = "very much". Finally, respondents answered questions about their job and other personal details (Kahneman et al., 2004).

DRM produces two main outcomes of interest. First of all it produces information about time use, which is of great interest in itself for better understanding the well-being of the population, but also for understanding patterns of activity, the importance of market and non-market work, etc. Secondly, it provides direct information on how people experience these activities. Next to providing direct information on the well-being of individuals it helps to provide insight in such diverse matters as the stress associated with care giving or working

on a job, the utility or disutility associated with activities around the house, the pleasure derived from social contacts, etc.

Well-being measures obtained via DRM may deviate considerably from what is obtained from global measures (belief-based generic judgments). For example, Kahneman et al. (2004) find that interaction with one's children ranks almost at the bottom of activities in terms of affect ratings (just above commuting, working, housework). In other studies where questions about the enjoyment of time spent with children are asked, time spent with one's children is found to be one of the most enjoyable ways to pass the time (e.g. Juster and Stafford, 1985). These discrepancies presumably reflect two influences. First, retrospective questions about time spent with one's children are answered on the basis of child-focused activities at the expense of numerous other situations in which the children are present but not focal. These other situations are not represented in memory as "spending time with the children," which gives rise to highly selective reports based on prototypical activities. Second, episodic reports suffer less from social desirability concerns than global reports (compare, "I don't enjoy spending time with my children" vs. "They were a pain last night").

Clearly experience sampling and DRM can provide unique new information. Internet interviewing can incorporate both experience sampling, by programming questionnaires in a suitable form for PDAs, and the DRM. The latter may be facilitated by an adaptation of calendar life history (CLH) methods, see below.

#### **B.6. Retrospective measurement of life histories**

One of the important difficulties faced by longitudinal surveys concerns how to deal with "initial conditions"—the lives of respondents before the baseline year of a survey. This is especially problematic for surveys that begin monitoring events at later ages. For instance, the Survey of Health, Ageing and Retirement in Europe (SHARE) only interviews respondents 50 and over. Knowing respondents' health or economic status only from the first year of the survey onwards may not be sufficient, since the entire prior histories of health and economic trajectories may matter for the trajectory of events after 50. This problem can sometimes be largely overcome on the economic side if administrative records are available and can be attached to a survey to fill in the missing economic history. It is usually not available for other domains of one's life, like health (either physical or mental health). Therefore, the only realistic option for retrieving histories in certain domains is to obtain the information directly from respondents. The value of information obtained from life histories may be great. For example, one of the most important scientific topics emerging in recent years has concerned the nature of the dynamic relation between economic resources and health (e.g. Adams, Hurd, McFadden, Merrill and Ribiero, 2003; and Smith, 2003). To identify the causal mechanisms, information on the timing of economic and health events over the complete life cycle would be very helpful to identify the causal mechanisms.

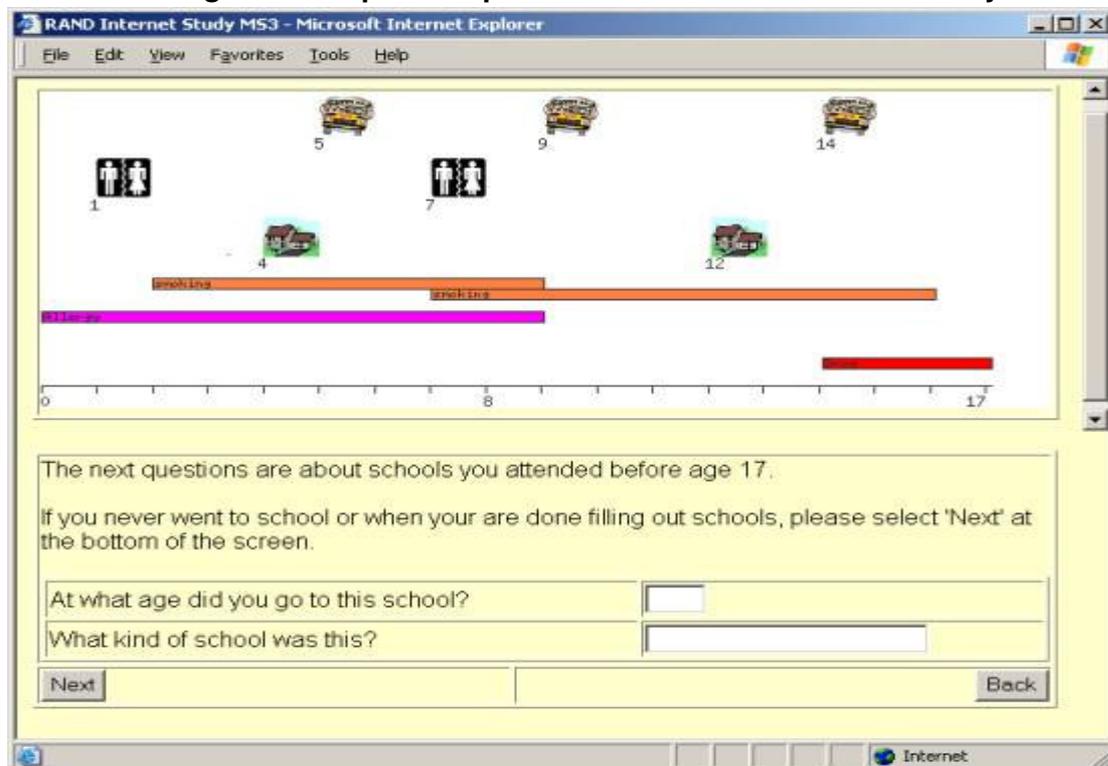
People do not reproduce events from the past perfectly (Rubin, 1996) and the characteristics of the individual, the type of data collected and the period of recall, all influence the accuracy of recall (Rubin and Baddeley, 1989; Sudman and Bradburn, 1973). An increased understanding of the nature of memory has been important for developing data collection techniques that improve accuracy of recall. It is thought that there is a network of themes and time that gives points of association for recollection. Belli (1998) explains that "memories are thematically and temporally structured within an hierarchical ordering". According to this approach there are three types of memories nested within each other: specific events that are nested within one or both of summarized events (class of similar events), and extended events (fundamental organizational structures such as 'Living with X', 'Working at A'). The process of recalling events has been shown to be top-down through these types of memories, or sequential through similar events over time, or the reconstruction of parallel

events. Surveys designed to tap such memory processes have resulted in improvements in recall (Sudman and Bradburn, 1973). For example, acknowledging important reference dates improves the placing of events (aided recall), working systematically through time (backwards or forwards) has the potential to reduce omissions, interviewer probes can prompt better responses, and covering contemporary issues for a topic prior to asking items that have a long recall period can help to locate memories. Wagenaar (1986) shows that spatial cues, such as identifying the people and places involved, in addition to temporal cues, can yield improvements in autobiographical recall.

Following this logic, the project “Internet interviewing and the HRS” is experimenting with various visual aides to prime people’s memory. The approach taken is that of calendar life history (CLH) methods. These CLHs create a time line in which dates are marked about significant salient life events such as dates of marriage, births of children, moves, job changes and the like. These markers can then be used to assist respondents in recalling the occurrence and dates of other events like their health histories. One advantage of using computerized methods in a personal interview (where one can show computer screens to a respondent) is that these calendars can be presented in a simple to understand graphical format which is much more transparent and easier to understand than what is possible to say in a telephone interview.

An example is contained in Figure 1 (taken from an interview that is part of the “Internet Interviewing and the HRS” study). The figure elicits health events in a respondent’s youth. First the respondent is asked about a number of life events, e.g. moving, going to school, separation of parents, whether a parent smoked. The results are then presented in a graphical manner as illustrated in the figure.

**Figure 1: Graphical representation of Calendar Life History**



The events mentioned are represented by respectively the picture of a house, a school bus, a torn picture of a man and woman, and the orange bars. When a respondent moves the mouse over a picture a text appears explaining what the picture represents. Next, questions about health events start. Showing the graphical representation serves two goals: (1) it allows respondents to make corrections and (2) one can next use the time line with the markers to ask questions about events of particular interest. The markers of important life events aids recall and helps to position events on the time line.

### **B.7. Biomarkers and other non-interview information**

A recent development is the collection of biomarkers (e.g. blood or saliva samples, blood pressure, peak flow tests) in social surveys (cf. Finch, Vaupel and Kinsella, 2001), which becomes increasingly feasible due to technological innovation. The collection of such data over the Internet in combination with more traditional survey information holds considerable promise for future research. A very different type of non-interview data comes from registers. If one is able to link survey data to register data several advantages emerge: (1) it saves interview time in cases where one can substitute register information for interview information (e.g. income or wealth information); (2) it can be used to validate survey information (see, e.g. Kapteyn and Ypma (2005), Pischke (1995), Rodgers, Brown and Duncan (1993); Bollinger (1998); Bound, Brown, Duncan and Rodgers (1994)); (3) if one uses registers as a sampling frame one can gauge (and possibly correct for) selective response.

### **B.8. Response rates and incentives**

High response rates are generally seen as a necessary condition for a representative survey, with results that are representative of the population of interest.<sup>4</sup> Response rates to surveys tend to be low in many European countries. The Netherlands is no exception. For instance, unit non-response to the first wave of the Socio-Economic Panel (SEP) was about 50%; the Netherlands Kinship Panel Study (NKPS) had a response rate to the first wave of less than 45% (i.e. a non-response rate of more than 55%); SHARE had a response rate of 54% in The Netherlands. On the other hand the Dutch version of the European Social Survey had a response rate of 68%. In the U.S. response rates tend to be higher; for instance the Health and Retirement Study reports a response rate of 82% for its first wave and response rates for later cohorts on the same order of magnitude; the National Longitudinal Study of Youth (NLSY) reports a response rate of 71% for the 1997 cohort. The way in which response rates are calculated is not always comparable across surveys, but there appears to be little doubt that the U.S. surveys attain considerably higher response rates.

One might argue that some of the differences reflect differences in culture. It seems unlikely however that all differences can be explained by culture; it would be hard to explain the wide variation in response rates in The Netherlands for instance. After all these surveys take place within a given culture. The literature on response determinants offers a number of possible explanations why response rates may differ. The first determinant is the amount of effort that goes into contacting potential respondents. This includes everything from interviewer training, the way respondents are addressed, the number of contact attempts, refusal conversion, etc. A second important difference may lie in the incentives offered to respondents; mostly in Europe incentives are modest or non-existent. The major U.S. social science surveys all reward respondents for their participation. For instance the HRS currently pays \$40 for a completed interview; while NLSY pays \$75. In her recent overview

---

<sup>4</sup> It should be mentioned though that attempts to increase response rates do not always seem to help much in decreasing selectivity; see e.g. Beukenhorst (1999, 2001; cited by Te Riele (2002)).

of biases introduced by non-response in surveys of Statistics Netherlands, Te Riele (2002) pays no attention to the role of financial incentives.

The literature on financial incentives appears to have reached broad consensus on the importance of financial incentives. A number of points are noteworthy: (1) prepaid incentives yield higher response rates than promised rewards after interview completion; (2) money is better than gifts; (3) response rates increase with increasing amounts of money; (4) the effect of incentives are approximately linear, i.e. the next euro produces about as much increase in response as the current one; (5) incentives give the biggest bang for the buck in surveys where response rates would be low without incentives. See for instance Singer and Kulka (2002). One may worry that the quality of the extra response bought with the incentives may be less than for other respondents. The evidence suggests that actually incentives generate a positive effect on quality (see e.g. Mack, Huggins, Keathley and Sundukchi, 1998).

Different respondents may react differently to incentives and indeed many respondents may not need any incentive. Jasso, Rosenzweig and Smith (1999) describe an experiment with the New Immigrant Pilot Survey in the U.S. in which initial refusals were offered incentives up to \$100 for completion of the interview. By varying the incentives randomly they are able to derive a table with different thresholds (reservation prices) for different respondents. The reservation price seems to rise sharply with income: high-income individuals need a higher incentive to be willing to participate. This is confirmed in other studies. For instance Mack et al. (1998) found that a \$20 incentive in the first wave of the Survey of Income and Program Participation (SIPP) panel was much more effective in recruiting and retaining black and low-income households than others.

A final issue is what the effects are of incentives in panels. One case of interest is formed by an experiment conducted with regard to the first wave of the HRS. In the first HRS wave (1992) the investigators were dissatisfied with the response rate they attained. They decided to approach the refusals again and instead of the standard \$10 incentive, they offered the respondents incentives up to \$100; thereby raising the response rate by 4 percentage points. Perhaps more importantly, in subsequent waves the initial refusals responded at almost the same rate as the other respondents, despite the fact that their incentive was now equal to what was paid to the others. See Lengacher, Sullivan, Couper and Groves (1995).

Te Riele (2002) provides a comprehensive overview of response rates and selectivity in a number of surveys conducted by Statistics Netherlands. The topic of a survey is one of the factors that may affect who is responding. The ensuing suggestion is that a general purpose survey may lead to less selectivity than a targeted survey. In most surveys, respondents with a lower socio-economic status (lower educational attainment, lower income, renters rather than home owners) tend to be underrepresented. Partly one may ascribe that to lower cognitive skills. The evidence discussed above suggests that financial incentives may be particularly effective among this group. Non-native respondents also tend to be underrepresented; however reweighing by country of origin seems to be effective in correcting bias. In other words, immigrants may be less likely to be interviewed, but those who are interviewed are representative of their compatriots.

### **B.9. Advantages and disadvantages of a panel**

Since an Internet panel is an important component of the proposed facility it is worthwhile to briefly discuss the pros and cons of using a panel.

### **B.9.1. Advantages**

*Quality control and preloading:* Panel data present several possibilities for quality improvement and quality checks. The simplest case is where the same question is asked in subsequent interviews and we know that the underlying fact cannot have changed or can only change in one direction (e.g. birth year cannot change and level of education cannot fall). It is well known that in practice, panel data often show several of these “impossible” changes. Where changes in one or both directions are possible, there may still be ways to determine their plausibility. In general these procedures fall under the heading of outlier detection (cf. e.g. Hadi, 1992, 1994; Zaman, Rousseeuw and Orhan, 2001).

Although it would be useful to detect likely errors and outliers in panel data, it is even more attractive to be able to correct these errors. The ideal strategy in such cases is to go back to the respondents to resolve apparent inconsistencies. This strategy is followed by the HRS in the so-called “2001 HRS Asset-Change Reconciliation Callback Project” (Hill, 2002). Project staff called 1481 households whose aggregate net worth inexplicably changed by more than \$150,000, or for whom one net worth component changed more than \$50,000 between the 1998 and 2000 waves. They achieved reconciliation for 1,255 households (2,479 net worth components) and, as a result, the variance in measured change for the entire sample of 11,583 households with the same financial respondents in both waves was cut in half. As a result, since 2002 the HRS interview contains an on-the-fly reconciliation module.

For several decades researchers have been aware of the possibility offered by computer assisted interviewing to use information obtained in one wave of a panel to condition interviewing in subsequent waves (e.g. Hill, 1994). Yet, the application in practice appears to have been limited. Partly this can be ascribed to the complexity of the ensuing programming task, and partly it may be due to doubts about the effect on the quality of the resulting data. An important concern is that changes may be underestimated because respondents may be tempted to simply repeat their response of previous waves. This has for instance led the Bureau of the Census to considerably delay dependent interviewing in the CPS (U.S. Department of Commerce, 1975; and Hill, 1994). Hill (1994) considers the 1985 and 1986 SIPP panels. Since each panel member is in the SIPP for 30 months, the observation periods of the 1985 and 1986 panels overlap for about 18 months. In the 1985 panel the interviewing across waves was independent whereas in the 1986 panel the interviewing was dependent. By comparing data on job changes, work hours, wages, and employers across the two panels one is able to assess the plausibility of observed changes in the two panels for the overlapping period. The main conclusion is that most of the observed change in the independent interviewing treatment was due to noise rather than to real change. Similar conclusions are reached when considering event histories.

Clearly, in our planned Internet panel, the use of dependent interviewing for the purpose of improving data quality has to be an important part of the services offered.

*Measurement of change:* Nijman and Verbeek (1990) investigate the cost effectiveness of using a panel, a series of cross sections, or a combination of both if one wants to estimate a linear combination of time effects. The outcome clearly depends on the relative cost of collecting data through a panel or through separate cross sections, as well as on the stochastic process that one is assuming. For a linear panel data model with individual effects they find that for the measurement of change, a panel is unambiguously more cost effective if the cost of collecting the data is the same for a panel and for a series of cross sections. Even if it were more costly to collect data through a panel, it may still be better to use the panel than a cross section, depending on the amount of unobserved individual heterogeneity in the data. This is intuitive, since unobserved heterogeneity contributes to sampling variation across independently drawn samples, whereas in a panel one can control for

unobserved heterogeneity. Assuming that Internet panels can be (substantially) cheaper than repeated cross sections would then suggest that a panel is more cost effective when it comes to the measurement of change; and probably also when it comes to the estimation of other linear combinations of time effects (like a trend).

*Cost:* Setting up any panel entails substantial fixed costs, including respondent recruiting, panel administration, address tracking, loyalty programs to keep panel members motivated, etc. This is no different for an Internet panel. However, in comparison with other interview modes (like CATI and CAPI) the marginal cost of conducting interviews with members of an Internet panel is low, and may consist mainly of incentive payments. Communication by e-mail is essentially free; information exchange with panel members can take place almost exclusively by using the Internet.

*Use of background variables:* This provides an additional efficiency gain, as it obviates the need to collect standard background variables at each interview. Furthermore, background variables can be used to stratify the sample and tailor interviews to the characteristics of a respondent.

### **B.9.2. Disadvantages**

*Attrition:* Respondents attrit from panels. To the extent that such attrition is non-ignorable for a parameter of interest, this requires special measures. Typically this requires modeling the mechanism that generates the attrition (or more generally the process that causes observations to be missing). The point here is not to discuss the extensive literature on selectivity, attrition, and missing data. Rather we want to make two points. First, also in a cross section non-response may be non-ignorable and hence create similar problems as attrition in panels. Typically, however, attrition in a panel comes on top of initial non-response (unit non-response to the first wave) and hence may be considered to be more serious. Second, a substantial amount of information already may have been collected on respondents before they drop out of the panel, which makes it easier to reliably model the process of attrition, and hence to correct for the selectivity bias that otherwise would be created by non-ignorable attrition. Additional information can be gleaned from refreshment samples. See e.g. Hirano, Imbens, Ridder and Rubin (2001). Refreshment samples are among the special groups that we intend to use to correct for possibly selective attrition.

*Learning effects:* Respondents in a panel after a while become experienced respondents, and as such their response to questions may differ systematically from the response of individuals who are not experienced respondents. This may be good or bad. To the extent that respondents learn how to interpret questions, they may make fewer errors than novice respondents. For example, CentERdata takes care to display question screens in a consistent way, so that confusion is minimized. On the other hand, experienced respondents may become strategic. For instance, they will learn that after an ownership question (e.g. “do you have a checking account”) usually comes a follow-up question if one answers “yes” (e.g. “what is the balance of that checking account?”). So they can learn that saying “no” reduces the burden of their task. To control for such phenomena, it is, once again, useful to draw refreshment samples and to systematically compare responses by new respondents (e.g. respondents replacing panel members who drop out) to the response by existing panel members.

## **C. Previous work by the principal investigators**

The proposed facility builds on a number of on-going projects in the area of interviewing methodology and data collection. The first one of these is the CentERpanel in The Netherlands directed by co-PI Marcel Das. He is the successor of Arie Kapteyn, who

founded CentERdata while he was a professor at Tilburg University. CentERdata is an institute running a representative Internet panel of some 2,000 households (the CentERpanel). Every weekend, the panel households answer questions posted for them on an Internet site, using their own computer or a WebTV provided to them by CentERdata. Numerous papers have been written by researchers from several countries using data from the CentERpanel. The technology allows economic experiments (e.g. in intertemporal decision making, job search, decisions under uncertainty) and experiments with new ways of presenting tasks. Various methodological experiments have been or are being carried out, including the use of unfolding brackets, graphical presentation of probabilities, methods of measuring consumption expenditures, the analysis of primacy and recency effects, stated preference experiments, the use of vignettes to measure health related work limitations, etc.

A large scale project in which the CentERpanel is used is “A Laboratory for the Study of Household Financial Decisions” of which Marcel Das is the principal investigator. The core of the study is the DNB Household Survey (DHS), consisting of annual interviews with participants in the panel. Every year the questionnaire is administered to the households during six weekends, yielding up to 2½ hours of interview time with the respondents. Areas covered include: demographics and work; housing and mortgages; health and income; assets and liabilities; and a large number of economic psychology topics. The waves 1993-2004 are currently available free of charge to any interested researcher subject only to a signed confidentiality statement. As a result of the lengthy questionnaire a unique amount of information on a wide variety of topics is available, without a counterpart in any existing socio-economic panel survey elsewhere. Next to the annual surveys, additional interview time is devoted to additional experiments to test theories of financial behavior or to gauge and improve data quality. Our proposal builds upon the successful experiences gained in this project.

The second project on which the facility builds is “Internet Interviewing and the HRS” of which Arie Kapteyn is principal investigator and on which Arthur van Soest is an investigator. The primary aim of the project is to inform the HRS about the potential of Internet interviewing; the secondary aim is to investigate issues for which Internet interviewing is particularly suited and which could lead to improvements in the HRS questionnaire. As part of the second aim various experiments are conducted that explore different ways of presenting information (e.g. by using pictographs or pie charts to represent probabilities) and different ways of framing decision problems (e.g. by constraining choices with a “built-in” intertemporal budget constraint). As part of the project three main data collection activities are undertaken:

First, a randomly drawn subsample of HRS respondents with Internet access is interviewed twice, in 2003 and 2005. Second, a panel of 1,000 individuals over the age of 40 with Internet access has been drawn from the U.S. population. The individuals in this panel are interviewed eight times over approximately a four-year period. Third, a telephone sample of individuals has been drawn, representative of the U.S. population over the age of 40. These individuals are interviewed once a year for four years. The telephone sample acts as a control group for the Internet panel and serves to study mode effects.

A third, recently started, project is the RAND Roybal Center for Financial Decision Making directed by Arie Kapteyn and co-directed by Arthur van Soest. This project will provide all CATI respondents in the “Internet Interviewing and the HRS” project with WebTVs. The resulting panel, called the “American Life Panel” emulates the CentERpanel, but is currently restricted to respondents 40 and over. Total sample size is 1,500 households. The Roybal Center aims to better understand how people reach decisions about issues affecting their economic status in old age and inform how public policy can educate or otherwise help

people align decisions with their long-term objectives. The focus is on assessing how people collect information, how they interpret the features of alternatives, how they think about risks and probabilities, what their preferences are, and how well they match their decisions to those preferences and interests. As part of the Roybal Center, respondents will be surveyed about three times per year with in-depth questions about a limited number of issues. We will develop decision support software and extensively use visual aids and graphical designs, thus fully exploiting the advantages offered by the Internet.

A fourth project relevant to the new facility is “Comparison of Survey and Register Data: The Swedish Case” led by Arie Kapteyn. In that project a sample of approximately 1,000 individuals over the age of fifty were drawn from the Swedish registers (a longitudinal administrative dataset called LINDA) and then surveyed about a number of life domains relevant to aging. The survey information was next linked to the information from the registers. Among other things the register information has been used to gauge the reliability of the survey data. As an example, Kapteyn and Ypma (2005) demonstrate that estimates of even simple econometric models can be very sensitive to deviations from the standard, but often untenable, assumption that administrative data are reported and matched without error. Furthermore they show that conclusions about the nature of measurement error in survey data critically depend on what is assumed about the quality of the administrative data. Assuming that the administrative data are error free leads one to conclude that measurement error in survey data is strongly mean reverting; that is, low income respondents exaggerate their income and high income respondents under report. That conclusion largely disappears, if one allows for a small amount of mismatching in the administrative data.

In recent years, Arthur van Soest has been involved extensively in the design of innovative survey questions, extensively exploiting the capabilities of the CentERpanel and the U.S. Internet panels discussed above. An example is a project on international comparison of self-reported health and work disability. It has often been suggested that people in different countries use very different response scales when evaluating (a certain aspect of) their health. What is “good health” to a typical respondent in one country, is often no more than “fair” in the eyes of a respondent in another country. To identify such response scale differences, Van Soest has, in several papers with Smith and Kapteyn (e.g. Kapteyn, Smith and Van Soest, 2004) applied the methodology of vignettes. Descriptions of hypothetical persons with a health problem are presented to the respondents, who are then asked to evaluate the health of these persons. Since the same hypothetical persons are presented in all countries, differences across countries in the evaluations identify differences in the response scales. Vignettes were fielded in the MS Internet survey in the U.S. as well as the CentERpanel. Applying the vignette methodology shows, for example, that about half of the difference in reported prevalence of work disability between the Netherlands and the U.S. is due to response scale differences. In follow up work, explanations for response scale differences are explored, e.g. reference groups and social norms.

Another example is an experiment on the ultimatum game fielded in the CentERpanel, joint work of Van Soest with Charles Bellemare and Sabine Kröger; see Bellemare et al. (2005). The ultimatum game is a popular game in experimental economics, aimed at measuring aversion against inequity. Whereas experimental economists almost exclusively use convenience samples of students, this project fielded the experiment in the CentERpanel, making the results representative of the Dutch adult population. A second innovation in the project is that subjects are asked for their expectations on what the opponent will do. Results suggest, among other things, that students tend to be much less inequity averse than the average Dutch respondent, and that the expectations on which decisions are based deviate substantially from rationality.

A third project in which Van Soest participates is SHARE, the Survey of Health, Ageing and Retirement in Europe. SHARE is a major effort to conduct large-scale surveys of the elderly in twelve (European) countries. The first wave was conducted in 2004. A unique feature of SHARE is its multi-disciplinary nature – with questions on health, socio-economic status, social relations and family transfers, well-being and job satisfaction, etc. See Börsch-Supan et al. (2005). Van Soest is the country team leader for The Netherlands and has worked on data validation and non-response issues. He has also developed an experimental module with vignettes on work disability and several health domains. A second wave of SHARE will be fielded in 2006. Van Soest is the Principal Investigator on a recently awarded project which will field an extension of SHARE wave 2, and introduces vignettes in new domains such as well-being, poverty and social exclusion.

CentERdata has gained wide-ranging expertise in the area of interview software development. Under supervision of Marcel Das, CentERdata is responsible for developing the survey instruments in SHARE. A generic CAPI version as well as all the country specific versions have been programmed by CentERdata. After several rounds of revisions of the generic instrument, the participating countries translated the question texts of their individual questionnaires using the Internet and a so-called Language Management Utility (LMU), developed by CentERdata. The translated question texts, interviewer instructions, answer categories, fill texts and other instrument texts (like error messages) from the (LMU) database were used to generate specific questionnaires for each country, based on the blueprint of the generic version. Next to the CAPI instruments, several tools were developed to facilitate the fieldwork. Although the proposed facility will not directly use software developed for SHARE, the experiences gained in this project are very relevant for developing software for the proposed facility. An initiative directly relevant for the new facility is the development of MMIC™, which is a joint project of RAND and CentERdata. Details can be found in Section D.6.

## **D. Description of the facility**

The best way to think of the proposed facility is to compare it to a laboratory in the sciences, with various components. Within the laboratory one can do a variety of experiments and test or quantify a variety of theories. Some of the experiments are aimed at calibrating or improving the measurement tools. Thus the laboratory provides both an environment for advanced experimentation and measurement, and a testing ground for improving measurement methods.

The components of our proposal comprise the building blocks for an infrastructure and the resources to actually use the infrastructure. In this section we describe the building blocks, while the next section is devoted to the necessary resources and how these will be allocated. This is not the place to describe the experiments that will be done with the facility; after all that is something the users have to decide. However, we will illustrate its potential uses by means of a number of examples. In addition we will pay particular attention to experiments to gauge and improve quality.

### **D.1. The Internet Panel**

We aim at a panel of about 5,000 households (about 7,500 respondents). In principle the respondents are available for interviews every week, but we anticipate a system where on average respondents will be interviewed at a lower frequency (to minimize the respondent burden). An interview (or experiment) is typically restricted to a maximum of 30 minutes per week and often will be shorter. Larger interviews can be broken up in 30 minute segments and administered over a period of several weeks. Respondents will receive an incentive

payment on the order of €15 per 60 minutes interview time, most likely at a staggered rate akin to frequent flier programs of airlines. That is, incentives will be higher for respondents who stay in the panel longer and who participate more frequently in interviews.

We do not intend to build on the current CentERpanel, which currently consists of 2,000 households. Instead, the existing panel will run in parallel with the new facility. Keeping the old panel alive allows for testing of learning effects (do the 'experienced' respondents in the existing panel answer differently than the respondents in the new panel) and it still feeds the long-term existing data series (like the DNB Household Survey). Moreover, having the two panels in parallel offers the possibility for testing of recruitment effects: the recruitment of the panel members for the new facility will differ from the existing one.

In the current set-up of the CentERpanel recruitment of panel members is done by telephone screening. A random sample of candidate panel members is interviewed by telephone. In the first interview a number of questions are asked about the demographic characteristics of the household. The interview is concluded with the question whether the interviewee would like to participate in survey research projects. If so, the household is included in a database of potential panel members.

If a household drops out of the panel, a new household is selected from the database of potential panel members. This is done on the basis of demographic characteristics (such that the panel will remain representative of the Dutch population). The selected household is asked whether the members of the household would like to become panel members, and if so, a number of additional questions are asked.

A major drawback of this method is the need for a landline phone connection. Nowadays many households only have a mobile phone. In the above procedure these households are excluded, as well as those households with no phone connection at all.

For the new panel we propose to draw a probability sample from population registries. Next a multi-mode approach is used to enlist panel members, including letters of invitation, phone calls and personal visits; the number of contact attempts will be at least equal to twelve and will involve both phone calls and personal visits. Initial refusals will be recontacted by experienced interviewers to attempt refusal conversion. Based on the literature discussed in Section B.8, a prepaid incentive will be given to a potential respondent at first contact (e.g. in a contact letter). We will also consider the possibility of varying initial incentives to entice respondents to participate; the fact that respondents get Internet access and high-speed connections (if they do not already have that) should provide a strong additional incentive for participation. Given the fact that lower income groups appear to be more sensitive to financial incentives than other groups we expect this to largely (if not completely) redress the tendency of lower socio-economic status groups to be under represented.

The described approaches aim at maximization of representativeness. The incentive of €15 per 60 minutes of interview time is meant to minimize attrition and selectivity. As mentioned before this amount is an average. Effects of (different values of) incentives will be experimented with during the first years of use of the facility.

## **D.2. Focus groups, control groups, and special samples**

Parallel to the Internet panel a system of recruiting focus groups, control groups and special samples will be developed:

1. Focus groups are primarily used to fulfill the classical role of questionnaire testing and development; they come in two forms: face-to-face in meetings and via the Internet. In

- the latter case members of the focus groups are probed in a semi-structured way to understand how they experience the tasks (interviews, experiments, etc.) given to them;
2. The uses of control groups include the study of mode effects (e.g. a group is being interviewed face-to-face and their answers are compared to the answers of panel members), panel effects (are there systematic differences between panel respondents and one-off participants and if so is that good or bad);
  3. Special samples may be drawn for a number of reasons, including: (1) the study of selectivity and representativity (e.g. interviewing the oldest old who are not Internet savvy to see if their responses are systematically different or whether Internet interviewees provide systematically different responses from otherwise observationally equivalent non-Internet interviewees); (2) groups that are too small to be sufficiently represented in the Internet sample or are difficult to interview (e.g. non-Dutch speaking residents).

### **D.3. Quality control procedures**

The quality control procedures will initially draw on three sources. First of all we will adhere to the procedures put in place by CentERdata in maintaining its CentERpanel. Three main themes make up the core of the data quality program:

1. Non-representativeness: do panel members differ from the Dutch population, and if yes, in what respect? How can this representativity be improved? The drawing of special samples and comparison with administrative data play important roles in this respect.
2. Measurement error: what is the quality of the data collected via the panel? To address this issue both the availability of administrative data and the possibility to check interwave consistency are crucial.
3. Panel management: in which respect is data quality influenced by the way the panel is managed and the way in which the panel members are approached?

With respect to the last mentioned theme: CentERdata's experience shows that a well-organized helpdesk facility and panel management system is of great importance to run a panel without any problems and to maintain respondent motivation. The regular distribution of newsletters, sending get-well cards, etc. are useful additional instruments in maintaining loyalty and motivation.

Secondly we will draw on the experiences in the "Internet Interviewing and the HRS" project. We will also rely heavily on the experts in that project, some of whom are listed as advisors on this proposal (see Section G).

In the third place a permanent program of methodological research will accompany the operation of the facility. Research topics include selectivity, panel effects, mode effects, analysis of key strokes and time stamps, presentation of tasks (e.g. varying visual displays), preloading, response feedback, range checks, social desirability, response maximization, etc. These studies will be published in academic journals, but at the same time they will help to continually improve the quality of the data collected.

### **D.4. Links with administrative data**

Data quality can be improved greatly if links with administrative data are possible. The availability of such data saves interview time; allows for cross-validation of survey information with administrative information; and allows for the analysis of determinants of non-response if one has administrative information on non-respondents. Furthermore, register information can be used to construct sample weights to correct for differential response.

As mentioned before, Netspar has recently entered into an agreement with Statistics Netherlands. Under the terms of the agreement, a host of administrative databases will be made available to researchers by way of remote access. (A researcher writes his/her own code and runs programs remotely at the Statistics Netherlands server; to do this one needs a dedicated secure computer at one's own institution. Currently there are two such computers, located at Tilburg University. If the experiment is successful the number of secure computers will be increased so that this facility will be available nation wide.) Furthermore Statistics Netherlands will also link administrative information to survey data and make the enriched data available via remote access. Examples of administrative information that will be made available are: income, labor market history, pension entitlements, wealth, and housing. An large amount of information is contained in Statistics Netherlands' Statistical Database (SSD). The SSD is a set of micro-linked and micro-integrated data files including demographics and socio-economic data.

#### **D.5. Use of non-interview measurements and special experimental set-ups**

A number of non-interview measures have been discussed in Section B, including DRM, experience sampling, and biomarkers. As indicated in Section C, various experiments have been carried out with the CentERpanel; we anticipate to extend this capability in the new facility and to combine experiments in the panel with experiments in a more classical laboratory setting. See also Section D.10 below for some examples.

#### **D.6. Software development**

The main "engine" of the new facility will be MMIC™. RAND and CentERdata are jointly developing a multi-mode interviewing capability that integrates various traditional modes of collecting interview data, including telephone interviewing, written interviewing, and personal interviewing. This may seem a fairly obvious thing to do, but there is very little software (and certainly no user-friendly and powerful software) that is able to manage the whole interview process from questionnaire design, sample management, and fieldwork monitoring to final dataset production.

In the development of MMIC™ we are concentrating heavily on self-interviewing using the Internet. We have powerful software available to program questionnaires, including visual displays and visual aides. Using the Internet opens up other possibilities, including real time availability of results and the use of alternative technology for interviewing, allowing respondents to participate using devices like PDAs, Smart phones and WebTVs from virtually anywhere in the world. The use of these devices enables new data collection methods such as medical measurements of blood pressure and heart rate at random intervals. MMIC™ supports the collection of such data.

##### *Advantages of MMIC™ over other software packages*

MMIC™ is designed to overcome many of the limitations inherent in existing survey processing suites like Blaise and Cases, particularly for the kind of large-scale CATI/CAPI questionnaires that are currently being fielded, such as SHARE. In addition, substantially reduced development times are expected to result from the more full-featured set of programming tools that will be provided and from the greatly expanded set of debugging features available in MMIC™. As a consequence, greater accuracy and responsiveness to the needs of the researchers can be obtained and programming overhead can be reduced.

MMIC™ should be seen as a new product, bypassing the shortcomings of existing packages and adding new functionality. MMIC™ will support two programming interfaces. The first, a (normal) top down programming structure, is build on a programming language that includes objects and integrates easy programming tools like a debugger and a variable viewer. The

MMIC™ programming language allows users to add descriptions to all items in the language, including questions and enumerated answer types. MMIC™ also supports easy ways to simplify the randomization for questions and answer categories.

The second interface implements a web or graphical user interface to program simple questionnaires without complicated routing. This visual questionnaire builder is intended to put greater control directly in the hands of researchers as well as an improved environment for translation work. MMIC™ will also provide extensive built-in non-Western language support. Metadata support for MMIC™ will include tools to facilitate the rapid production of human-readable representations of the instrument for review and documentation, which otherwise are very time-consuming to produce for a large survey. Perhaps most importantly, MMIC™ will offer comparatively better data extraction capabilities than other packages. These include the ability to easily move data, along with corresponding metadata, directly into current statistical analysis packages such as STATA. Finally, due to the fact that much of the MMIC™ program code will be available for modification and continued development by users, MMIC™ promises to be more flexible than existing packages and has much greater potential to adapt to future changes in the needs of survey research.

MMIC™ integrates the layout of the questionnaire with the metadata and exports to a uniform central (meta) database. Data entry programs and a sample management system for all modes will be included as well as support for survey control utilities.

Furthermore, there are (advanced) plans to include tools in MMIC™ for metadata viewing, generating codebooks and exporting well-documented files of the survey responses that are ready to load into STATA and other statistical analysis software, or spreadsheet/database programs.

#### *Hardware and software requirements for MMIC™*

MMIC™ is platform (and hardware) independent. It can be run on any device that can run a java virtual machine such as laptops with Linux or Windows or even PDAs.

Although MMIC™ is still under development, it is already being used in a number of applications, including Internet interviewing in the RAND projects and the fielding of SHARE, which is being conducted in 12 countries with 17 different languages and 5 different scripts (including Cyrillic, Greek, Arabic, and Hebrew). We propose further development of MMIC™ as part of the facility and in response to the needs of the researchers using the facility.

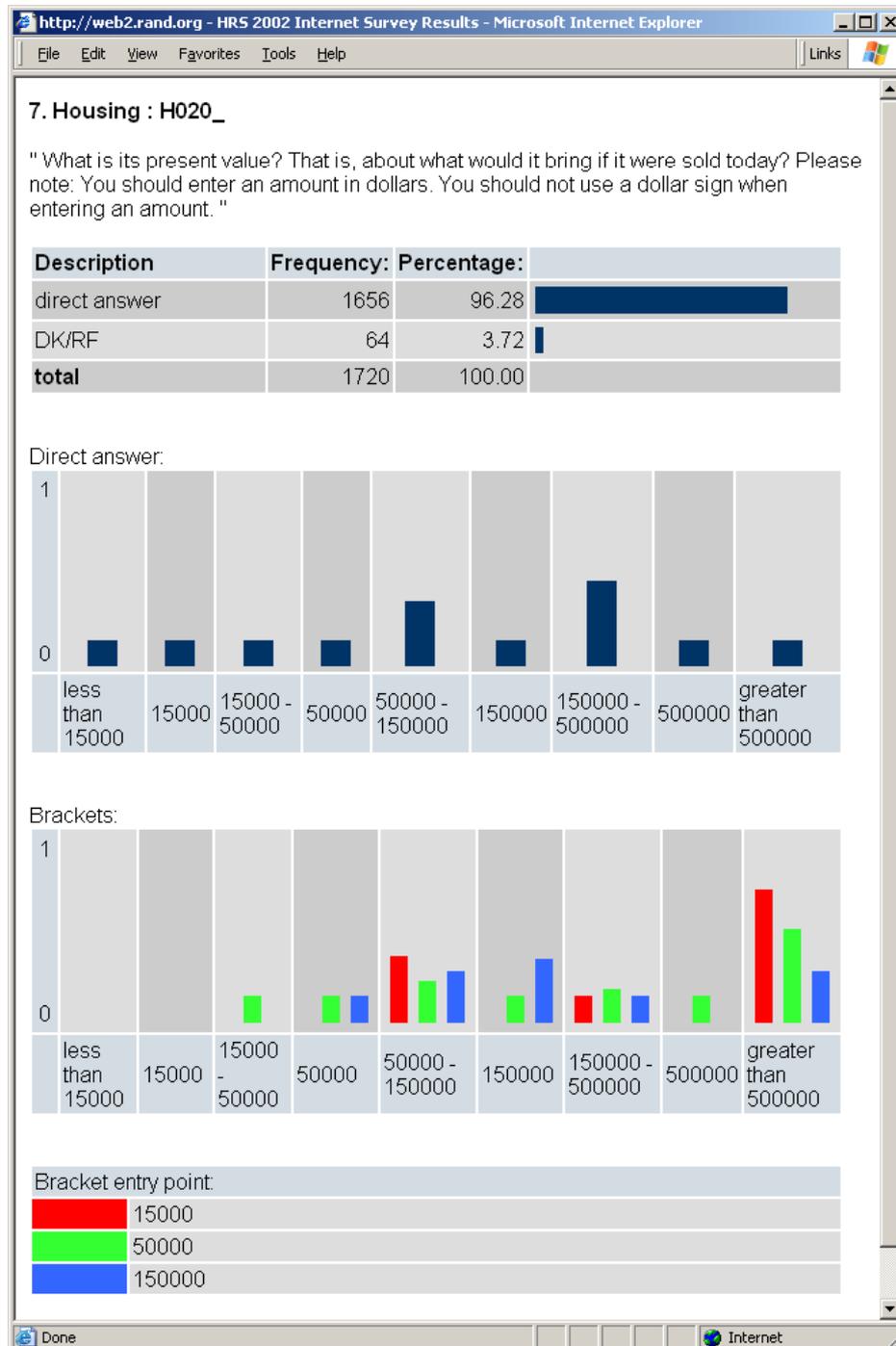
#### **D.7. Data dissemination**

Data collected in the facility will be freely available subject only to privacy and confidentiality restrictions. MMIC™ facilitates the real time availability of data. That is, whenever a respondent completes a survey, the resulting data are added to a database, and a researcher has access to the database in real time. In addition, during the field period simple descriptive statistics are updated in real time (like simple frequencies or graphs).

In Figure 2 we show a somewhat complicated example of what such statistics look like, taken from the “Internet interviewing and the HRS” project. Respondents are asked the value of their house. Most respondents (96.28%) answer this question directly. Respondents who answer “don’t know” or “refuse” get follow-up unfolding bracket questions, where the entry points vary randomly across respondents. The top graph gives the histogram of the direct responses, whereas the bottom graph presents the answers to the bracket questions, with colors indicating the entry points. We observe that in particular respondents with expensive houses are more likely to answer “don’t know” or “refuse”. There does not seem to be a

dramatic “anchoring effect”, that is the answers to the bracketing questions are not systematically related to the choice of entry point.

**Figure 2: An example of a real-time presentation of descriptive statistics**



Once the field period is closed the final dataset is made available to the researcher who initiated the survey. At the same time a codebook with a description of the data is produced. Both the data and the codebook are made available on the Web and can be downloaded by any researcher, subject to signing a data confidentiality statement. In this we follow CentERdata’s practice with respect to the DNB Household Survey (previously the CentER

Savings Survey), of which the data are also made available to the scientific community as soon as they are available (<http://www.uvt.nl/centerdata/dhs/>).

In the development of data dissemination protocols and procedures we expect to collaborate closely with DANS (Data Archiving & Networked Services). DANS is a natural partner for archiving the data and to take responsibility for further dissemination of data beyond the immediate posting of data on the web-site of the laboratory.

#### **D.8. Access procedures and user feedback**

The underlying philosophy of the use of the facility is to maximize openness to the scientific community. Not only will data be freely available, we propose to form a committee of overseers from different universities in The Netherlands to select experiments to be carried out in the laboratory. A similar model has been used in the “Laboratory for the Study of Household Financial Decisions”, funded by the Netherlands Organisation for Scientific Research (NWO). In that project a board of 10 overseers from different universities in The Netherlands reviews proposals received by e-mail from interested researchers. Researchers can submit proposals at any time during the year and typically a decision is made in about a week. The board of overseers appoints a “supervisor” from among its midst to monitor progress of the project and to offer advice. This “light” procedure proves to work very well and to be highly efficient. In a rather short period of time several proposals were approved. Examples are a study of risk preferences by Chris Veld (Tilburg University) and Yulia V. Veld-Merkoulova (Erasmus University Rotterdam), a study of short-run variations in households’ financial market expectations by Rob Alessie (Utrecht School of Economics, Utrecht University), Maarten van Rooij (Dutch Central Bank), Michael Hurd (RAND), and Joachim Winter (University of Mannheim), and a study on the elicitation of risk attitudes by Adam Booij and Gijs van de Kuilen (University of Amsterdam). Based on the experiences in the last few years, this approach stimulates: (1) collaboration of researchers from different universities and institutes in The Netherlands, and (2) collaboration between Dutch researchers and researchers from abroad.

In principle, use of the facility is free; we do not anticipate cost sharing or any form of matched funding. It is tempting to require co-funding of any sort, but in practice this greatly complicates the organization of experiments and slows down scientific progress (see for instance, *Kennisambitie en Researchinfrastructuur*, 2005, pp. 29-30).

An important question is if access should be limited to researchers affiliated with Dutch universities and research institutes. We believe the answer to be “NO”. The experience with the “Laboratory for the Study of Household Financial Decisions” teaches us that there is keen international interest in such a facility and the quality of research can only benefit from more researchers gaining experience in using such a facility. Conceivably, a review committee would give priority to researchers from Dutch institutes in cases where capacity is limited.

The facility will also be available for government and non-profit institutions. However, this may not crowd out the use of the facility by academic researchers nor may it increase respondent burden too much. Government and non-profit institutions will be charged a fee, comparable to the current (reduced) rates for scientific use of the CentERpanel. This fee covers the handling of the questionnaire (programming/testing, supervision, and data delivery) and will partly be used to increase the budgets for focus and control groups, special samples, the use of non-interview measurements and special experimental set-ups, and experiments with incentives and other measures to maximize response rates (see also Section E). No use by commercial entities is envisaged.

A user committee will be formed to provide feedback on the operation of the facility and to suggest improvements and new investments. We plan to organize yearly workshops or conferences. Topics will be selected in consultation with the co-applicants and user groups. Experts from around the world will be invited to participate.

#### **D.9. A parallel facility in the U.S.**

As mentioned in Section C, RAND is building an on-line panel of 1,500 households in the U.S. with respondents over 40. The set-up is quite similar to the CentERpanel and already researchers have exploited the common set-up to do internationally comparative research (e.g. Kapteyn, Smith and Van Soest, 2004). This proposal does **not** request funds for a parallel facility in the U.S. We anticipate a gradual growth of the U.S. facility by acquiring funds from U.S. sources, including the National Institutes of Health and the National Science Foundation. The RAND facility also uses MMIC™, which greatly facilitates the implementation of similar experiments. One can easily migrate survey instruments from one panel to the other; the only thing needed is translation of the texts of the instrument. Having a parallel facility in a different country importantly enhances the scope for cross-cultural and cross-institutional research. For instance, the study by Kapteyn, Smith and Van Soest (2004) considers differences in self-reported work disabilities, using vignettes (cf. King et al., 2004). They find that differences in response scales to the same questions are an important explanation for differences in self-reported work limitations across the two countries.

Ideally, similar facilities would emerge in other countries as well. We believe that successful facilities in The Netherlands and in the U.S. will spawn similar initiatives in other countries. This will give researchers in The Netherlands a leading role in the development of scientific infrastructure in other countries.

#### **D.10. Examples of data collection and experiments**

We anticipate the laboratory to make available a total of approximately 300 minutes interview time per respondent per year (next to about 150 minutes for government and non-profit institutions who would be paying customers; see D.8 above and Section E). Partly this can be filled with questions that are the same every year and partly questions may vary according to the experiment or research topic a researcher is interested in. To fix thoughts, currently respondents to the CentERpanel spend about 1000 minutes a year answering questions or participating in experiments. The respondents to the CentERpanel essentially do not get paid (they only receive a reimbursement for being online), whereas we propose to pay respondents (on average) €15 per 60 minutes of interview time. The power of the Internet is that the 300 minutes (5 hours) can be spread out over many interviews.

Although this proposal does not aim at prescribing the content of the surveys and experiments, it is useful to consider some possibilities. For instance, one can imagine that part of the 5 hours will be filled with topics that are covered in on-going or recently terminated surveys. Examples would be the Socio-Economic Panel (SEP, conducted by Statistics Netherlands until 2002), a General Social Survey (GSS, with content similar to the European Social Survey), the Netherlands Kinship Panel Study (NKPS), the DNB Household Survey (DHS), the European Values Study, and the National Election Survey (Nationaal Kiezersonderzoek). These surveys cover a broad range of social and economic issues. They also have considerable overlap. In fact, by combining these surveys (and using administrative data so that one can skip most questions about household composition, income, labor market, etc.), one would probably be able to conduct a major part of these surveys in about 3 hours total. This would then leave about 2 hours for other topics.

This example brings out a number of features of the proposed set-up:

1. The facility is very cost effective. Despite generous compensation of respondents and high tech infrastructure the total cost of combined surveys (like mentioned above) would be substantially below the total cost of the separate surveys, while there still remains ample room for further experiments and data collection. The generous incentives and extensive quality control should moreover yield much richer and more reliable data.
2. There may be a concern that the interview burden will lead to some undesirable response effects. Our experience until now points in the opposite direction. Experienced panel members are simply better respondents and as a result the data are less noisy. Having said that, the facility allows for various checks of data quality as indicated in Section D.5 and these checks will help to validate the data we are collecting.
3. Since so much information is already collected in a systematic way, additional surveys can be very cost effective. There is no need to collect any background variables, since these are already known.

Examples of experiments have already been discussed in Section C. Here we mention one additional example in the area of saving for retirement that is actually being considered in the RAND panel.

In making decisions, individuals often look to others for assistance. This may be particularly prevalent when decisions are as complex as saving for retirement. One might learn something from observing the behavior of others, or one might seek advice from someone who has faced the same decision problem. In the experiment, the role of information extracted from the observed choices made by other investors is examined. Initially subjects will face a simple decision problem that replicates one faced by an investor: a choice between a riskless asset and a risky one. The novelty of the way the experiment is set up is not only that we can let a population representative sample perform tasks that are usually given to subjects in laboratory settings; we will also allow for respondents to see the choices made by others. In a later stage interaction between respondents is allowed by setting up an Internet discussion board allowing respondents to share information and discuss strategies. In yet a later stage we plan to let respondents trade in “shadow stocks” i.e. they can affect pseudo trades in existing stocks and their compensation will be a function of how well the stocks are doing in the market.

## **E. Budget and organization**

This section presents the budget (E.1.) and the organizational structure (E.2.).

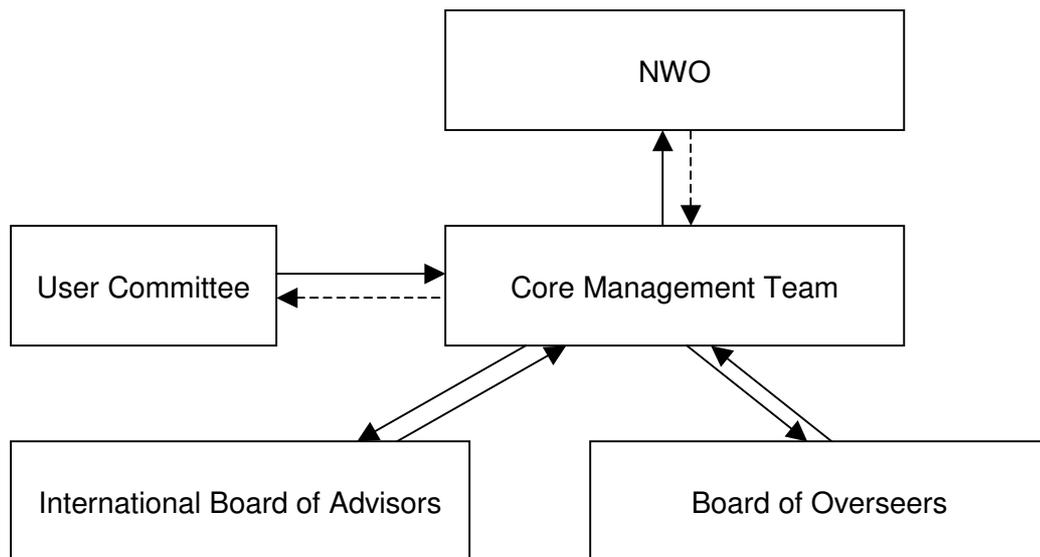
### **E.1. Budget**

*<details are available upon request>*

### **E.2. Organizational structure**

Figure 3 summarizes the organizational structure. The principal investigators (PIs) will form the Core Management Team of the facility. They report to NWO about progress, communicate with the Board of Overseers and keep in close contact with the International Board of Advisors. Feedback given by a User Committee will be communicated with the Core Management Team.

**Figure 3: Graphical representation of the organizational structure**



A brief and somewhat tentative description of the main tasks of the various organizational units is given below:

International Board of Advisors:

- advises on the design of the facility;
- provides feedback on newly planned investments and makes recommendations for changes and additions;
- reviews both the scientific and societal contribution of the facility.

The Board of Overseers:

- oversees the design of the facility;
- invites and selects proposals for experiments and data collection (including proposals for structural long term data collection; e.g. a new SEP);
- organizes feedback from social scientists in The Netherlands (and possibly elsewhere) about the desired design of the facility.

The Board of Overseers will (initially) consist of the co-applicants and will be chaired by one of the PIs. For organizational efficiency, we anticipate a maximum number of 10 members. The Board meets at least once a year to discuss the progress and latest developments. In between meetings, frequent communication will take place via e-mail, conference calls or video conferencing.

The Core Management Team will divide its tasks as follows:

Marcel Das bears responsibility for the supervision of the Internet panel (D.1.), links with administrative data (D.4), and data dissemination (D.7). He will also supervise the user feedback part (D.8).

Arie Kapteyn supervises the use of focus groups, control groups, and special samples (D.2), as well as the use of non-interview measurements and special experimental set-ups (D.5).

Arthur van Soest will chair the Board of Overseers and is responsible for the access procedures (D.8). He will supervise the quality control procedures (D.3).

In view of the existing agreement between CentERdata and RAND, Das and Kapteyn will supervise the further development of MMIC™ (software development, D.6).

## **F. Comparison with “BIG” criteria**

1. *Science Case:* Measurement and experimentation in the social sciences has been hampered by relatively primitive and costly technology. The proposal aims at a unique new environment for social science research. This will provide optimal conditions for scientific breakthroughs. Providing optimal conditions is of course only a necessary and not a sufficient condition. We believe that the group of principal investigators, co-applicants and international advisors are eminently qualified to make sure that the best possible researchers will take advantage of the laboratory. A crucial condition for this is that the facility is open to researchers from anywhere in the world. Quality should be the dominant consideration in deciding who uses the facility, more than nationality.
2. *Talent Case:* The proposed facility goes well beyond anything existing elsewhere in the world and yet is entirely feasible. It will not only provide Dutch researchers with a comparative advantage, it will make it extremely attractive for foreign researchers to use the facility. Even if these foreign researchers will not move to The Netherlands as a result, social sciences in The Netherlands will become a focal point in the scientific community. As a result we expect research in the social sciences in The Netherlands to become more exciting and hence it to be more attractive for researchers to work in the country.
3. *Innovation Case:* An important aspect of the proposed facility is quick turn-around. This will allow for much faster response by the scientific community to policy or business issues. Already, the CentERpanel is capable of producing data in a few weeks time. This allows for quick responses to the need for addressing urgent societal or business problems. In general, one can imagine a variety of societal and business problems that can be addressed in the laboratory. As a signal of this, we notice that the president of the Dutch Central Bank has agreed to be a member of our advisory board. CentERdata already collaborates with the Central Bank in conducting the annual DNB Household Survey. The new facility will only expand the possibilities for understanding financial decision making of Dutch citizens; a highly relevant societal issue in a time where individual responsibility for one's own financial well-being is increasing fast. The findings in the laboratory should therefore be of direct relevance for the several large financial conglomerates in The Netherlands. This is only an example; there are many others. In general the panel will accumulate an unprecedented amount of information on individuals; information of unprecedented quality moreover. We have earlier mentioned topics like the interaction between socio-economic status and physical and mental health, well-being, and social networks. These and other topics are directly relevant for policy and the laboratory can provide a solid and easily accessible knowledge base on which to base policy. The plan to also let government and non-profit organizations use the facility strengthens the societal relevance of the laboratory.
4. *Partnership Case:* The facility is open to all researchers. It is a public good with maximal scientific and societal pay-off. We are working on a parallel facility in the U.S. We anticipate that in the future other countries will follow our example, which will only increase the scientific and societal benefits. In the mean time, based on the experiences with the “Laboratory for the Study of Household Financial Decisions” mentioned in Section C, we anticipate researchers from around the world to use of the facility. This will

give The Netherlands a prominent position, as argued under point 2 above, but at the same time it will foster cooperation between researchers from different countries and different disciplines. One anticipated effect will be a bigger international network of Dutch researchers.

5. *Business Case:* As argued repeatedly, the proposed set-up is highly cost effective. It would be much more costly to collect the same amount of information in more traditional or decentralized ways and the quality of the information would be inferior. For example, the total annual cost of the laboratory is approximately € 2 million. Carrying out three or four of the surveys mentioned in Section D.10 separately already costs more than € 2 million, while the opportunities for analysis are considerably less. Moreover, integrating the questionnaires in one panel, still leaves substantial room for further experiments and data collection.
6. *Technical Case:* The proposed set-up builds on extensive experience of the principal investigators. We have experience with each of the elements of the proposed facility, both technically and organizationally. The main difference with our current activities is the larger scale of the proposed facility and the integration of various elements, which actually simplifies the organization. We will of course keep developing new capabilities. The technical structure of the laboratory, including the software underlying the data collection, can easily accommodate new types of information and new ways of collecting data.

The principal investigators did not apply for an “Investment Subsidy NWO Large” (deadline September 1, 2005) since the budget needed for the new facility would not have fitted in this program.

## **G. Principal investigators, co-applicants and consultants**

### **Principal investigators:**

**Dr. J.W.M. (Marcel) Das** is director of CentERdata since September 2000. Before joining CentERdata he was a senior researcher at CentER Applied Research, an institute for applied economic research at Tilburg University. As a director of CentERdata he has managed a large number of research projects, ranging from small to large-scale projects. In 2003 he received a grant from NWO for the project entitled “A Laboratory for the Study of Household Financial Decisions”. Since 2002 he supervises CentERdata’s role in the SHARE project, which involves the development of various survey instruments (for details, see Das, Vis and Weerman, 2005).

As a senior researcher at CentER Applied Research Das participated in a substantial number of contract research projects, ranging from analyzing the impact of trade- and exercise limitations on the valuation of employee stock options (commissioned by the Ministry of Finance) to the analysis of games with elements of skill (commissioned by e.g. the examining judge at the court of Amsterdam, The Netherlands). Some of the contract research projects were continued in a more fundamental setting which led to publications in refereed journals and books (Van Soest and Das, 2000, 2001). All projects in the area of gambling (with elements of skill) were collected in Van der Genugten, Das and Borm (2001).

As an applied econometrician Das has worked on the analysis of subjective data, which resulted in several publications in international refereed journals (Das and Van Soest, 1997, 1999, Das and Donkers, 1999, and Das, Dominitz, and Van Soest, 1999). His current research interests include design issues in web-based interviewing and micro-econometric modelling of household behavior.

**Prof. dr. A.H.O. (Arthur) van Soest** is both professor of econometrics at Tilburg University and senior economist at RAND. He is an applied econometrician with considerable experience in analyzing subjective data. He has worked on explaining income expectations and on comparing income expectations with realizations (Das and Van Soest, 1997, 1999; Das, Dominitz and Van Soest, 1999) using panel data models. He has analyzed stated preference data and other subjective data on risk aversion and time preferences (Donkers and Van Soest, 1999; Donkers, Melenberg and Van Soest, 2001) and has studied deviations between actual working hours and desired working hours and the predictive value of such deviations for future changes in hours worked (Euwals and Van Soest, 1999; Euwals, Melenberg and Van Soest, 1998). Other studies using panel data include Alessie, Hochguertel and Van Soest (2004), analyzing dynamic panel data models for household portfolio composition, and Dustmann and Van Soest (2001, 2002), studying the relation between immigrant wages and language proficiency using panel data, focusing on the role of measurement error.

Much of his recent work aims at applying flexible semiparametric models to cross-section and panel data, such as Charlier, Melenberg and Van Soest (2000, 2001); Dustmann and Van Soest (2004); Van Soest, Das and Gong (2002) and Gong, Van Soest and Zhang (2005).

Van Soest was involved in the first wave of SHARE as country team leader for The Netherlands and as working group leader for Data Validation and Data Base Management. Together with Smith and Kapteyn, he worked on work disability vignettes for an NIA funded project for the NBER disability group, and designed the vignette questionnaire for the additional sample funded by NIA. He is PI on a project on work disability using vignettes that is part of an NIA funded program on Health and Economic status in Older Populations.

**Prof. dr. ir. A. (Arie) Kapteyn** joined the RAND corporation as a senior economist in January of 2001, from Tilburg University in The Netherlands. At Tilburg University he served the university in many capacities, including dean of the Faculty of Economics and Business Administration, founder and director of CentER and of CentERdata. He also was director of CentER Applied Research. He is currently the director of RAND's Labor and Population program, of the Population Research Center and of the Roybal Center for Financial Decision Making. He is a member of the core management team of SHARE. Kapteyn is a pioneer in the use of subjective variables in economic modeling (e.g. Kapteyn and Wansbeek, 1985).

His current interests include household saving behavior, particularly for retirement (e.g. Alessie, Lusardi, Kapteyn, 1999; Kapteyn and Teppa, 2003), experimental analysis of financial decision making (e.g. Gneezy, Potters, and Kapteyn, 2003), the measurement of disability using vignettes, the relation between health and retirement, data quality and the pros and cons of linking survey data with administrative data (Kapteyn and Ypma, 2005). He is co-author of a widely cited chapter in the Handbook of Econometrics on latent variables models (Aigner, Hsiao, Kapteyn and Wansbeek, 1984). He was a member of the recent National Research Council panel on a research agenda and new data for an aging world. He is the PI of the "Internet Interviewing and the HRS". In 1994 he was elected fellow of the Econometric Society.

**Co-applicants:**

**Prof. dr. J.G. (Jelke) Bethlehem** is doing scientific research on survey methodology at Statistics Netherlands and is professor in the Department of Quantitative Economics of the University of Amsterdam. His research interests are in survey research methodology,

nonresponse correction techniques, computer-assisted data collection, electronic dissemination of statistical information, and statistical graphics.

**Prof. dr. ir. B.G.C. (Benedict) Dellaert** is Meteor Research Chair and professor at the Department of Marketing, Maastricht University. He is an expert in the area of stated preference techniques and interactive choice experiments. His research interests are in consumer decision making, consumer-producer interaction, and retailing and tourism.

**Prof. dr. J.A.P. (Jacques) Hagenaars** is professor in the Department of Methodology and Statistics at Tilburg University. He is chair of the board of the Interuniversity Graduate School of Psychometrics and Sociometrics (IOPS) and member of the Kuratorium of GESIS, the German social science infrastructure. His research interests are in comparative values studies, survey analysis, categorical data analysis and especially categorical latent variable models.

**Prof. dr. M. (Matthijs) Kalmijn** is professor in the Department of Social Cultural Sciences at Tilburg University. He holds a professorship with the label family and the life course. Kalmijn is a senior member of 'Research Team Netherlands Kinship Panel Study'. His research focuses on several subareas in sociology: (a) marriage formation, partner choice, and divorce, (b) intergenerational solidarity and intergenerational reproduction, (c) the relation between demographic and socioeconomic dimensions of the life course.

**Prof. dr. P. (Peter) Kooreman** is a professor of economics at the University of Groningen. His research focuses on applied microeconomics and empirical econometrics. Areas of application include consumer and household behavior, public choice, and health, environmental, and transportation economics. He has published widely in international economic journals.

**Prof. dr. J.P. (Johan) Mackenbach** received a Medical Doctor's degree and a PhD in Public Health from Erasmus University in Rotterdam, The Netherlands. He is also a registered epidemiologist and public health physician. He is a member of the Dutch Health Council and of the Dutch Health Research Council. His research interests are in social epidemiology, medical demography, and health services research.

**Prof. dr. W.E. (Willem) Saris** is professor at the faculty of Social and Behavioral Sciences of the University of Amsterdam and has also a position as ICREA research professor at the Ramon Llull University in Barcelona. He is chairman of the European Survey Research Association (ESRA) and member of the central coordinating team of the European Social Survey. His specialization is methodology for social science research. He developed the telepanel data collection method in 1986, which is a fore-runner of Web surveys. His recent work concentrates on the evaluation and improvement of survey methods.

#### **International Board of Advisors:**

**Prof. dr. M.P. (Mick) Couper** is a Senior Associate Research Scientist in the Survey Research Center, and an Adjunct Associate Professor in the Department of Sociology, both at the University of Michigan. He is also a Research Associate Professor in the Joint Program in Survey Methodology (JPSM). He has over fifteen years of experience in the design, implementation and analysis of survey research on a variety of topics. His current research focuses on nonresponse, the role of the interviewer, and computer assisted survey data collection (including CATI, CAPI, audio-CASI, and Web surveys). He is the co-author (with Robert Groves) of *Nonresponse in Household Interview Surveys* (Wiley, 1998), and chief editor of *Computer Assisted Survey Information Collection* (Wiley, 1998). He has published numerous other articles in a variety of journals.

**Prof. dr. D.A. (Don) Dillman** is Regents' Professor and the Thomas S. Foley Distinguished Professor of Government and Public Policy in the Departments of Sociology and Community and Rural Sociology at Washington State University. He also serves as Deputy Director for Research and Development in the Social and Economic Sciences Research Center

(SESRC). He is recognized internationally as a major contributor to the development of modern mail, telephone and Internet survey methods. In 1970, he was founding coordinator of the SESRC's Public Opinion Laboratory (1970-1973), one of the first university-based telephone survey laboratories in the United States. In 1991 he was appointed (under the Intergovernmental Personnel Act) as the senior survey methodologist in the Office of the Director, U.S. Bureau of the Census, a position he held until 1995, where he provided leadership for the development of new questionnaire designs and procedures for the 2000 Decennial Census and other government surveys.

**Prof. dr. D. (Daniel) Kahneman** is Eugene Higgins Professor of Psychology, Princeton University, Professor of Public Affairs at Princeton's Woodrow Wilson School, and a Fellow of the Center for Rationality, Hebrew University, Jerusalem. His many honors include the 2002 Nobel Prize in Economics, the Grawemeyer Prize in Psychology, the Distinguished Scientific Contributions Award of the American Psychological Association and numerous honorary degrees. He is a member of the National Academy of Sciences and a fellow of the American Academy of Arts and Sciences, the Econometric Society, American Psychological Association, and the American Psychological Society, among others. His research focuses on judgment and decision making, broadly conceived, and the determinants of well-being. His recent work addressed conceptual and empirical distinctions between experienced and remembered utility and decision utility.

**Prof. dr. D.L. (Daniel) McFadden** is a Professor of Economics at the University of California, Berkeley. He has designed and administered a number of household surveys and is currently directing the Berkeley Internet Virtual Laboratory (Ivlab), an experimental Internet facility at UC Berkeley. McFadden has published extensively on a wide variety of topics, including the analysis of choice based samples. In 2000 he was awarded the Nobel prize in economics for developing methods for the econometric analysis of individual and household data that are firmly grounded on formal models of economic behavior.

**Prof. dr. N. (Norbert) Schwarz** is Professor of Psychology at the University of Michigan and Research Professor in the Survey Research Center and the Research Center for Group Dynamics at Michigan's Institute for Social Research. He also holds appointments as Professor of Marketing in the Ross School of Business at the University of Michigan and Research Professor of Survey Methods in the University of Michigan/University of Maryland Joint Program in Survey Methods. His research interests focus on human judgment and cognition, including the interplay of feeling and thinking, the socially situated nature of cognition, and the implications of basic cognitive and communicative processes for public opinion, consumer behavior and social science research.

**Dr. mr. A.H.E.M. (Nout) Wellink**, President of De Nederlandsche Bank (DNB, Dutch Central Bank) since July 1997, has been a member of the Governing Board since 1 January 1982. Before joining DNB, he followed a career at the Ministry of Finance from 1970 until 1982, holding successive posts as member of the economics staff, Executive Director of the General Financial and Economic Policy Department and Treasurer-General. From 1965 to 1970, Wellink worked as a student research assistant and later as a researcher at Leiden University, where he had been reading Dutch Law from 1961. Wellink obtained his doctorate from the Economics Faculty at the Erasmus University, Rotterdam, in 1975. He completed his secondary education (science stream) in 1961.

## H. References

- Adams, P., Hurd, M.D., McFadden, D., Merrill, A. and T. Ribiero (2003), "Healthy, Wealthy and Wise? Tests for Direct Causal Paths between Health and Socioeconomic Status," *Journal of Econometrics* 112, 3-56.
- Aigner, D.J., C. Hsiao, A. Kapteyn and T.J. Wansbeek (1984), "Latent Variable Models in Econometrics," in: Z. Griliches and M.D. Intriligator (eds.), *Handbook of Econometrics Vol. II*, North-Holland, Amsterdam, 1321-1393.
- Alessie, R., S. Hochguertel and A. Van Soest (2004), "Ownership of Stocks and Mutual Funds: A Panel Data Analysis," *Review of Economics and Statistics* 86, 3, 783-796.
- Alessie, R.J.M., A. Lusardi and A. Kapteyn (1999), "Saving After Retirement: Evidence from Three Different Surveys," *Labour Economics* 6, 277-310.
- Bellemare, C., S. Kröger and A. Van Soest (2005), "Actions and Beliefs: Estimating Distribution-Based Preferences Using a Large Scale Experiment with Probability Questions on Expectations, IZA discussion paper DP1666, Institute for the Future of Labor, Bonn.
- Belli, R.F. (1998), "The Structure of Autobiographical Memory and the Event History Calendar: Potential Improvements in the Quality of Retrospective Reports in Surveys," *Memory* 6, 383-406.
- Berrens, R.P., A.K. Bohara, H. Jenkins-Smith, C. Silva and D.L. Weimer (2001), "Replacement Technology or Meaningless Data? How Close Are Meaningful Internet Surveys," working paper, University of New Mexico.
- Beukenhorst, D.J. (1999), "Vertekening door Non-Respons in het WBO," internal CBS report, Sector Waarnemingsmethodologie, Heerlen.
- Beukenhorst, D.J. (2001), "Rendement van het Mixed-Mode Design van POLS en Vertekening," internal CBS report, Sector Methoden en Ontwikkeling, Heerlen.
- Börsch-Supan, A., A. Brugiavini, H. Jürges, J. Mackenbach, J. Siegrist and G. Weber (eds.) (2005), *Health, Ageing and Retirement in Europe. First Results from the Survey of Health, Ageing and Retirement in Europe*, Mannheim Research Institute for the Economics of Aging (MEA), Mannheim.
- Bollinger, C.R. (1998), "Measurement Error in the Current Population Survey: A Nonparametric Look," *Journal of Labor Economics* 16, 3, 576-594.
- Bound, J., C. Brown, G.J. Duncan and W.L. Rodgers (1994), "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data," *Journal of Labor Economics* 12, 3, 345-368.
- Chang, L. and J.A. Krosnick (2003a), "National Surveys via RDD Telephone Interviewing vs. the Internet: Comparing Sample Representativeness and Response Quality," working paper, Ohio State University.
- Chang, L. and J.A. Krosnick (2003b), "Comparing Oral Interviewing with Self-Administered Computerized Questionnaires: An Experiment," working paper, Ohio State University.
- Charlier, E., B. Melenberg and A. Van Soest (2000), "Estimation of a Censored Regression Panel Data Model Using Conditional Moment Restrictions Efficiently," *Journal of Econometrics* 95, 1, 25-56.
- Charlier, E., B. Melenberg and A. Van Soest (2001), "An Analysis of Housing Expenditures Using Semiparametric Models and Panel Data," *Journal of Econometrics* 101, 1, 71-107.
- Couper, M.P. (2000), "Web Surveys. A Review of Issues and Approaches," *Public Opinion Quarterly* 64, 464-494.
- Couper, M.P., A. Kapteyn, M. Schonlau and J. Winter (2004), "Noncoverage in an Internet Survey," *Social Science Research*, forthcoming.
- Crawford, S., S. McCabe, M.P. Couper and C. Boyd (2002), "From Mail to Web: Improving Response Rates and Data Collection Efficiencies," presented at the International Conference on Improving Surveys, Copenhagen, Denmark.
- Das, M. and B. Donkers (1999), "How Certain are Dutch Households About Future Income? An Empirical Analysis," *Review of Income and Wealth* 45, 325-338.

- Das, M. and A. Van Soest (1997), "Expected and Realized Income Changes: Evidence from the Dutch Socio-Economic Panel," *Journal of Economic Behavior and Organization* 32, 137-154.
- Das, M. and A. Van Soest (1999), "A Panel Data Model for Subjective Information on Household Income Growth," *Journal of Economic Behavior and Organization* 40, 409-426.
- Das, M., J. Dominitz and A. Van Soest (1999), "Comparing Predictions and Outcomes: Theory and Application to Income Changes," *Journal of the American Statistical Association* 94, 75-85.
- Das, M., C. Vis and B. Weerman (2005), Developing the Survey Instruments for SHARE, in: A. Börsch-Supan and H. Jürges (eds.), *The Survey of Health, Ageing and Retirement in Europe – Methodology*, Mannheim, MEA.
- Dennis, J.M. and R. Li (2003), "Effects of Panel Attrition on Survey Estimates," paper presented at the 2003 meeting of American Association for Public Opinion Research (AAPOR).
- Dominitz, J. and C.F. Manski (1996), "Eliciting Student Expectations of the Returns to Schooling," *Journal of Human Resources* 31, 1-26.
- Donkers, B. and A. Van Soest (1999), "Subjective Measures of Household Preferences and Financial Decisions," *Journal of Economic Psychology* 20, 6, 613-642.
- Donkers, B., B. Melenberg and A. Van Soest (2001), "Estimating Risk Attitudes Using Lotteries - A Large Sample Approach," *Journal of Risk and Uncertainty* 22, 2, 165-195.
- Dustmann, C. and A. Van Soest (2001), "Language Fluency and Earnings: Estimation with Misclassified Language Indicators," *Review of Economics and Statistics* 83, 4, 663-674.
- Dustmann, C. and A. Van Soest (2002), "Language and the Earnings of Immigrants," *Industrial and Labor Relations Review* 55, 3, 473-492.
- Dustmann, C. and A. Van Soest (2004), "An Analysis of Speaking Fluency of Immigrants Using Ordered Response Models with Classification Errors," *Journal of Business and Economic Statistics* 22, 3, 312-321.
- Euwals, R., B. Melenberg and A. Van Soest (1998), "Testing the Predictive Value of Subjective Labour Supply Data," *Journal of Applied Econometrics* 13, 5, 567-586.
- Euwals, R. and A. Van Soest (1999), "Desired and Actual Labor Supply of Unmarried Men and Women in The Netherlands," *Labor Economics* 6, 95-116.
- Finch, C.S., J.W. Vaupel and K. Kinsella (eds.) (2001), *Cells and Surveys: Should Biological Measures be Included in Social Science Research?*, National Research Council, Washington D.C.
- Gneezy, U., J.J.M., Potters and A. Kapteyn (2003), "Evaluation Periods and Asset Prices in a Market Experiment," *Journal of Finance* 58, 821-837.
- Gong, X., A. van Soest and P. Zhang (2005), "The Effects of the Gender of Children on Expenditure Patterns in Rural China: A Semiparametric Analysis," *Journal of Applied Econometrics* 20, 4, 509-527.
- Groves, R.M. (1989), *Survey Errors and Survey Costs*, New York: Wiley.
- Hadi, A.S. (1992), "Identifying Multiple Outliers in Multivariate Data," *Journal of the Royal Statistical Society, series B* 54, 761-771.
- Hadi, A.S. (1994), "A Modification of a Method for the Detection of Outliers in Multivariate Samples," *Journal of the Royal Statistical Society, series B* 56, 393-396.
- Hill, D.H. (1994) "The Relative Empirical Validity of Dependent and Independent Data Collection in a Panel Survey," *Journal of Official Statistics* 10, 359-380.
- Hill, D.H. (2002), "Wealth Dynamics: Reducing Noise in Panel Data," working paper, University of Michigan.
- Hill, D., M. Perry and R.J. Willis (2004), "Do Internet Surveys Alter Estimates of Uncertainty and Optimism about Survival Chances?," working paper, University of Michigan.
- Hirano, K., G. Imbens, G. Ridder and D. Rubin (2001), "Combining Panel Data Sets with Attrition and Refreshment Samples," *Econometrica* 69, 1645-1659.
- Hurd, M. and A. Kapteyn (2003), "Health, Wealth and the Role of Institutions," *Journal of Human Resources* 38, 386-415.

- Jasso, G., M.R. Rosenzweig and J.P. Smith (1999), "The Effects of Interview Payments and Periodicity on Sample Selection and Attrition and on Respondent Memory: Evidence from the Pilot Study of the New Immigrant Survey," RAND working paper.
- Juster, F.T. and F.P. Stafford (eds.) (1985), "*Time, Goods, and Well-Being*, University of Michigan Press, Ann Arbor, MI.
- Kahneman, D., A.B. Krueger, D.A. Schkade, N. Schwarz and A.A. Stone (2004), "A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method (DRM)," *Science*, 1776-1780.
- Kapteyn, A., J.P. Smith and A. Van Soest (2004), Self-reported Work Disability in the US and The Netherlands, RAND working paper WR-206.
- Kapteyn, A. and F. Teppa (2003), "Hypothetical Intertemporal Consumption Choices," *The Economic Journal* 113, C140-C152.
- Kapteyn, A. and A. Van Soest (2004), "Asset Ownership and Asset Amounts: Comparing HRS 2002, HRS Internet 2003, RAND Internet MS1 and RAND CATI 2004," RAND working paper.
- Kapteyn, A. and T.J. Wansbeek (1985), "Estimation in a Linear Model With Serially Correlated Errors When Observations Are Missing," *International Economic Review* 26, 469-490.
- Kapteyn, A. and J. Ypma (2005), "Measurement Error and Misclassification: A Comparison of Survey and Register Data," RAND working paper WR-283, Labor and Population Program.
- Kennisambitie en Researchinfrastructuur – Investeren in Grootschalige Kennisinfrastructuur* (2005), Innovatieplatform, July, 29-30.
- King, G., C. Murray, J. Salomon and A. Tandon (2004), "Enhancing the Validity and Cross-cultural Comparability of Measurement in Survey Research," *American Political Science Review* 98, 1, 567-583.
- Lengacher, J.E., C.M. Sullivan, M.P. Couper and R.P. Groves (1995), "Once Reluctant, Always Reluctant? Effects of Differential Incentives on Later Survey Participation in a Longitudinal Study," presented at the Annual Conference of the American Association for Public Opinion Research, Fort Lauderdale, FL, May 18-21.
- Link, M.W. and R.W. Oldendick (1999), "Call Screening: Is It Really a Problem for Survey Research?," *Public Opinion Quarterly* 63, 577-589.
- Mack, S., V. Huggins, D. Keathley and M. Sunduckchi (1998), "Do Monetary Incentives Improve Response Rates in the Survey of Income and Program Participation?," *Proceedings of the Section of Survey Research Methods*, American Statistical Association.
- Nijman, Th. E. and M. Verbeek (1990), "Estimation of Time Dependent Parameters in Linear Models Using Cross Sections, Panels, or Both," *Journal of Econometrics* 46, 333-346.
- Oldendick, R.W. and M.W. Link (1994), "The Answering Machine Generation: Who Are They and What Problem do They Pose for Survey Research?," *Public Opinion Quarterly* 58, 264-273.
- Piekarski, L., G. Kaplan and J. Prestegaard (1999), "Telephony and Telephone Sampling," paper presented at the Annual Conference of the American Association for Public Opinion Research, St. Petersburg, FL.
- Pischke, J.-S. (1995), "Measurement Error and Earnings Dynamics: Some Estimates from the PSID Validation Study," *Journal of Business & Economic Statistics* 13, 3, 305-314.
- Rodgers, W.L., C. Brown and G.J. Duncan (1993), "Errors in Survey Reports of Earnings, Hours Worked, and Hourly Wages," *Journal of the American Statistical Association* 88, 24, 1208-1218.
- Rubin, D.C. (1996) *Remembering Our Past. Studies in Autobiographical Memory*, Cambridge, Cambridge University Press.
- Rubin, D.C. and A.D. Baddeley (1989), "Telescoping is Not Time Compression: A Model of the Dating of Autobiographical Events," *Mem Cognit.* 17, 6, 653-661.
- Scollon, C.N., C. Kim-Prieto and E. Diener (2003), "Experience Sampling: Promises and Pitfalls, Strengths and Weaknesses," *Journal of Happiness Studies* 4, 5-2003.
- Singer, E. and R.A. Kulka (2002), "Paying Respondents for Survey participation," in: M. Ver Ploeg, R.A. Moffitt and C.F. Citro (eds.), *Studies of Welfare Populations: Data Collection and Research Issues*, Committee on National Statistics, National Research Council.

- Smith, J.P. (2003), "Consequences and Predictors of New Health Events," working paper 10063, National Bureau of Economic Research (NBER).
- Sudman, S. and Bradburn, N.M. (1973), "Effects of Time and Memory Factors on Response in Surveys," *Journal of the American Statistical Society* 68, 344, 805-815.
- Te Riele, S. (2002), "Vertekening door Non-Respons," Centraal Bureau voor de Statistiek, working paper.
- Tourangeau, R., Rips, L.J. and Rasinski, K. (2000), *The Psychology of Survey Response*, Cambridge, Cambridge University Press.
- U.S. Department of Commerce, Social and Economic Statistics Administration (1975), "Comparison of Month to Month Changes in Industry and Occupation Codes with Respondent's Report of Change: CPS Job Mobility Study," Response Research Staff Report No. 75-5.
- Van der Genugten, B.B., M. Das and P.E.M. Borm (2001), *Behendig gokken in en rond het casino* (in Dutch), Academic Service, Schoonhoven.
- Van Soest, A. and M. Das (2000), "Family Labour Supply and Proposed Tax Reforms in the Netherlands," in: T. Callan (ed.), *Taxes, Transfers and Labour Market Responses: What Can Microsimulation Tell Us?*, The Economic and Social Research Institute, Dublin.
- Van Soest, A. and M. Das (2001), "Family Labor Supply and Proposed Tax Reforms in the Netherlands," *De Economist* 149, 191-218.
- Van Soest, A., M. Das and X. Gong (2002), "A Structural Labour Supply Model with Flexible Preferences," *Journal of Econometrics*, Annals issue on Information and Entropy Econometrics 107, 345-374.
- Wagenaar, W.A. (1986), "A Study of Autobiographical Memory over Six Years," *Cognitive Psychology* 18, 225-252.
- Zaman, A., P.J. Rousseeuw and M. Orhan (2001), "Econometric Applications of High-Breakdown Robust Regression Techniques," *Economics Letters* 71, 1-8.