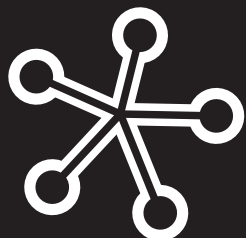




**A WEBDATANET working paper**

# **Mind the gap in the digital data tsunami**

*Pablo de Pedraza and Francisco Serrano*



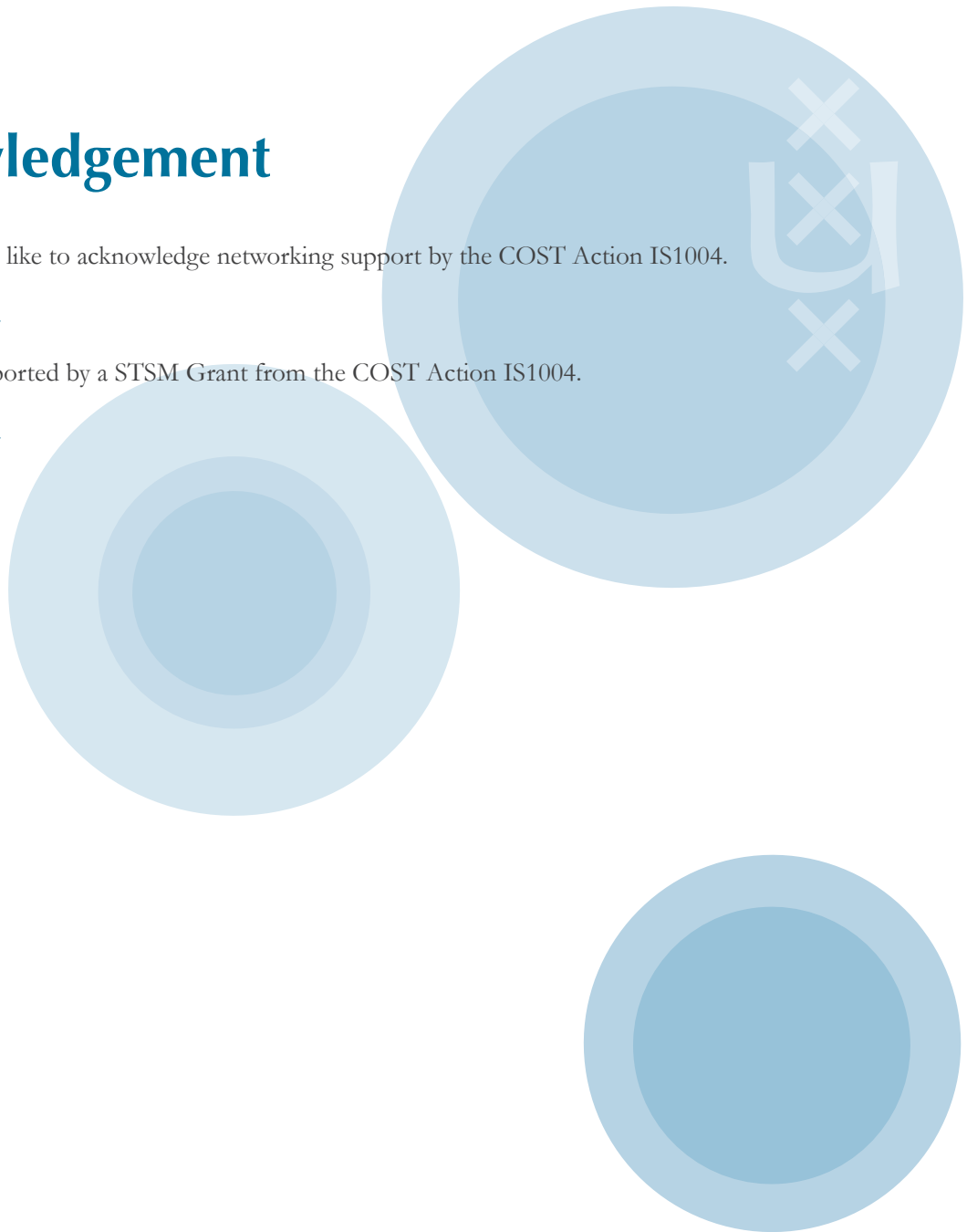
# Acknowledgement

The authors would like to acknowledge networking support by the COST Action IS1004.

[www.webdatanet.eu](http://www.webdatanet.eu)

This work was supported by a STSM Grant from the COST Action IS1004.

[www.webdatanet.eu](http://www.webdatanet.eu)



November 2014

© Pablo de Pedraza, Francisco Serrano

General contact: [aias@uva.nl](mailto:aias@uva.nl)

## **Bibliographical information**

Pedraza, P. de, Serrano, F. (2014). Mind the gap in the digital data tsunami: Representativeness of web-based data collection methods in socio-political studies. University of Amsterdam, AIAS Working Paper 154.

ISSN online: 2213-4980

ISSN print: 1570-3185

Information may be quoted provided the source is stated accurately and clearly. Reproduction for own/internal use is permitted.

This paper can be downloaded from our website [www.uva-aias.net](http://www.uva-aias.net) under the section: Publications/Working papers.

# **Mind the gap in the digital data tsunami:**

**Representativeness of web-based data  
collection methods in  
socio-political studies**

**Pablo de Pedraza**

*AIAS, University of Amsterdam*

**Francisco Serrano**

*Complutense University of Madrid*

**WEBDATANET working paper**

**WP 154**



# Table of contents

|   |           |
|---|-----------|
| ABSTRACT.....   | 7         |
| <b>1 INTRODUCTION.....</b>  | <b>9</b>  |
| <b>2 TOTAL SURVEY ERROR, COVERAGE &amp; DIGITAL DIVIDE, AND COVERAGE DIMENSIONS .....</b> | <b>11</b> |
| 2.1 TSE paradigm .....  | 11        |
| 2.2 Digital divide and Coverage bias of web surveys.....                                  | 12        |
| 2.3 Definition and conceptualization of Internet penetration .....                        | 14        |
| 2.4 Coverage Scale and Dimensions .....   | 15        |
| <b>3 DATA &amp; ESTIMATIONS.....</b>  | <b>19</b> |
| <b>4 RESULTS (SEE TABLE IN APPENDIX) .....</b>  | <b>21</b> |
| <b>5 CONCLUSIONS AND FURTHER RESEARCH .....</b>   | <b>25</b> |
| <b>6 APPENDIX . ODDS RATIO FOR EACH LOGIT MODEL (*p&lt;.05;**p&lt;.01;***p.001).....</b>  | <b>27</b> |
| REFERENCES .....  | 31        |
| AIAS WORKING PAPERS .....   | 35        |



# Abstract

In this paper, we study the representativeness of web-based methods by identifying individual demographic and socio-political determinants of several dimensions of Internet penetration and types of use. As a theoretical framework, we use both the Total Survey Error (TSE) and the theory of the Digital Divide. Models are built using the Spanish CIS databank: Internet usage and penetration variables are used as dependent and demographics and socio-political attitudes as independent variables. Overall findings confirm the existence of key differences between online and offline populations. Women, the elderly, the less educated, lower social classes, rural and religious populations are less likely to have access or to use the Internet. No differences, however, were found by gender regarding Access and E-mail usage; using the Internet for management purposes is more likely among the 25-44 years old than among the 18-25 set. Regarding socio-political attitudes, Internet users are less religious, have a stronger political identity and a more critical view of the political system and mainstream parties. Further research on Internet penetration definitions -places of connection, devices for access and skills of users- is needed given the growing complexity of the scenarios that characterize the access and use of the Internet today.





# 1 Introduction

## Relevance of the study: survey methods and digital data

In the last few years, the number of studies and research projects involving web-based data collection methods has rapidly grown. Researchers from both the qualitative and quantitative *worlds* of social sciences increasingly tend to either produce, collect or analyze data using the Internet. Initially, this interest was largely fostered by the potentialities associated with web surveys in relation to offline -face to face and telephone- surveys. Apart from cost- and time-saving advantages, several authors have also shown other benefits of conducting web surveys, such as a reduction in the social desirability bias, the possibility of making international comparisons, generation of rich paradata and growing landline penetration (Couper 2000; Dillman et al. 2009).

Lately, non-reactive and other methods of web-based data collection different from surveys are becoming widely known among social scientists. The rise of new technology-related research trends such as *opt-in panels* and *big data*, together with the growing perception that it will be possible to collect data on everything that people do, is putting the pressure on and raising concerns about its implications in survey research and the role they can play in the future (Couper 2013). According to some authors, the proliferation of digital data, regardless of whether we call it the “data revolution” or the “digital data tsunami” (Prewitt 2013) may be moving survey methods towards a risk of marginalization which may require an adaption to the new situation (Savage and Burrows 2007).

Some studies have already shown the pros and cons of both types of data (Couper 2013). Survey data has been characterized as *case poor and variable rich*, in opposition to digital data which is defined *case rich and variable poor* (Prewitt 2013; Couper 2013). This suggests that both types of data and the methods associated with them can be compatible and thus integrated harmoniously into the research process, producing mutually enriching methodological knowledge. In this regard, it has been pointed out that surveys can be useful tools for understanding many of the data generated on the Internet, for evaluating their quality and for understanding not only what can and what cannot be concluded from them but also why (Couper 2013). Moreover, the possibility to innovate and develop new modes and methods will only evolve if digital data insights are sufficiently explored.

This paper aims to offer an example of how theory and concepts taken from traditional survey methodology can be applied to non-survey web-based research methods and data. Its ultimate goal is to provide insights and contribute to the current debate about the adequacy, limits, qualities, strengths and weaknesses of data generated on the Internet, either as a byproduct of other activities, such as administrative and transactional activities and social media, or as a result of participation in opt-in or probabilistic survey panels. To do so, this paper builds upon the “total survey error paradigm” and a general idea of representativeness that focuses on the *coverage and selection bias* (Bethlehem 2010) to study the composition of the different dimensions of Internet penetration –access, basic use, frequency and types of use (management- information communication and culture consumption)- regarding socio-demographic and socio-political variables.

On a substantive level, this study focuses on and uses data collected for opinion polls and electoral surveys, for which socio-demographic and socio-political attitudinal bias are of critical importance. For example, voting predictions based on web-based data may fail- if substantial differences in terms of political attitudes are found between individuals with access, different types of users and the rest of the target population. The study of the determinants of different levels of Internet usage –i.e. advanced use or social media use- is also useful in assessing the representativeness of other methods of online research, where the preeminence of advanced users in the production of data has been shown (Couper, 2013). In this sense, our approach will follow the perspective already suggested by the literature (Lavrakas, 2013, Couper 2013, WebDataNet 2009) about the application of the total survey error approach to research methods and techniques other than surveys.

The rest of the paper is organized as follows. In section 2 we introduce the Total Survey Error paradigm, focusing on the coverage error and we relate it to the digital divide theory, from which we draft the coverage dimensions. In section 3 we present the data and the models used for the analysis. In section 4 we describe the main findings extracted from the analysis and in section 5 we make some concluding remarks and suggest further paths for research.

## 2 Total survey error, coverage & digital divide, and coverage dimensions

### 2.1 TSE paradigm

As early as 1941 W. Edwards Deming devised the first ever list about the main problems facing survey research. That list already included nonresponse, sampling, interviewer influence, mode effects and data preparation and management (Groves 2010). Since then, deep changes have profoundly transformed the methodological picture in survey research, affecting researcher interests, populations studied, data collection format as well as theoretical and epistemological assumptions and approaches. In the last 20 years, the most important changes have had to do with the inclusion of automated processes of data collection, management and analysis brought about by the generalization of software tools in the survey field (Berzelak, Lozar Manfreda, & Vehovar 2006). Despite these transformations, Deming's old scheme still represents the basis of the Total Survey Error (TSE) paradigm, as we know it today. This commonly accepted approach to TSE is characterized by a double process of inference, representativeness and measurement, and includes four main sources of survey errors: sampling, nonresponse, measurement and coverage. Recent studies advocate a unified perspective in the application of the TSE methodological framework, applicable to all kinds of research methods in social science, and not only to surveys or quantitative methods (Lavrakas 2013, Couper 2013).

*Sampling error* arises from the fact that *not all members of the frame population are measured. If the selection process were repeated, a slightly different set of sample persons would be obtained* (Couper 2000). In regard to *sampling error* applied to web surveys, statistical inference is only possible when a sample frame of the target population is available and there is a known, non-zero probability for each sampling unit to be included in the sample frame. The main strategies for facing the absence of sample frames have focused on the design of dual sampling frames, the integration of offline sampling frames with online data collection, the creation of online panels with offline recruitment methods or the implementation of post-survey statistical adjustments (Pedraza, Tijdens, Muñoz de Bustillo, Steinmetz 2010; De Leeuw 2011). As regards *nonresponse error*, the continuous fall in response rates represents a major concern in survey research as it is a common feature among all survey data collection modes (Dillman 2009). In online surveys, responses rates are even lower

than those of other modes. Previous studies show that response rates in web surveys stand at 30-35%, about 10 points lower than face-to-face and telephone surveys (Shih & Fan 2008; Lozar Manfreda et al. 2008). Mixed evidence is found in the literature regarding the utility of using incentives or sequential mix-mode designs as solutions to deal with the nonresponse problem. (Görizt 2006; De Leeuw 2011). *Measurement issues* represent the area of web survey research with the greatest potential for practical implementation of general population web surveys. Online survey data collection is characterized by the possibility of including interactive features that provide self-administered and computer-assisted guidance to the interviewer (De Leeuw 2011). Nowadays software tools allow surveyors the possibility to create customized and personalized questionnaires with routes and filters adapted to the respondent, greater flexibility when designing the format of the questions, help functions (De Leeuw 2011), and above all, less social desirability bias derived from the absence of a human interviewer. (De Leeuw & Hox ; Manfreda, Vehovar & Berzelak 2011).

## 2.2 Digital divide and Coverage bias of web surveys

The usefulness of methodological research and analyses of coverage bias, in particular, are of special importance in general population studies, in which the conclusions drawn must be applicable to the whole population and not just to those who have access to or use the web. Despite the popularity attained by online research methods in the last few years, general population offline surveys are still the preferred choice of public institutes and statistical offices. Under-coverage is the main obstacle for the inclusion of web surveys in official bodies as not all the population has access to the Web, and even if they have, in many cases, they might not be very familiar with it and even reluctant to use it (Dillman 2009). Even though nowadays it is possible to track a big portion of what users do on the Internet through non-reactive data or the so-called organic data (Groves 2011) or big data (Couper 2013), the fact is that not everybody uses the Internet and those who do connect to the Web do not make the same uses of it. This is something to bear in mind if we are to understand the potential of Internet data to obtain findings applicable to the general population. For instance, there seems to be a gap among more active Internet users who produce most of the social media data and the rest of more passive (or even lurkers) web user groups (Couper 2013). The existing non-random differences among population groups with and without Web access or among different type of users are the basis of the *coverage and selection bias* (Bethlehem 2010). Concerns raised by the research community about the representativeness of such studies call for further methodological research that can contribute to making social scientists aware of the potential and limitations of the use of web-based data.

Digital divide studies have traditionally been used as a theoretical source to explain these differences through socio-demographic variables such as gender, age, education, income, type of habitat (metropolitan vs. rural), nationality, etc. In addition, other variables related to attitudes, political views and lifestyles have lately been recognized as relevant when identifying and understanding the differences between people who use the Internet and those who do not (Vehovar et al. 1999, Bandilla et al. 2003, Schonlau et al. 2004, Lee 2006, Pedraza et al. 2010). Another key feature is that, as Internet access increases, it becomes necessary to go further and study differences between types of users, in terms of experience, knowledge, frequency, types of contents visited, etc. According to this feature, survey methodology needs to consider not only Internet access variables, meaning having a web connection in the household, but also other variables referring to different types of uses, in order to set the basis for a comprehensive analysis of coverage bias (Lynn 2013).

*Coverage error* in surveys has been defined as a function of the mismatch between the target population and the frame population (Couper 2000). Coverage becomes a central issue for the quality of the survey when this mismatch reflects a difference in rates of Internet access among different demographic groups (Mohorko, De Leeuw and Hox, 2011). The *digital divide theory* has been used as the preferred theoretical source to approach coverage bias in web surveys.

In this regard, three main trends have been identified by the majority of recent studies (Riccardini & Fazio, 2002; Robinson, DiMaggio, Hargittai, 2003; Wilson, Waller & Reiser, 2003; DiMaggio, Hargittai, Celeste & Shafer, 2004; Zhang, Callegaro & Thomas, 2009) as the key features that have characterized the digital divide process over the last few years. These trends provide useful knowledge that could be applicable to the study of web survey coverage and other methods of online data research. 1. The influence of socio-demographic variables as determinants of access and usage of Internet still persists, although the existing divide for each of these variables is decreasing at different paces and even increasing for some of the variables. 2. Beyond socio-demographics, the study and analysis of Internet access determinants has to include the so-called *webographics* or attitudinal variables related to attitudes towards different aspects of life to fully comprehend the nature of the differences between those with and without Internet. 3. As the divide in Internet Access is progressively decreasing, it is necessary to go further and study the differences between users, in terms of experience, knowledge, frequency, types of contents visited, etc.

The use of the *digital divide theory* to study web survey coverage bias is thus based on the identification between Internet penetration (which could serve as the reference for an *ideal* but not yet existing sampling frame) and web survey coverage (De Leeuw 2011). This identification assumes both that people covered in the frame population are those with Internet or Web access (Couper 2000), and that those with access are Web users and thus reachable to fill out a web survey. However, and as mentioned before, the picture of web survey coverage and Internet penetration seems more complex.

## 2.3 Definition and conceptualization of Internet penetration

One of the main problems of using Internet penetration to study coverage bias is the lack of agreement found among sources/researchers in the conceptualization and measurement of Internet penetration. The existence of different empirical approaches and indicators in the measurement of the Digital Divide (Dillman 2009; Dolnicar, Vehovar, and Sicherl 2003; Riccardini & Fazio, 2002) yields very dissimilar numbers depending on the indicator or criteria chosen by each source to measure Internet *access* or *usage*. If we talk about Internet access we may be referring, for instance, to variables as diverse as *types of connections* (regular phone connection, broadband, public WiFi, etc.), *places where the connection is available* (household, work, public places such libraries, cafes, etc.) or *devices used for the connection* (desktop computer, laptop, PDA's, mobile, tablet etc.). If we consider penetration as *use of the Internet* we could refer to *frequency of use* (Last 12 months, last 3 months, daily use, etc.), *type of use* (management, information, communication, culture consumption, social networks, etc.) or *level of use* (basic, intermediate, proficiency etc.).

In Spain, overall figures on Internet access and usage in 2011 show levels of Internet penetration that range from 40% to almost 70% of the total population (EGM, Tatum, Eurobarometer 2005-2011 Internet World Stats, INE), but with a growth of 440 % for the period between 2000 and 2011 ([Internet World Stats](#)). As an example of the above-mentioned problems in the measurement of Internet access/usage, the penetration in Spain varies from 42% of the total population if the indicator selected is *usage yesterday* to 67% if the indicator is *usage in the last three months* or to 61% if we take into account *broadband access*.

In order to provide the analyses with an operative concept of Internet penetration, we set as starting point a broad and general definition of web-based data coverage flexible enough to include any of the above dimensions of Internet penetration: *the portion of the target population that can be reached to fill out a survey –or that can generate data- through the Internet (Couper 2000)*. Such a conceptualization makes it possible to develop several dimensions of Internet access and uses and therefore sev-

eral scales and dimensions of the idea of coverage. By using this definition, several scenarios are taken into consideration. For instance, individuals without Internet access would have a zero probability to be reached to fill out a web questionnaire unless provided with the necessary device and connection. Even those having Internet but not using or not willing to use it, have in principle, no chance to complete an online survey or to generate web-based data. Additionally, those individuals who use the Internet more often are more active and are more familiar with the web environment and would be, in principle, more likely to be accessible to answer the web survey or to create web data.

## 2.4 Coverage Scale and Dimensions

Taking into account the literature about the digital divide and based on the above definition of coverage in web surveys –*the portion of the target population that can be reached to fill out a survey through the Internet (Couper 2000)* – a model was developed for the analysis of coverage and its dimensions in web-based data. We employ a **coverage scale** that is composed of different dimensions of coverage ranging from **Minimum coverage** –having broadband Internet connection in the household- to what is called **Total coverage** –providing an e-mail address- each of which refers to a dependent variable to be used in the statistical analysis (see Table 2). The dimensions included in the model and its description are as follows:

### **Dimension 0. – Internet Access**

*Broadband access in the household* was considered as one of the extreme poles of the coverage scale referring to minimum degree of coverage. Other typologies of Internet connection were discarded as they cannot guarantee a satisfactory completion process of questionnaires that include multimedia – videos, audio, pictures, etc. - without involving nonresponse or dropout episodes.

### **Dimension 1. Internet Use**

#### *Dimension 1.1 - Basic use*

This dimension indicates the minimum level of Internet usage that ensures that the questionnaire can be taken by the respondent. People who have used the Internet in the last 12 months are supposed to have enough of a basic command of Internet usage to fill out a standard questionnaire. This is used to find out the differences in relation to those households with Internet access where not all its members have the aforementioned skills.

### *Dimension 1.2. - Daily use*

According to this dimension, Internet users are those who connect to the Internet on a daily basis. It was considered that there are significant socio-demographic and attitudinal differences worth exploring between those who reported having connected to the Internet at least once in the last year, and those who navigate the web every day.

### ***Dimension 2. Types of use***

This dimension refers to specific uses of the Internet. Within this dimension, a set of five sub-types of uses was established: management, information search, communication, audiovisual consumption, and providing email in a face-to-face survey.

#### *Dimension 2.1. - Management*

Within this dimension a distinction was established between 2.1.1. – *Online bank management* and 2.1.2. – *Other management activities*. Both activities denote at least advanced skills in Internet usage.

#### *Dimension 2.2. - Search for information*

This dimension refers to the usage of Internet for informational purposes, i.e. Wikipedia, newspapers, etc.

#### *Dimension 2.3. – Communication*

This dimension is classified in two sub-dimensions:

2.3.1. – *Use of E-mail*

2.3.2. – *Use of social networks, chats, etc.*

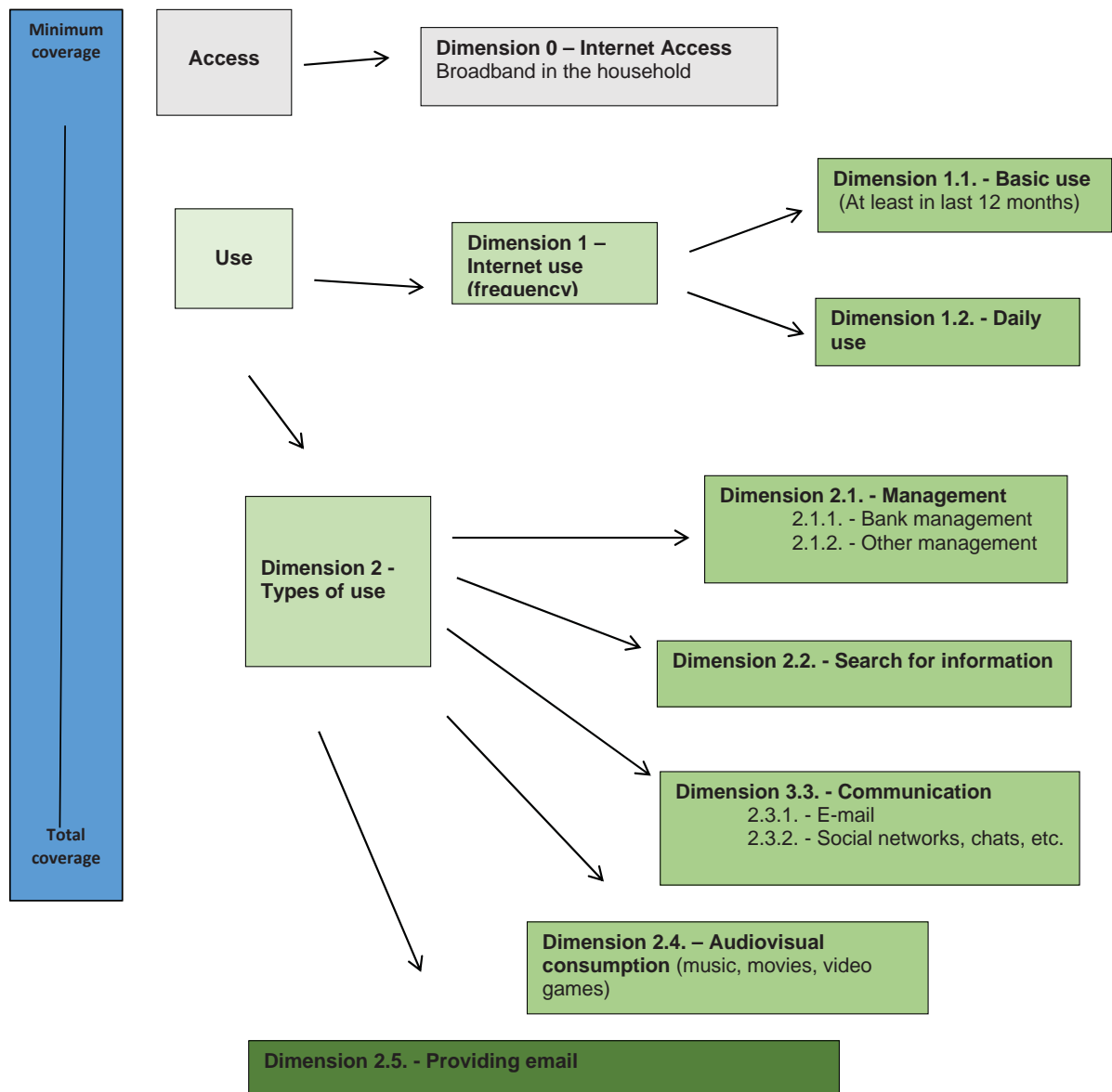
#### *Dimension 2.4. – Audiovisual consumption (music, movies, video games)*

#### *Dimension 2.5. - Providing email*

Dimension 2.5, is considered especially important for web surveys. Although it can be considered a type of use, it goes beyond coverage. It can provide information about attrition in sequential mix modes and the sample that could be obtained in a second email wave. Moreover, it provides information, for example, about coverage bias in a practical implementation of a face-to-face panel with an E-mail follow-up. Moreover, it is a type of use that could also gather some relevant data regarding nonresponse and measurement errors (i.e. social desirability)



Figure 1. Dimensions of the **Coverage scale** /dependent variables for the analysis





## 3 Data & estimations

### *Data source*

In order to fulfill our goal of studying the determinants of coverage dimensions taking into account demographic and political attitudes, the estimations were conducted using secondary data taken from the CIS database. The CIS (*Centro de Investigaciones Sociológicas* [www.cis.es](http://www.cis.es)) is an independent public research institute assigned to the **Spanish Ministry of the Presidency**. The CIS normally carries out monthly or quarterly face-to-face (N=2500) public opinion and political surveys in Spain. It uses a multistage stratified cluster sampling procedure, with random selection of sampling units (census sections) with probabilities of selection proportional to the resident population.

### *Data analysis and building up of the models*

We estimate 10 logistic regressions where the dependent variables are each of the aforementioned coverage dimensions and estimate the impact of socio-political attitudes after accounting for a set of demographic variables. The estimations were done in order to contrast the research hypothesis.

*Table 3. The model built for each of the coverage dimensions*

|   |
|---|
| <p><b><u>Dependent variable</u></b></p> <ul style="list-style-type: none"> <li>- Coverage dimension x</li> </ul> <p><b><u>Independent variables</u></b></p> <ul style="list-style-type: none"> <li>- Socio-demographic variables + religiosity and ideological self-placement + <b>set of socio-political attitude variables</b> commonly included on CIS questionnaires</li> </ul> |
|---|

The N of each model varies in relation to the availability of variables in each dataset. Data interpretation and comparison between dimensions should be done cautiously. It is neither the goal of this study to show the concrete effects of the determinants on each dimension nor to compare the models quantitatively, but rather to confirm the main effects previously found in the field literature and to show the main trends found in each model with and without the socio-political variables.

### *Selection of CIS studies and variables*

The data and studies were selected according to the availability and comparability of relevant variables. The binary **dependent variables** correspond to each of the dimensions included in the coverage scale:

Table 4. *Dependent variables according to coverage dimensions*

|                    |   |
|--------------------|---|
| Dimension 0.       | - Internet Access: <i>Broadband connection in the household</i> |
| Dimension 1.       | - Basic use: <i>Use of the Internet in the last 12 months</i>   |
| Dimension 2.       | - Daily use: <i>Use of the Internet everyday</i>                |
| Dimension 3. 1. 1. | - Type of use: <i>Online bank management</i>                    |
| Dimension 3. 1. 2. | - Type of use: <i>Other management activities</i>               |
| Dimension 3. 2.    | - Type of use: <i>Search for information</i>                    |
| Dimension 3. 3. 1. | - Type of use: <i>E-mail</i>                                    |
| Dimension 3. 3. 2. | - Type of use: <i>Social networks, chats, etc.</i>              |
| Dimension 3. 4.    | - Type of use: <i>Audiovisual consumption</i>                   |
| Dimension 3. 5.    | - Type of use: <i>Providing E-mail</i>                          |

The **independent variables** were selected according to two groups of variables. The first group refers to the set of *socio-demographic* variables commonly accepted as coverage determinants. The second group of independent variables is composed of variables related to *socio-political attitudes* that are also often included in the CIS studies. This set of variables was included in order to learn more in detail the differences between the haves and the have-nots, going beyond socio-demographic characteristics. This set of variables can be of great help for political or opinion studies as they can offer relevant information about bias in relation to central attitudinal variables. They can also be used for post-stratification techniques, for the exploration of potential bias of studies using, for example, data from social networks to predict election results or any other politically oriented analyses conducted using data generated as a byproduct of any of the types of uses of the coverage dimensions.

## 4 Results (see Table in Appendix)

### *New features of coverage bias by socio-demographics*

In general and despite the growing penetration of Internet usage and access, results show that coverage still represents an important methodological issue, although there are interesting aspects that should be taken into account when studying web surveys and web data coverage bias. Overall findings confirm the main trends drawn in previous studies on the digital divide for socio-demographic characteristics. There are still some social groups that are less likely to have access, to use Internet everyday/in the last 12 months or to make some specific uses of the Internet. Women, elderly, less educated, lower social class, rural area and religious tend to show a negative effect in relation to most of the coverage dimensions included in the scale. However, this does not mean -following Couper's definition- that these groups cannot be reached via web surveys (or that they do not generate web-based data) as some of them report an equal probability of occurrence for specific uses, and in some cases, as occurs with women, they do not report less probability of having Internet access than men. In this sense, there are new features revealed by the analysis that should be taken into consideration when studying coverage bias for web surveys.

Sex and age report no significant effect on the likelihood of having broadband Internet access at home. However, differences between both men and women and younger and older groups remain significant in most models that refer to type of usage. This should alert researchers when considering individuals with access as coverage in web surveys -those reachable via the Internet- as they are likely to find cases in which individuals living in households with Internet access, i.e. housewives and the elderly, actually have very little chance to use the Internet, and thus could be underrepresented in the sample with respect to parameters of the general population.

For management uses of the web, groups aged between 25 to 44 seem to have a higher probability of doing banking and other online administrative arrangements than the age group of 18 to 24 year olds. Results also report that there is no digital divide by gender in the use of E-mail and social networks, chats, etc., a fact that can be relevant for the study of coverage bias in two ways: first, because social networks can play an important role as places in the contacting and recruitment stages of the survey process, and second, that data generated in social networks is not biased in terms of gender.

In contrast, occupation (other than student), social status (other than upper classes) and education (having lower levels of education) tend to report a less significant effect as determinants of specific uses in comparison with access. The divide among social classes is thus lower in relation to the uses individuals make of the Internet than having broadband access to the web at home. This suggests that there may be individuals from upper classes who have access but do not use the web, or, groups from lower social strata that do not have broadband in the household but use the Internet in other places.

### ***Specific coverage bias by socio-political and attitudinal variables***

Both Internet users and holders have a stronger and more active political identity than those who do not use/have Internet. Individuals from both categories tend to put themselves closer to the poles of the ideological self-placement scale –extreme left and extreme right- and less in centered positions of the continuum –center, center left and center right- which may indicate that a strong ideological polarization exist among Internet users/holders. Following this trend, we can appreciate in the analysis that having Internet access or being a user of any of the dimensions included in the coverage scale is somehow linked with a more critical view towards the political establishment and mainstream political parties. In relation to expectancies, Internet users/holders are in general more optimistic about the future of the national economy than those who do not use/have access to the web. Another interesting feature regarding political orientation is that, according to the analysis, there is a great diversity of political views among those who use the social networks, chats, etc. This may have implications in the implementation of new strategies of sampling/recruitment through social networks (Twitter, Facebook, Google +, etc.) and predictions made using non-reactive data from social networks as the results show that bias by ideological view is not relevant.

As regards religiosity, the analysis reveals that believing in God has a negative effect over all dimensions of coverage included in the scale, even after controlling for socio-demographics. Thus, when approaching a web survey or analyzing digital data aimed at the general population, researchers should take into account that Internet users/holders might be, in general, less religious than the rest of the population.

Another common feature observed throughout the analysis in all dimensions and models is that there is a strong negative effect between *don't know* and *don't answer* categories and the uses/access to Internet. This might entail methodological implications related not only to coverage bias, but in a broader sense also to social desirability bias and nonresponse, as Internet users/holders seem more eager to reveal political and religious preferences and, in general terms, to express an opinion and to answer.

***Bias by providing e-mail: Who are the ones who want to take a web survey?***

In light of the results for this dimension, we find that those who provide an e-mail address to a face to face interviewer, and therefore accept to participate in a future web survey, do not show significant bias for variables like gender, vote in last elections, ideology and occupation –except for housewives. The bias in this case does not follow the same trend as for other definitions in some variables. For instance, the size of the population is not a determinant of the likelihood of providing an email. As expected, those who reply *don't know/don't answer* or place themselves in the center of the ideological self-placement scale are less likely to provide an email address than those who admit to being right or left wing.





## 5 Conclusions and further research

The TSE paradigm that has been evolving over decades can still be used today as a theoretical framework to assess the quality of both census/survey data and other types of data. This paper is an example of how traditional theories on survey research can be adapted and used to explore and understand the problems that may emerge from web-based data. We have used data from a traditional survey and the TSE to show the type of bias that may be generated from web surveys and non-reactive web data. There is a need to continue exploring the “digital data tsunami” using these tools, especially if we are to use data generated in the Internet for scientific purposes in the name of the public interest of the whole population. There is still a long way to go and deeper explorations of Internet coverage and its dimensions are needed. In this paper, we have been able to conclude the following.

Coverage bias by socio-demographics still represents a key issue when using web-based data to infer about the general population in a country like Spain. At the same time, it may be advisable to reconsider the notion of coverage of web-based data, not only as Internet penetration in terms of households with Internet access, but also as individuals that actually use the Internet regardless of whether it takes place at home or in places other than households (libraries, cafés, public Wi-Fi, a mobile device etc.). Using the same logic, it is also important to consider that people who report having web access may not have the skills to answer a survey or to generate data as a by-product of Internet activity since they do not actually use the Internet. Both of the above scenarios can lead to an underrepresentation of certain social groups in Internet-generated data, such as housewives and the elderly, within the coverage operationalization.

In this regard, future research is needed to clarify the relationship between household *composition-accesses* and the possible bias for socio-demographics. Bringing further knowledge to the phenomena of the digital divide, could help web data experts and web surveyors when implementing potential solutions to the problem of coverage bias and representativeness. For instance, if the solution chosen for a web survey is to provide Internet access and equipment to those households without them, the bias may not necessarily be resolved as the access to the survey would be still restricted to those with enough computer skills within the households.

Other direct methodological consequences can be extracted from the findings of this study. Regarding weighting, a better understanding of the determinants of each dimension of coverage provides a better way to apply weighting techniques and to determine what the key variables are. In relation to social desirability

and nonresponse, results sound promising as it appears that Internet users are more eager to provide more substantive answers than offline populations who in principle are more likely to answer *don't know/don't answer* and to choose socially expected response options. It may also be interesting to focus further research on whether this is a shared characteristic among most Internet users and how this may be affecting the social desirability bias in other modes of data collection such as personal, mail and phone surveys.

Another important point is the diversity in terms of political views and ideology found in social networks and email user populations. This is of special interest for political and opinion research as a way to explore new possibilities. First, to develop survey recruitment strategies through email lists already gathered or by approaching social media users. Second, to foster opinion research focusing on non-reactive data collected from social networks.

As Internet penetration and usage skills are expected to grow rapidly, it will be necessary to carry out further research in areas such as mobile research and new devices used to connect to the web, or on topics such as places of connection, or more generally, new habits and trends in web use. As mobile penetration is growing even faster than landline penetration, especially in countries with medium or lower rates of landline access, the potential for mixing landline-mobile web surveys in terms of questionnaire design or nonresponse behavior could also be an interesting line of study.

Finally, the need to contemplate the different methods of collecting and analyzing data through the web is now fundamental in online research. Web surveys and non-reactive data seem to be compatible tools for research that can gather a great deal of data to be used through mutually profitable methodological studies. As long as Internet access and use remain unequally distributed among general audiences in most countries, efforts should focus on studying and disentangling the causes of such disparities and consider them when inferring results to the general population. The above findings should also be taken into consideration when setting definitions of Internet penetration, as the current scenario is far more complex than a mere a division of the population between haves and have-nots in terms of access. Researchers should activate a sort of methodological vigilance to be aware at every moment of socio-demographic and socio-political differences between populations that produce webdata –in either web surveys or non-reactive data- and the rest of passive individuals.

# 6 Appendix . Odds ratio for each logit model (\*p<.05;\*\*p<.01;\*\*\*p<.001).

|                     | D0 – Internet Access | D1.1 – Basic use | D1.2 – Daily use | D2.1.1 – Banking management | D2.1.2 Other management | D2.2 – Search for information | D2.3. 1 – Use e-mail | D2.3.2. – Use of Social networks | D2.4 – Audiovisual Consumption | D2.5 – Providing e-mail |
|---------------------|----------------------|------------------|------------------|-----------------------------|-------------------------|-------------------------------|----------------------|----------------------------------|--------------------------------|-------------------------|
| <b>Region</b>       |                      |                  |                  |                             |                         |                               |                      |                                  |                                |                         |
| <i>Madrid</i>       |                      |                  |                  |                             |                         |                               |                      |                                  |                                |                         |
| Aragón              | .669                 | .583*            | .586**           | .697                        | 1.233                   | 1.025                         | .696                 | .690                             | 1.135                          | .566                    |
| Asturias            | .405***              | .547**           | .661*            | .857                        | .993                    | 1.350                         | .710                 | .784                             | 1.135                          | .259**                  |
| Baleares            | .788                 | .700             | .902             | .856                        | 1.438                   | .787                          | .620                 | .587                             | 1.468                          | 1.886*                  |
| Canarias            | .645*                | .612*            | .609**           | 1.409                       | 1.069                   | .947                          | .864                 | 1.038                            | .971                           | 1.750*                  |
| Cantabria           | .516*                | .554*            | .511**           | .213*                       | .254                    | .741                          | .754                 | .874                             | .441                           | .715                    |
| C. Mancha           | .662*                | .494***          | .690*            | .568                        | 1.408                   | .752                          | .649                 | .626                             | 1.092                          | .742                    |
| C. León             | .410***              | .473***          | .563***          | .791                        | .791                    | .950                          | .936                 | .930                             | .851                           | .599*                   |
| Cataluña            | .751*                | .766             | .831             | .956                        | .842                    | .953                          | .956                 | .933                             | .888                           | 1.168                   |
| Valencia            | .432***              | .482***          | .743**           | 1.100                       | .946                    | .702                          | .630**               | .757                             | .823                           | .903                    |
| Extremadura         | .366***              | .474**           | .486***          | .594                        | .859                    | .722                          | .316**               | .411*                            | .738                           | .576                    |
| Galicia             | .404***              | .460***          | .465***          | .555                        | .602                    | .817                          | .628*                | .590*                            | .837                           | .673                    |
| Andalucía           | .590***              | .495***          | .577***          | .610*                       | .893                    | .697*                         | .653*                | .717                             | .879                           | .736                    |
| Murcia              | .436***              | .333***          | .458***          | .771                        | 1.195                   | .808                          | .645                 | .800                             | .797                           | .790                    |
| Navarra             | 1.081                | 1.267            | .665             | .609                        | .486                    | .601                          | .487                 | .337*                            | .687                           | 1.317                   |
| Pais Vasco          | .491**               | .428***          | .488***          | 1.153                       | .582                    | .925                          | 1.064                | .556*                            | 1.462                          | .578*                   |
| La Rioja            | .617                 | .589             | .554             | .728                        | 1.029                   | .620                          | .435                 | .118*                            | .500                           | .721                    |
| <b>Habitat size</b> |                      |                  |                  |                             |                         |                               |                      |                                  |                                |                         |
| <b>Up to 2000</b>   |                      |                  |                  |                             |                         |                               |                      |                                  |                                |                         |
| 2001-10000          | 1.209                | 1.224            | 1.199            | 2.544*                      | 1.189                   | 1.346                         | 1.240                | 1.219                            | .963                           | .759                    |
| 10001-50000         | 1.724**              | 1.389*           | 1.383**          | 1.988                       | 1.409                   | 1.482*                        | 1.234                | 1.420                            | 1.160                          | .806                    |
| 50001-100000        | 2.320***             | 1.263            | 1.294            | 2.575*                      | 1.700                   | 1.487                         | 1.160                | 1.232                            | 1.180                          | 1.045                   |
| 100001-400000       | 2.014***             | 1.683***         | 1.613***         | 1.957                       | 1.247                   | 1.258                         | 1.072                | 1.312                            | .922                           | .712                    |
| 400001-1000000      | 2.440***             | 1.811**          | 1.469*           | 2.701*                      | 2.019                   | 1.562                         | 1.547                | 1.276                            | 1.182                          | .718                    |
| More than 1000000   | 1.319                | 1.262            | 1.510**          | 2.274*                      | 1.439                   | 1.282                         | 1.093                | 1.225                            | .999                           | .569*                   |
| <b>Sex</b>          |                      |                  |                  |                             |                         |                               |                      |                                  |                                |                         |
| <b>Female</b>       |                      |                  |                  |                             |                         |                               |                      |                                  |                                |                         |
| Male                | .997                 | 1.503***         | 1.592***         | 1.434**                     | 1.228                   | 1.222*                        | 1.191*               | 1.183                            | 1.440**                        | 1.166                   |
| <b>Age</b>          |                      |                  |                  |                             |                         |                               |                      |                                  |                                |                         |
| <b>18 to 24</b>     |                      |                  |                  |                             |                         |                               |                      |                                  |                                |                         |
| 25 to 34            | .578**               | .278***          | .627***          | 1.956**                     | 1.742*                  | .746*                         | .769                 | .576***                          | .565***                        | .669**                  |
| 35 to 44            | .649**               | .125***          | .370***          | 2.076**                     | 1.872**                 | .691*                         | .739*                | .263***                          | .377***                        | .434***                 |
| 45 to 54            | .937                 | .066***          | .242***          | 1.288                       | 1.596                   | .529***                       | .497***              | .172***                          | .157***                        | .305***                 |
| 55 to 64            | .543***              | .025***          | .152***          | .929                        | .923                    | .315***                       | .365***              | .121***                          | .110***                        | .197***                 |
| Over 65             | .191***              | .009***          | .054***          | .546                        | .426                    | .117***                       | .133***              | .070***                          | .021***                        | .073***                 |
| <b>Education</b>    |                      |                  |                  |                             |                         |                               |                      |                                  |                                |                         |
| <b>No studies</b>   |                      |                  |                  |                             |                         |                               |                      |                                  |                                |                         |
| Primary school      | 3.028***             | 4.315***         | 4.721***         | 1.340                       | 3.974                   | 8.690**                       | 12.143*              | 3.511*                           | 3.646                          | 2.260***                |
| High school         | 5.561***             | 12.984***        | 12.030***        | 4.169                       | 12.383*                 | 17.316***                     | 27.072**             | 5.801*                           | 6.766                          | 3.776***                |

|                              |          |           |           |         |          |           |          |         |         |         |
|------------------------------|----------|-----------|-----------|---------|----------|-----------|----------|---------|---------|---------|
| Bachelor's degree            | 6.953*** | 33.412*** | 22.455*** | 5.837*  | 19.655** | 20.552*** | 34.574** | 7.503** | 7.384   | -       |
| Don't answer                 | 5.351*   | 6.156*    | 4.344**   | 4.201   | 9.280    | 14.197**  | 16.383*  | 3.614   | 9.805   | 7.007** |
| <u>Occupation</u>            | -        | -         | -         | -       | -        | -         | -        | -       | -       | -       |
| Student                      | .572*    | .595      | .518***   | 2.364*  | .870     | 1.000     | .780     | .692    | .648    | .879    |
| Employee                     | .668     | .508      | .486***   | 1.901   | .720     | .970      | .665     | 1.156   | .851    | 1.306   |
| Small business owner         | .392***  | .335*     | .466***   | 2.020   | .802     | .91       | .777     | .828    | .892    | .886    |
| Unemployed                   | .483**   | .261**    | .255***   | 1.389   | .569     | .539*     | .294***  | .286**  | .340**  | .359**  |
| Domestic or non-paid         | .359*    | .187**    | .142***   | 1.389   | .569     | .539*     | .294***  | .286**  | .340**  | .359**  |
| Farmer                       | .393***  | .299**    | .314***   | 1.225   | .771     | .737      | .497*    | .620    | .762    | .355    |
| Retired/pensioner            | .454     | .608      | .564      | 4.096*  | 1.135    | 1.112     | 1.218    | 1.382   | 1.055   | .680    |
| Other                        | -        | -         | -         | -       | -        | -         | -        | -       | -       | -       |
| <u>Socio-economic status</u> | -        | -         | -         | -       | -        | -         | -        | -       | -       | -       |
| Middle-high                  | .652**   | .732*     | .737***   | 1.019   | .958     | .941      | .959     | .987    | 1.053   | 1.022   |
| New middle                   | .582***  | .423***   | .554***   | .785    | .669     | .650*     | .641**   | .509**  | .640    | .612**  |
| Old middle                   | .492***  | .359***   | .405***   | .441*** | .576**   | .643**    | .584***  | .598**  | .741    | .648**  |
| Qualified workers            | .411***  | .271***   | .380***   | .391*** | .596*    | .612**    | .538***  | .619*   | .585*   | .572**  |
| Unqualified workers          | -        | -         | -         | -       | -        | -         | -        | -       | -       | -       |
| <u>Ideology</u>              | -        | -         | -         | -       | -        | -         | -        | -       | -       | -       |
| Left                         | .895     | .935      | .953      | 1.117   | .842     | 1.049     | 1.060    | .815    | .804    | .850    |
| Center                       | .808     | .792*     | .850      | 1.167   | 1.015    | 1.282     | 1.347    | .853    | 1.312   | 1.122   |
| Right                        | .675**   | .603***   | .696***   | .601    | .793     | .959      | 1.049    | .892    | .937    | .717*   |
| Don't know                   | .687**   | .725*     | .756**    | .920    | .579*    | .763      | .764     | .828    | .852    | .522*** |
| Don't answer                 | -        | -         | -         | -       | -        | -         | -        | -       | -       | -       |
| <u>Religiosity</u>           | -        | -         | -         | -       | -        | -         | -        | -       | -       | -       |
| Believer                     | 1.356**  | 1.705***  | 1.494***  | 1.422** | 1.535**  | 1.200     | 1.209    | 1.145   | 1.205   | 1.31**  |
| Non-believer                 | 1.170    | 1.752*    | 1.569**   | 1.064   | 1.611    | 1.137     | 1.271    | .859    | 1.341   | .706    |
| Don't answer                 | -        | -         | -         | -       | -        | -         | -        | -       | -       | -       |
| <u>Opinion economy</u>       | -        | -         | -         | -       | -        | -         | -        | -       | -       | -       |
| Positive                     | -        | -         | -         | -       | -        | -         | -        | -       | -       | -       |
| Fair                         | -        | 1.248     | 1.063     | 1.906   | .895     | 1.264     | .981     | .850    | .851    | -       |
| Negative                     | -        | 1.228     | .984      | 1.142   | .782     | 1.089     | 583      | .882    | .942    | -       |
| Very negative                | -        | 1.110     | .965      | 1.092   | .666     | 1.019     | .824     | .849    | .702    | -       |
| Don't know                   | -        | .370      | 1.098     | 1.561   | .932     | 2.572     | 1.732    | .753    | 1.232   | -       |
| Don't answer                 | -        | .189      | .462      | -       | -        | -         | -        | -       | -       | -       |
| <u>Opinion politics</u>      | -        | -         | -         | -       | -        | -         | -        | -       | -       | -       |
| Positive                     | -        | -         | -         | -       | -        | -         | -        | -       | -       | -       |
| Fair                         | -        | .851      | 1.067     | .973    | 1.038    | .914      | .808     | .675*   | .856    | -       |
| Negative                     | -        | .981      | 1.177     | 1.427   | 1.481    | 1.198     | 1.069    | .869    | .987    | -       |
| Very negative                | -        | 1.091     | 1.283*    | 2.418** | 1.851*   | 1.593*    | 1.402    | .866    | 1.173   | -       |
| Don't know                   | -        | .466**    | 1.165     | .300    | .577     | .537*     | .918     | .842    | .884    | -       |
| Don't answer                 | -        | .261*     | .589      | 3.139   | 4.575    | .625      | 1.663    | .312    | .626    | -       |
| <u>Expectancies economy</u>  | -        | -         | -         | -       | -        | -         | -        | -       | -       | -       |
| Better                       | -        | .831*     | .854*     | .680**  | .679**   | .694***   | .637***  | .734*   | .746*   | -       |
| Same                         | -        | .825*     | .825*     | .416*** | .554***  | .394***   | .367***  | .444*** | .496*** | -       |
| Worse                        | -        | 1.022     | .754*     | .660    | .594     | .606*     | .562***  | .538*   | .469*   | -       |
| Don't know                   | -        | 1.827     | 1.886     | 2.934   | -        | 1.030     | 2.169    | -       | -       | -       |
| Don't answer                 | -        | -         | -         | -       | -        | -         | -        | -       | -       | -       |
| <u>Expectancies politics</u> | -        | -         | -         | -       | -        | -         | -        | -       | -       | -       |
| Better                       | -        | -         | -         | -       | -        | -         | -        | -       | -       | -       |

|                                 |                |               |               |                |                |               |               |                |
|---------------------------------|----------------|---------------|---------------|----------------|----------------|---------------|---------------|----------------|
| Same                            | 1.009          | 1.324         | 1.041         | 1.168          | 1.337*         | 1.264         | 1.259         | -              |
| Worse                           | 1.190          | .981          | .846          | 1.064          | 1.266          | 1.400         | 1.374         | -              |
| Don't know                      | .961           | 1.401         | 1.072         | 1.433          | 1.163          | 1.191         | .995          | -              |
| Don't answer                    | 1.788          | .298          | .377          | 1.160          | .299           | 1.761-        | .732          | -              |
| Main problem facing the country | -              | -             | -             | -              | -              | -             | -             | -              |
| <b>Unemployment</b>             | -              | -             | -             | -              | -              | -             | -             | -              |
| Economy                         | .988           | .813          | 1.030         | .858           | .840           | .877          | .882          | -              |
| Politicians, government         | 1.333*         | 1.074         | 1.261         | 1.052          | 1.083          | 1.203         | 1.233         | -              |
| Immigration                     | .689           | 1.253         | 1.006         | 1.006          | .990           | 1.286         | .859          | -              |
| Other                           | 1.011          | 1.021         | 1.207         | .934           | .975           | 1.074         | 1.036         | -              |
| Don't know                      | .649           | .283          | .888          | .804           | 1.496          | 1.570         | 1.336         | -              |
| Don't answer                    | .413           | -             | -             | 1.874          | .394           | 1.464         | .993          | -              |
| Main problem affecting you      | -              | -             | -             | -              | -              | -             | -             | -              |
| <b>Unemployment</b>             | -              | -             | -             | -              | -              | -             | -             | -              |
| Economy                         | .925           | 1.029         | 1.083         | 1.054          | 1.039          | .919          | 1.204         | -              |
| Politicians, government         | .939           | 1.210         | 1.254         | 1.042          | 1.087          | .953          | 1.176         | -              |
| Immigration                     | 1.257          | 1.096         | 1.339         | 1.277          | 1.674          | .920          | 1.102         | -              |
| Other                           | .896           | .780          | .742          | .907           | .900           | .737*         | .915          | -              |
| Don't know                      | .849           | .841          | .476*         | .810           | .787           | .646          | .549*         | -              |
| Don't answer                    | .870           | 1.039         | .770          | .800           | .969           | .812          | .824          | -              |
| Vote in last elections          | -              | -             | -             | -              | -              | -             | -             | -              |
| <b>Yes</b>                      | -              | -             | -             | -              | -              | -             | -             | -              |
| No                              | .865           | 1.180         | .958          | 1.486**        | 1.542**        | 1.725***      | 1.730***      | .972           |
| Don't remember                  | 1.662          | 1.590         | .345          | 1.871          | 1.647          | .684          | 2.900         | .428           |
| Don't know                      | 1.114          | .848          | 2.434*        | 2.131*         | 2.079*         | 1.272         | 1.483         | .476*          |
| Party voted                     | -              | -             | -             | -              | -              | -             | -             | -              |
| <b>PSOE</b>                     | -              | -             | -             | -              | -              | -             | -             | -              |
| PP                              | 1.005          | .871          | 1.122         | 1.130          | 1.197          | 1.207         | .920          | .918           |
| IU                              | .879           | 1.175         | .876          | .819           | 1.003          | .778          | 1.210         | 1.568*         |
| CIU                             | 1.733*         | 1.141         | 1.178         | .767           | .690           | 1.177         | .958          | 1.628          |
| PNV                             | 1.085          | .757          | .917          | .751           | .993           | 1.056         | .424          | .584           |
| UPyD                            | 1.503          | 1.280         | 1.960         | 2.150          | 1.965          | 2.212         | 2.442*        | 1.775          |
| ERC                             | 3.987**        | 1.169         | .963          | .905           | .798           | 1.317         | .903          | 1.319          |
| BNG                             | 4.906**        | 2.937         | .962          | 1.535          | 2.788          | 3.317         | 1.052         | 2.186          |
| CC                              | .974           | -             | -             | .357           | .593           | 1.183         | -             | 1.042          |
| NA BAI                          | .280*          | 2.322         | 5.485*        | 2.360          | 2.037          | .972          | 1.548         | .501           |
| OTHER                           | 1.502          | 1.288         | 1.394         | 1.763*         | 1.926*         | 2.472**       | 1.248         | 1.524          |
| Blank Ballot                    | 1.092          | 1.327         | 1.665         | 1.844*         | 2.029*         | 2.115*        | 1.135         | .797           |
| Did not vote                    | -              | -             | -             | -              | -              | -             | -             | -              |
| Don't remember                  | 1.076          | 2.285*        | 1.230         | 1.376          | 2.284**        | 2.501**       | 1.657         | .492           |
| Don't answer                    | 1.216          | .992          | 1.231         | 1.250          | 1.386*         | 1.002         | .910          | .673*          |
| <b>Number of obs</b>            | <b>9156</b>    | <b>4755</b>   | <b>4749</b>   | <b>4815</b>    | <b>4815</b>    | <b>4808</b>   | <b>4758</b>   | <b>4806</b>    |
| <b>LR chi2(65)</b>              | <b>6396.50</b> | <b>720.98</b> | <b>647.31</b> | <b>1093.40</b> | <b>1140.45</b> | <b>856.20</b> | <b>772.56</b> | <b>1055.51</b> |
| <b>Prob &gt; chi2</b>           | <b>0.0000</b>  | <b>0.0000</b> | <b>0.0000</b> | <b>0.0000</b>  | <b>0.0000</b>  | <b>0.0000</b> | <b>0.0000</b> | <b>0.0000</b>  |
| <b>Pseudo R2</b>                | <b>0.4888</b>  | <b>0.2364</b> | <b>0.2081</b> | <b>0.2001</b>  | <b>0.2250</b>  | <b>0.2211</b> | <b>0.2236</b> | <b>0.2234</b>  |



# References

- Alvira Martín, F. (2004): *La encuesta: una perspectiva general metodológica*, Madrid: CIS.
- Bandilla, W., Bosnjak, M., & Altdorfer, P. (2003): Survey administration effects? A comparison of web-based and traditional written self-administered surveys using the ISSP environment module. *Social Science Computer Review*, 21, 235-243.
- Berzelak, J., Lozar Manfreda, K., Vehovar, V. (2006): *Software Tools for Web Surveys: The More You Pay, the More You Get*. Presentation at Applied Statistics.
- Bethlehem, J. (2010): Selection Bias in Web Surveys. *International Statistical Review*, 78, 2, 161–188
- Bethlehem, J. and Biffignandi, S. (2012). *Handbook of Web Surveys*. Hoboken, NJ: Wiley.
- Blair E, Blair J. (2006): Dual Frame Web-Telephone Sampling for Rare Groups. *Journal of Official Statistics*, 22, 2, pp. 211-220.
- COST Action IS1004 Webdatanet (2013): [www.webdatanet.eu](http://www.webdatanet.eu)
- CIS Centro de Investigaciones Sociológicas database (2012): [http://www.cis.es/cis/opencms/EN/8\\_cis/index.html](http://www.cis.es/cis/opencms/EN/8_cis/index.html)
- Couper, M. (2000): Web surveys: A review of issues and approaches. *Public Opinion Quarterly*, Vol. 64(4), pp. 464-481.
- Couper, M. (2011): The future modes of data collection. *Public Opinion Quarterly*, Vol. 75, No. 5, 2011, pp. 889–908.
- Couper, M. (2013). Is the Sky Falling? New Technology, Changing Media, and the Future of Surveys. *Survey Research Methods* 7 (3) pp. 145-156
- De Leeuw (2011): Online surveys: Past, present and future. Presentation at the 2nd Webdatanet MC & WGs meeting, Amsterdam 30 November - 2 December 2011. Available at <http://webdatanet.cbs.dk/images/presentations/cost2011%20websurveys%20pdf.pdf>
- De Leeuw, E.D. Hox, J.J (2011): Internet surveys as part of a mixed-mode design. Pp. 45-76 en M. Das, P. Ester & L. Kaczmirek (Eds.). *Social and Behavioral Research and the Internet*. NY: Routledge.
- Deming, W.E. (1944): On Errors in Surveys. *American Sociological Review* 9:359-369.
- Díaz de Rada, Vidal. (2010): Comparación entre los resultados proporcionados por encuestas telefónicas y personales: el caso de un estudio electoral. Colección “Opiniones y actitudes”. Núm. 66. CIS. Madrid.
- Dillman, Don A., Jolene D. Smyth, and Leah Melani Christian. (2009): *Internet, Mail, and Mixed mode Surveys: The Tailored Design Method*. 3rd Edition, Hoboken, Wiley. Nueva York.
- DiMaggio, P., Hargittai E., Celeste, C. & Shafer, S. (2004): From Unequal Access to Differentiated Use: A Literature Review and Agenda for Research on Digital Inequality. *Encuesta General de Medios* (2012): [www.egm.aimc.es/aimc/](http://www.egm.aimc.es/aimc/)
- Eurobarometer (2012): [http://ec.europa.eu/public\\_opinion/cf/index\\_en.cfm](http://ec.europa.eu/public_opinion/cf/index_en.cfm)
- Görzitz, A. (2006): Incentives in Web Studies: Methodological Issues and a Review. *International Journal of Internet Science*, 2006, 1 (1), 58-70
- Groves, R.; Fowler, F.; Couper, M.; Lepkowski, J.; Singer, E.; Tourangeau, R. (2004): *Survey Methodology*. Wiley. New York.

- Groves, R. M. (2011): Three Eras of Survey Research. *Public Opinion Quarterly*, Vol. 75, No. 5, 2011, pp. 861–871
- Ferneer, H. Sonck, N. Stoop, I (2011): Smartphone Data Collection Among LISS Panel Members. Presentación en MESS Workshop 2011, Amsterdam.
- Fuchs, M. and Busse, B. (2009). The Coverage Bias of Mobile Web Surveys Across European Countries. *International Journal of Internet Science*, 4, 21-33. Available at: [http://www.ijis.net/ijis4\\_1/ijis4\\_1\\_fuchs\\_pre.html](http://www.ijis.net/ijis4_1/ijis4_1_fuchs_pre.html)
- Informe Tatum (2012): [www.tatum.es](http://www.tatum.es)
- Instituto Nacional de Estadística (INE) (2012): <http://www.ine.es/>
- Internet World Stats (2012): <http://www.internetworldstats.com>
- Kaczmirek, L. (2008). Human-Survey Interaction: Usability and Nonresponse in Online Surveys. Mannheim: MADOC (Mannheim Electronic Document Server). Doctoral thesis.
- Kalgraff, K.; Kolsrud, K. (2011): Experiences with mixed mode mail & web-enquêtes in probability samples with known individuals. Presentation at the 2nd Webdatanet MC & WGs meeting, Amsterdam 30 November - 2 December 2011.
- Prensky, M. (2001): Digital natives, digital immigrants. *On the Horizon*, 9(5), 1-6.
- Lavrakas. J. P. (2013): Applying a Total Error Perspective for Improving Research Quality in the Social, Behavioral, and Marketing Sciences. *Public Opinion Quarterly* 77 (3): 831-850.
- Lee S., (2006): Propensity Score Adjustment as a Weighting Scheme for Volunteer Panel Web Surveys. *Journal of Official Statistics*, 22:2, 329-349.
- Lozar Manfreda, K.; Bosnjak, M.; Berzelak, J.; Haas, I.; Vehovar, V.; (2008): Web Surveys versus Other Survey Modes – A Meta-Analysis Comparing Response Rates. *International Journal of Market Research*, 50, 1, pp. 79-104.
- Lozar Manfreda, K.; Berzelak, J.; Haas, I.; Vehovar, V.; (2011). Nonresponse in web surveys. *International Encyclopedia of Statistical Science*, 2011
- Lynn, P. (2013): Issues of Coverage and Sampling in Web Surveys for the General Population: An Overview. Synthesis paper for NCRM Web Survey Network opening conference. Available at <http://www.cls.ioe.ac.uk/shared/get-file.ashx?itemtype=document&id=1595>
- Mohorko, A., De Leeuw, E., Hox, J. (2013): Internet Coverage and Coverage Bias in Europe: Developments Across Countries and Over Time. *Journal of Official Statistics*. Volume 29, Issue 4, Pages 609–622.
- Pedraza, P. de.; Tijdens, K; Muñoz de Bustillo, R.; Steinmetz, S. (2010): A Spanish Continuous Volunteer Web Survey: Sample Bias, Weighting and Efficiency. *Revista Española de Investigaciones Sociológicas (Reis)* N.º 131, 2010, pp. 109-130
- Prewitt, K. (2013) The 2012 Morris Hansen lecture: Thank you Morris, et al., for Westat et al. *Journal of Official Statistics*, 29 (2), 223-231.
- Riccardini, F., Fazio, M. (2002): Measuring the Digital Divide. Paper presented at the IAOS Conference on Official Statistics and the New Economy, 2002. Available at: <http://www.websm.org/db/12/170/rec/>
- Robinson, J. P., Dimaggio, P., Hargittai E. (2003): New Social Survey Perspectives on the Digital Divide. *It & Society*, Volume 1, Issue 5, Summer 2003, Pp. 1-22.
- Sánchez Carrión, J. J., Segovia Guisado, J. M. y Sánchez Meseguer, P., (2012): Las Encuestas En Internet en Arroyo M. y Sádaba I. *Metodología de la investigación social, Síntesis*, Madrid: 79-108.



- Savage, M.; Burrows, R. (2007): The Coming Crisis of Empirical Sociology. *Sociology* vol. 41 no. 5 885-899
- Schonlau, M.; Zapert, K.; Payne Simon, L.; Sanstad, K.; Marcus, S.; Adams, J.; Spranca, M.; Kan, H-J.; Turner, R.; Berry, S. (2004): A comparison between a propensity weighted web survey and an identical RDD survey. *Social Science Computer Review*.
- Shih, T.; Fan, X. (2008) Comparing Response Rates from Web and Mail Surveys: A Meta-Analysis *Field Methods* August 20: 249-271.
- Steinmetz, S.; Tijdens, K.; Pedraza, P. de, (2009): Comparing different weighting procedures for volunteer-web surveys. *AIAS Working Papers* 09/76.
- Toepoel, V., Dillman, D. A. (2011): Words, Numbers, and Visual Heuristics in Web Surveys: Is There a Hierarchy of Importance? *Social Science Computer Review*, 29, 2, pp. 193-207.
- Torres, Cristobal. (2007): Inhibidores de uso de las TIC en la sociedad española. Informe elaborado para Telefónica, Ministerio de Industria, Turismo y Comercio, red.es.
- Vehovar, V.; Batagelj, Z.; Manfreda, K. (1999): Web surveys: can the weighting solve the problem? *Proceedings of the Survey Research Methods Section, American Statistical Association* (1999), pp. 962-967. Available at [http://www.amstat.org/sections/SRMS/proceedings/papers/1999\\_168.pdf](http://www.amstat.org/sections/SRMS/proceedings/papers/1999_168.pdf)
- Webdatanet Memorandum of Understanding (2011). Available at: <http://webdatanet.cbs.dk/images/presentations/mou%20is1004-webdatanet.pdf>
- WebSM (2013): <http://www.websm.org/>
- Wilson, K. R., Wallin J. S., Reiser C. (2003): Social Stratification and the Digital Divide. *Social Science Computer Review*, Vol. 21 No. 2, 133-143.
- Zhang, C., Callegaro M., Thomas, M. (2009) More than the Digital Divide? Investigating the Differences between Internet and Non-Internet Users on Attitudes and Behaviors. Prepared for the Midwest Association for Public Opinion Research (MAPOR) 2008 Conference. 2009 version. Available at: <http://www.knowledgenetworks.com/ganp/docs/Digital%20Divide%20-%20MAPOR%20abstract%20-%20Zhang-Callegaro-Thomas%20-%202012-08.pdf>



## AIAS Working Papers (€ 7,50)

Recent publications of the Amsterdam Institute for Advanced Labour Studies. They can be downloaded from our website [www.uva-aias.net](http://www.uva-aias.net) under the subject Publications.

- 153 [Mobile CATI: An Overview of Issues](#)  
November 2014 - Ana Slavec and Daniele Toninelli
- 152 [Wat levert keuzevrijheid in het pensioenstelsel op?](#)  
November 2014 - Paul de Beer, Johan De Deken, David Hollanders, Sijbren Kuiper, Wiemer Salverda, Natascha van der Zwan
- 151 [Measurement error in occupational coding: an analysis on SHARE data](#)  
November 2014 - Michele Belloni, Agar Brugiavini, Elena Meschi, Kea Tjeldens
- 150 [Who has access to mobile devices in an online commercial panel? An analysis of potential respondents for mobile surveys](#)  
November 2014 - Melanie Revilla, Daniele Toninelli, Carlos Ochoa, Germán Loewe
- 149 [Reviewing the measurement and comparison of occupations across Europe](#)  
August 2014 - Kea Tjeldens
- 148 [‘Werkgeverskeuze en Pensioen: Een Institutionele Analyse](#)  
September 2014 - Natascha van der Zwan
- 147 [Why Dutch women work part-time: A Oaxaca-decomposition of differences in European female part-time work rates](#)  
July 2014 - Nick Deschacht and Kea Tjeldens
- 146 [National Labour Rights for Women](#)  
July 2014 - Janna Besamusca and Kea Tjeldens
- 145 [Labour-market institutions and the dispersion of wage earnings](#)  
May 2014 - Wiemer Salverda and Danielle Checchi
- 144 [Does desire for social status promote solidarity? Investigating the role of egalitarian versus inegalitarian societal contexts](#)  
May 2014 - Marii Paskov
- 143 [Big Data and virtual communities: methodological issues](#)  
April 2014 - M<sup>a</sup> Rocio Martínez-Torres, Sergio L. Toral and Nicoletta Fornara
- 142 [Skills and occupational needs: labour market forecasting systems in Italy](#)  
March 2014 - Giovanni Castiglioni and Kea Tjeldens
- 141 [Who should earn what? A Q methodological study on notions of justice of wage differences](#)  
November 2013 - Wout Scholten and Margo Trappenburg
- 140 [Who has an informal job and how is that job paid? A job-based informality index for nine sub-Saharan African countries](#)  
November 2013 - Kea Tjeldens, Janna Besamusca and Maarten van Klaveren
- 139 [Positive Action in EU Gender Equality Law: Promoting More Women in Corporate Decision Making?](#)  
October 2013 - Nuria Elena Ramos Martín
- 138 [Multiple barriers and bridges to work](#)  
October 2013 - Tomáš Sirovátka and Els Sol
- 137 [Governance of EU labour law. EU’s working time directive and it’s implementation in the Netherlands](#)  
September 2013 - Els Sol, Nuria Ramos
- 136 [Benchmark. Towards evidence-based work first](#)  
September 2013 - Els Sol, Julie Castonguay, Hanneke van Lindert, Yvonne van Amstel
- 135 [De bevoegdheden van werkgevers en werknemers om een pensioenuitvoerder te kiezen](#)  
October 2013 - Sijbren Kuiper
- 134 [Economic valuation in Web surveys; A review of the state of the art and best practices](#)  
August 2013 - Angeliki, N. Menegaki, Konstantinos P. Tsagarakis
- 133 [Do Spanish firms support initial vocational training? Company behaviour in low-coordinated institutional frameworks](#)  
August 2013 - David Fernández Guerrero
- 132 [Interactive applets on the Web for methods and statistics](#)  
August 2013 - Ulf-Dietrich Reips, Gary McClelland
- 131 [Can creative web survey questionnaire design improve the response quality?](#)  
July 2013 - Julijana Angelovska, Petroula M. Mavrikiou

- 129 Children, Elder Care and the Probabilities Spanish Women have of Holding Decent Works  
July 2013 - Alberto Villacampa González, Pablo de Pedraza García
- 128 Collectieve Zeggenschap in het Nederlandse Pensioenstelsel: De Beroepspensioenvereniging  
August 2013 - Natascha van der Zwan
- 127 More or less strangers. Social distance as reflected in news media reporting on the young, the old and the allochthon  
December 2012 - Dorota Lepianka
- 126 Development of the public-private wage differential in the Netherlands 1979 – 2009  
December 2012 - Ernest Berkhout and Wiemer Salverda
- 125 Solidarity in a multicultural neighbourhood. Results of a field experiment  
December 2012 - Paul de Beer and Maarten Berg
- 124 Conditions and motives for voluntary sharing. Results of a solidarity game experiment  
December 2012 - Paul de Beer and Maarten Berg
- 123 “Gone Fishing” Modeling Diversity in Work Ethics  
October 2012 - Annette Freyberg-Inan and Rüya Gökhan Koçer
- 122 Skill-based inequality in job quality  
August 2012 - Haya Stier
- 121 Occupational segregation and gender inequality in job quality  
August 2012 - Haya Stier and Meir Yaish
- 120 The impact of attitudes and work preferences on Dutch mothers’ employment patterns  
April 2012 - Justine Ruitenbergh and Paul de Beer
- 119 “He would never just hit the sofa” A narrative of non-complaining among Dutch Mothers.  
A qualitative study of the influences of attitudes on work preferences and employment patterns of Dutch mothers  
April 2012 - Justine Ruitenbergh
- 118 Collective redress and workers’ rights in the EU  
March 2012 - Jan Cremers and Martin Bulla
- 117 Forthcoming: An individual level perspective on the concept of flexicurity  
Antonio Firinu
- 116 Comparative study of labour relations in African countries  
December 2011 - Rüya Gökhan Koçer and Susan Hayter
- 115 More flexibility for more innovation?  
December 2011 - Eva Wachsen and Knut Blind
- 114 De loonkloof tussen mannen en vrouwen. Een review van het onderzoek in Nederland  
December 2011 - Kea G. Tijdens en Maarten van Klaveren
- 113 European social dialogue as multi-level governance. Towards more autonomy and new dependencies  
September 2011 - Paul Marginson and Maarten Keune
- 112 Flexicurity: a new impulse for social dialogue in Europe?  
September 2011 - Maarten Keune
- 11-111 Health workforce remuneration. Comparing wage levels, ranking and dispersion of 16 occupational groups in 20 countries  
August 2011 - Kea Tijdens and Daniel H. de Vries
- 11-110 Over- and underqualification of migrant workers. Evidence from WageIndicator survey data  
July 2011 - Kea Tijdens and Maarten van Klaveren
- 11-109 Employees’ experiences of the impact of the economic crisis in 2009 and 2010  
July 2011 - Kea Tijdens, Maarten van Klaveren, Reinhard Bispinck, Heiner Dribbusch and Fikret Öz
- 11-108 A deeper insight into the ethnic make-up of school cohorts: Diversity and school achievement  
January 2011 - Virginia Maestri
- 11-107 Codebook and explanatory note on the EurOccupations dataset about the job content of 150 occupations  
January 2011 - Kea Tijdens, Esther de Ruijter and Judith de Ruijter
- 10-106 The Future of Employment Relations: Goodbye ‘Flexicurity’ – Welcome Back Transitional Labour Markets?  
2010 - Günther Schmid
- 11-105 Forthcoming:  
This time is different?! The depth of the Financial Crisis and its effects in the Netherlands.  
Wiemer Salverda

- 11-104 Forthcoming:  
Integrate to integrate. Explaining institutional change in the public employment service - the one shop office  
Marieke Beentjes, Jelle Visser and Marloes de Graaf-Zijl
- 11-103 Separate, joint or integrated? Active labour market policy for unemployed on social assistance and unemployment benefits  
2011 - Lucy Kok, Caroline Berden and Marloes de Graaf-Zijl
- 10-102 Codebook and explanatory note on the WageIndicator dataset a worldwide, continuous, multilingual web-survey on work and wages with paper supplements  
2010 - Kea Tjeldens, Sanne van Zijl, Melanie Hughie-Williams, Maarten van Klaveren, Stephanie Steinmetz
- 10-101 Uitkeringsgebruik van Migranten  
2010 - Aslan Zorlu, Joop Hartog and Marieke Beentjes
- 10-100 Low wages in the retail industry in the Netherlands. RSF project Future of work in Europe / Low-wage Employment: Opportunity in the Workplace in Europe and the USA  
2010 - Maarten van Klaveren
- 10-99 Pension fund governance. The intergenerational conflict over risk and contributions  
2010 - David Hollanders
- 10-98 The greying of the median voter. Aging and the politics of the welfare state in OECD countries  
2010 - David Hollanders and Ferry Koster
- 10-97 An overview of women's work and employment in Zimbabwe  
Decisions for Life Country Report  
2010 - Maarten van Klaveren, Kea Tjeldens, Melanie Hughie-Williams and Nuria Ramos
- 10-96 An overview of women's work and employment in Belarus  
Decisions for Life Country Report  
2010 - Maarten van Klaveren, Kea Tjeldens, Melanie Hughie-Williams and Nuria Ramos
- 10-95 Uitzenden in tijden van crisis  
2010 - Marloes de Graaf-Zijl and Emma Folmer
- 10-94 An overview of women's work and employment in Ukraine  
Decisions for Life Country Report  
2010 - Maarten van Klaveren, Kea Tjeldens, Melanie Hughie-Williams and Nuria Ramos
- 10-93 An overview of women's work and employment in Kazakhstan  
Decisions for Life Country Report  
2010 - Maarten van Klaveren, Kea Tjeldens, Melanie Hughie-Williams and Nuria Ramos
- 10-92 An overview of women's work and employment in Azerbaijan  
Decisions for Life Country Report  
2010 - Maarten van Klaveren, Kea Tjeldens, Melanie Hughie-Williams and Nuria Ramos
- 10-91 An overview of women's work and employment in Indonesia  
Decisions for Life Country Report  
2010 - Maarten van Klaveren, Kea Tjeldens, Melanie Hughie-Williams and Nuria Ramos
- 10-90 An overview of women's work and employment in India  
Decisions for Life Country Report  
2010 - Maarten van Klaveren, Kea Tjeldens, Melanie Hughie-Williams and Nuria Ramos
- 10-89 Coordination of national social security in the EU – Rules applicable in multiple cross border situations  
2010 - Jan Cremers
- 10-88 Geïntegreerde dienstverlening in de keten van Werk en Inkomen  
2010 - Marloes de Graaf-Zijl, Marieke Beentjes, Eline van Braak
- 10-87 Emigration and labour shortages. An opportunity for trade unions in new member states?  
2010 - Monika Ewa Kaminska and Marta Kahancová
- 10-86 Measuring occupations in web-surveys. The WISCO database of occupations  
2010 - Kea Tjeldens
- 09-85 Multinationals versus domestic firms: Wages, working hours and industrial relations  
2009 - Kea Tjeldens and Maarten van Klaveren
- 09-84 Working time flexibility components of companies in Europe  
2009 - Heejung Chung and Kea Tjeldens

- 09-83 [An overview of women's work and employment in Brazil](#)  
Decisions for Life Country Report  
2009 - Maarten van Klaveren, Kea Tijdens, Melanie Hughie-Williams and Nuria Ramos
- 09-82 [An overview of women's work and employment in Malawi](#)  
Decisions for Life Country Report  
2009 - Maarten van Klaveren, Kea Tijdens, Melanie Hughie-Williams and Nuria Ramos
- 09-81 [An overview of women's work and employment in Botswana](#)  
Decisions for Life Country Report  
2009 - Maarten van Klaveren, Kea Tijdens, Melanie Hughie-Williams and Nuria Ramos
- 09-80 [An overview of women's work and employment in Zambia](#)  
Decisions for Life Country Report  
2009 - Maarten van Klaveren, Kea Tijdens, Melanie Hughie-Williams and Nuria Ramos
- 09-79 [An overview of women's work and employment in South Africa](#)  
Decisions for Life Country Report  
2009 - Maarten van Klaveren, Kea Tijdens, Melanie Hughie-Williams and Nuria Ramos
- 09-78 [An overview of women's work and employment in Angola](#)  
Decisions for Life Country Report  
2009 - Maarten van Klaveren, Kea Tijdens, Melanie Hughie-Williams and Nuria Ramos
- 09-77 [An overview of women's work and employment in Mozambique](#)  
Decisions for Life Country Report  
2009 - Maarten van Klaveren, Kea Tijdens, Melanie Hughie-Williams and Nuria Ramos
- 09-76 [Comparing different weighting procedures for volunteer web surveys. Lessons to be learned from German and Dutch Wage indicator data](#)  
2009 - Stephanie Steinmetz, Kea Tijdens and Pablo de Pedraza
- 09-75 [Welfare reform in the UK, the Netherlands, and Finland. Change within the limits of path dependence.](#)  
2009 - Minna van Gerven
- 09-74 [Flexibility and security: an asymmetrical relationship? The uncertain relevance of flexicurity policies for segmented labour markets and residual welfare regimes](#)  
2009 - Aliko Mouriki (guest at AIAS from October 2008 - March 2009)
- 09-73 [Education, inequality, and active citizenship tensions in a differentiated schooling system](#)  
2009 - Herman van de Werfhorst
- 09-72 [An analysis of firm support for active labor market policies in Denmark, Germany, and the Netherlands](#)  
2009 - Moira Nelson
- 08-71 [The Dutch minimum wage radical reduction shifts main focus to part-time jobs](#)  
2008 - Wiemer Salverda
- 08-70 [Parallele innovatie als een vorm van beleidsleren: Het voorbeeld van de keten van werk en inkomen](#)  
2008 - Marc van der Meer, Bert Roes
- 08-69 [Balancing roles - bridging the divide between HRM, employee participation and learning in the Dutch knowledge economy](#)  
2008 - Marc van der Meer, Wout Buitelaar
- 08-68 [From policy to practice: Assessing sectoral flexicurity in the Netherlands](#)  
October 2008 - Hesther Houwing / Trudie Schils
- 08-67 [The first part-time economy in the world. Does it work?](#)  
Republication August 2008 - Jelle Visser
- 08-66 [Gender equality in the Netherlands: an example of Europeanisation of social law and policy](#)  
May 2008 - Nuria E.Ramos-Martin
- 07-65 [Activating social policy and the preventive approach for the unemployed in the Netherlands](#)  
January 2008 - Minna van Gerven
- 07-64 [Struggling for a proper job: Recent immigrants in the Netherlands](#)  
January 2008 - Aslan Zorlu
- 07-63 [Marktwerking en arbeidsvoorwaarden – de casus van het openbaar vervoer, de energiebedrijven en de thuiszorg](#)  
July 2007 - Marc van der Meer, Marian Schaapman & Monique Aerts

- 07-62 Vocational education and active citizenship behaviour in cross-national perspective  
November 2007 - Herman G. van der Werfhorst
- 07-61 The state in industrial relations: The politics of the minimum wage in Turkey and the USA  
November 2007 - Ruġa Gökhan Koçer & Jelle Visser
- 07-60 Sample bias, weights and efficiency of weights in a continuous web voluntary survey  
September 2007 - Pablo de Pedraza, Kea Tjeldens & Rafael Muñoz de Bustillo
- 07-59 Globalization and working time: Work-Place hours and flexibility in Germany  
October 2007 - Brian Burgoon & Damian Raess
- 07-58 Determinants of subjective job insecurity in 5 European countries  
August 2007 - Rafael Muñoz de Bustillo & Pablo de Pedraza
- 07-57 Does it matter who takes responsibility?  
May 2007 - Paul de Beer & Trudie Schils
- 07-56 Employment protection in dutch collective labour agreements  
April 2007 - Trudie Schils
- 07-54 Temporary agency work in the Netherlands  
February 2007 - Kea Tjeldens, Maarten van Klaveren, Hester Houwing, Marc van der Meer & Marieke van Essen
- 07-53 Distribution of responsibility for social security and labour market policy  
Country report: Belgium  
January 2007 - Johan de Deken
- 07-52 Distribution of responsibility for social security and labour market policy  
Country report: Germany  
January 2007 - Bernard Ebbinghaus & Werner Eichhorst
- 07-51 Distribution of responsibility for social security and labour market policy  
Country report: Denmark  
January 2007 - Per Kongshøj Madsen
- 07-50 Distribution of responsibility for social security and labour market policy  
Country report: The United Kingdom  
January 2007 - Jochen Clasen
- 07-49 Distribution of responsibility for social security and labour market policy  
Country report: The Netherlands  
January 2007 - Trudie Schils
- 06-48 Population ageing in the Netherlands: demographic and financial arguments for a balanced approach  
January 2007 - Wiemer Salverda
- 06-47 The effects of social and political openness on the welfare state in 18 OECD countries, 1970-2000  
January 2007 - Ferry Koster
- 06-46 Low pay incidence and mobility in the Netherlands - Exploring the role of personal, job and employer characteristics  
October 2006 - Maite Blázquez Cuesta & Wiemer Salverda
- 06-45 Diversity in work: The heterogeneity of women's labour market participation patterns  
September 2006 - Mara Yerkes
- 06-44 Early retirement patterns in Germany, the Netherlands and the United Kingdom  
October 2006 - Trudie Schils
- 06-43 Women's working preferences in the Netherlands, Germany and the UK  
August 2006 - Mara Yerkes
- 05-42 Wage bargaining institutions in Europe: a happy marriage or preparing for divorce?  
December 2005 - Jelle Visser
- 05-41 The work-family balance on the union's agenda  
December 2005 - Kilian Schreuder
- 05-40 Boxing and dancing: Dutch trade union and works council experiences revisited  
November 2005 - Maarten van Klaveren & Wim Sprenger
- 05-39 Analysing employment practices in western european multinationals: coordination, industrial relations and employment flexibility in Poland  
October 2005 - Marta Kahancova & Marc van der Meer

- 05-38 [Income distribution in the Netherlands in the 20<sup>th</sup> century: long-run developments and cyclical properties](#)  
September 2005 - Emiel Afman
- 05-37 [Search, mismatch and unemployment](#)  
July 2005 - Maite Blazques & Marcel Jansen
- 05-36 [Women's preferences or delineated policies? The development of part-time work in the Netherlands, Germany and the United Kingdom](#)  
July 2005 - Mara Yerkes & Jelle Visser
- 05-35 [Vissen in een vreemde vijver: Het werven van verpleegkundigen en verzorgenden in het buitenland](#)  
May 2005 - Judith Roosblad
- 05-34 [Female part-time employment in the Netherlands and Spain: an analysis of the reasons for taking a part-time job and of the major sectors in which these jobs are performed](#)  
May 2005 - Elena Sirvent Garcia del Valle
- 05-33 [Een functie met inhoud 2004 - Een enquête naar de taakhoud van secretaresses 2004, 2000, 1994](#)  
April 2005 - Kea Tijdens
- 04-32 [Tax evasive behavior and gender in a transition country](#)  
November 2004 - Klarita Gërxhani
- 04-31 [How many hours do you usually work? An analysis of the working hours questions in 17 large-scale surveys in 7 countries](#)  
November 2004 - Kea Tijdens
- 04-30 [Why do people work overtime hours? Paid and unpaid overtime working in the Netherlands](#)  
August 2004 - Kea Tijdens
- 04-29 [Overcoming marginalisation? Gender and ethnic segregation in the Dutch construction, health, IT and printing industries](#)  
July 2004 - Marc van der Meer
- 04-28 [The work-family balance in collective agreements. More female employees, more provisions?](#)  
July 2004 - Killian Schreuder
- 04-27 [Female income, the ego effect and the divorce decision: evidence from micro data](#)  
March 2004 - Randy Kesselring (Professor of Economics at Arkansas State University, USA) was guest at AIAS in April and May 2003
- 04-26 [Economische effecten van Immigratie – Ontwikkeling van een Databestand en eerste analyses](#)  
Januari 2004 - Joop Hartog & Aslan Zorlu
- 03-25 [Wage Indicator – Dataset Loonwijzer](#)  
Januari 2004 - Kea Tijdens
- 03-24 [Codeboek DUCADAM dataset](#)  
December 2003 - Kilian Schreuder & Kea Tijdens
- 03-23 [Household consumption and savings around the time of births and the role of education](#)  
December 2003 - Adriaan S. Kalwij
- 03-22 [A panel data analysis of the effects of wages, standard hours and unionisation on paid overtime work in Britain](#)  
October 2003 - Adriaan S. Kalwij
- 03-21 [A two-step first-difference estimator for a panel data tobit model](#)  
December 2003 - Adriaan S. Kalwij
- 03-20 [Individuals' unemployment durations over the business cycle](#)  
June 2003 - Adriaan Kalwei
- 03-19 [Een onderzoek naar CAO-afspraken op basis van de FNV cao-databank en de AWVN-database](#)  
December 2003 - Kea Tijdens & Maarten van Klaveren
- 03-18 [Permanent and transitory wage inequality of British men, 1975-2001: Year, age and cohort effects](#)  
October 2003 - Adriaan S. Kalwij & Rob Alessie
- 03-17 [Working women's choices for domestic help](#)  
October 2003 - Kea Tijdens, Tanja van der Lippe & Esther de Ruijter
- 03-16 [De invloed van de Wet arbeid en zorg op verlofregelingen in CAO's](#)  
October 2003 - Marieke van Essen
- 03-15 [Flexibility and social protection](#)  
August 2003 - Ton Wilthagen



- 03-14 [Top incomes in the Netherlands and the United Kingdom over the Twentieth Century](#)  
September 2003 - A.B. Atkinson & dr. W. Salverda
- 03-13 [Tax evasion in Albania: An institutional vacuum](#)  
April 2003 - Klarita Gërxhani
- 03-12 [Politico-economic institutions and the informal sector in Albania](#)  
May 2003 - Klarita Gërxhani
- 03-11 [Tax evasion and the source of income: An experimental study in Albania and the Netherlands](#)  
May 2003 - Klarita Gërxhani
- 03-10 [Chances and limitations of “benchmarking” in the reform of welfare state structures - the case of pension policy](#)  
May 2003 - Martin Schludi
- 03-09 [Dealing with the “flexibility-security-nexus: Institutions, strategies, opportunities and barriers](#)  
May 2003 - Ton Wilthagen & Frank Tros
- 03-08 [Tax evasion in transition: Outcome of an institutional clash - Testing Feige’s conjecture](#)  
March 2003 - Klarita Gërxhani
- 03-07 [Teleworking policies of organisations- The Dutch experience](#)  
February 2003 - Kea Tijdens & Maarten van Klaveren
- 03-06 [Flexible work - Arrangements and the quality of life](#)  
February 2003 - Cees Nierop
- 01-05 [Employer’s and employees’ preferences for working time reduction and working time differentiation – A study of the 36 hours working week in the Dutch banking industry](#)  
2001 - Kea Tijdens
- 01-04 [Pattern persistence in European trade union density](#)  
October 2001 - Danielle Checchi & Jelle Visser
- 01-03 [Negotiated flexibility in working time and labour market transitions – The case of the Netherlands](#)  
2001 - Jelle Visser
- 01-02 [Substitution or segregation: Explaining the gender composition in Dutch manufacturing industry 1899 – 1998](#)  
June 2001 - Maarten van Klaveren & Kea Tijdens
- 00-01 [The first part-time economy in the world. Does it work?](#)  
2000 - Jelle Visser

## AIAS Working Paper Series

The AIAS working paper series consists of several publications of AIAS staff and AIAS guests on a wide variety of subjects in the fields of labour economics, sociology of work, labour law, and health and safety.

ISSN online 2213-4980

ISSN print 1570-3185

## Information about AIAS

AIAS is an institute for multidisciplinary research and teaching at the University of Amsterdam. Founded in 1998, it brings together the University's expertise in labour studies.

AIAS research focuses on the analysis of labour markets, social security and governance. It combines various disciplinary approaches along three perspectives: Societal regulations & coordination of markets, Individual transactions in markets and Societal and individual effects. Some of our research programmes are:

- GINI Growing Inequalities' Impacts
- Equalsoc network of Excellence (Economic Changes, Quality of Life and Social Cohesion)
- Solidarity in the 21<sup>st</sup> Century
- Flex Work Research Centre
- WageIndicator

AIAS offers various in-company courses in the field of HRM, inequality and solidarity, labour market development, labour relations etc.

Annually AIAS organizes conferences about ongoing research and current trends. Furthermore, several (lunch) seminars and workshops take place during the year, offering interesting opportunities for the exchange and deliberation of research on labour issues from all over the world. AIAS has a major collection of academic socio-economic data in the field of labour relations, labour organizations, employment and working conditions in the Netherlands and abroad. AIAS and its staff contribute to society on many subjects, for different audiences and in varying formats (articles, books, reports, interviews, presentations etc...). Next to this Working Paper Series, we also have the series 'Labour markets and industrial relations in the Netherlands' and the GINI Discussion Papers which also addresses a great variety of topics.



## Amsterdam Institute for Advanced labour Studies

University of Amsterdam

Postal address: PO Box 94025 • 1090 GA Amsterdam • The Netherlands

Visiting address: Nieuwe Prinsengracht 130 • 1018 VZ Amsterdam • The Netherlands

Tel +31 20 525 4199 • Fax +31 20 525 4301

aias@uva.nl • www.uva-aias.net