



# Social Science Survey Data Sets in the Public Domain

## Access, Quality, and Importance

(By David HOWELL)

Summarized by  
Jeff ASUNCION & Hilton LAZO



# SOCIAL SCIENCE SURVEY DATA SETS IN THE PUBLIC DOMAIN

---

Access, Quality, and Importance

(By DAVID HOWELL)



University of the Philippines  
CENTER for INTEGRATIVE  
and DEVELOPMENT STUDIES

Co-organized with



DEPARTMENT OF POLITICAL SCIENCE

UNIVERSITY OF THE PHILIPPINES DILIMAN COLLEGE OF SOCIAL SCIENCES AND PHILOSOPHY

100 Years of Leadership, Excellence, and Community in Political Science



# Contents

About the Writers	iv
An Introduction	1
Assessing the Quality of Surveys	4
Datasets and Archives: Some Features	6
Philippine and International Data Archives	10
Using and Citing Data	12
Why Collect and Share Data Together?	13
Summing Up	13
Notes	14
Bibliography	14

Summarized by Jeffrey Asuncion (Publications Officer) and Hilton Lazo (Research Associate). Based from a public presentation given by David Howell last September 16, 2014 and organized by the University of the Philippines-Center for Integrative and Development Studies (UP-CIDS) in cooperation with the Social Weather Stations (SWS) and UP Diliman Department of Political Science (UP DPS).

### About the Writers

As CIDS Publications Officer, JEFFREY ASUNCION has his hands full with revitalizing the *Public Policy Journal* and of coming out with the center's other publications. He is currently finishing his MA History studies.

HILTON LAZO is a Research Associate of CIDS and is currently focusing on issues surrounding the country's energy sector. He is also a Juris Doctor candidate at the UP College of Law.

# SOCIAL SCIENCE SURVEY DATA SETS IN THE PUBLIC DOMAIN

---

Access, Quality, and Importance

(By DAVID HOWELL)

Summarized by Jeff ASUNCION and Hilton LAZO

## An Introduction

Surveys are not just mere sets of numbers. They trace in figures of a story, a pattern, and a trend of what is happening and might be occurring at a certain place and during a certain period. To the trained eye, surveys are useful references in making informed predictions of phenomena that could affect routine political and economic exercises.

Contrary to popular belief, analyzing surveys is very technical and requires experience. Hence, talks by renowned political science experts can help laymen, students, and trained professionals alike to understand further the intricacies of survey preparation and analysis. One such presentation was by American political scientist David Howell. A Michigan-based elections analyst, Howell shared his knowledge and expertise on the proper preparation, use, and dissemination of surveys. As part of his Philippine visit, Howell spoke in front of an audience of students, statisticians, and academics at CIDS last September 16. In the lecture, titled *Social Science Survey Data Sets in the Public Domain: Access, Quality, and Importance*, he expounded on the attributes of high quality surveys, gave examples of accessible datasets and archives, revealed

the correct handling and citation of data gathered from surveys, emphasized the training of people in properly using these information, and highlighted the importance of collaboration in data collection and through data sharing. The lecture was organized by CIDS in partnership with UP-DPS and SWS. The same lecture was also held at Cebu City's University of San Carlos, and at Holy Name University in Tagbilaran, Bohol. Also attending the lecture were CIDS Executive Director Edna Co, UP DPS Chairperson Jorge V. Tigno, and Vladymir Lacudine of SWS. Drs. Co and Tigno are also both fellows of SWS, the three-decades-old statistical institute.

Howell is a key researcher at the University of Michigan's (UMich) Center for Political Studies which is known for its quantitative researches on values, electoral systems, national election studies, and the archiving of election-related data. Howell has spent more than two decades at UMich's Institute of Social Research and had been affiliated with the American National Election Studies. His current scholarly interests are on research project development, program evaluation, international research capacity building, and studies on electoral systems. He is also involved in the Comparative Study of Electoral Systems, or CSES, whose research modules formed a considerable part of his lecture on surveys and free, Internet-based datasets.

CSES describes itself as a collaborative research program among election study teams from various countries. Helping run the Center's secretariat are UMich's Center for Political Studies and GESIS Leibniz Institute for the Social Sciences, with considerable assistance from the German education ministry, German state governments, UMich, and the U.S.-based National Science Foundation. CSES is a unique analytical project because it involves extensive cross-national collaboration and a three-layered data gathering process:

The CSES is composed of three tightly linked parts. First, a common module of public opinion survey questions is included in each participant country's post-election study. These "micro" level data include vote choice,

candidate and party evaluations, current and retrospective economic evaluations, evaluation of the electoral system itself, in addition to standardized sociodemographic measures. Second, district level data are reported for each respondent, including electoral returns, turnout, and the number of candidates. Finally, system or “macro” level data report aggregate electoral returns, electoral rules and formulas, and regime characteristics. This design allows researchers to conduct cross-level, as well as cross-national analyses, addressing the effects of electoral institutions on citizens’ attitudes and behavior, the presence and nature of social and political cleavages, and the evaluation of democratic institutions across different political regimes.<sup>1</sup>

**Comparative Study of Electoral Systems (CSES)**

CSES is a collaborative program of research among election study teams from around the world. Participating countries include a common module of survey questions in their post-election studies. The resulting data are deposited along with voting, demographic, district and macro variables. The studies are then merged into a single, free, public dataset for use in comparative study and cross-level analysis.

The research agenda, questionnaires, and study design are developed by an international committee of leading scholars of electoral politics and political science. The design is implemented in each country by their foremost social scientists.

The CSES is composed of three tightly linked parts: First, a common module of public opinion survey questions is included in each participant country's post-election study. These “micro” level data include vote choice, candidate and party evaluations, current and retrospective economic evaluations, evaluation of the electoral system itself, in addition to standardized sociodemographic measures. Second, district level data are reported for each respondent, including electoral returns, turnout, and the number of candidates. Finally, system or “macro” level data report aggregate electoral returns, electoral rules and formulas, and regime characteristics. This design allows researchers to conduct cross-level, as well as cross-national analyses, addressing the effects of electoral institutions on citizens’ attitudes and behavior, the presence and nature of social and political cleavages, and the evaluation of democratic institutions across different political regimes.

The CSES is unique among comparative post-electoral studies because of the extent of cross-national collaboration at all stages of the project. The research agenda, the survey instrument, and the study design are developed by the CSES Planning Committee, whose members include leading scholars of electoral politics from around the world.

**Tweets**

**Ctr for Political St** @umiscps 17 Oct  
CPS Blog: The fifth module of the Comparative Study of Electoral Systems (CSES): @csestweets cpsblog.isr.umich.edu/?p=1190 Retweeted by CSES

**Pedro Magalhães** @PCMagalhaes 15 Oct  
Where's Waldo @csestweets? A very nice conference, thanks to Bernhard WeBels and the @WZB\_Berlin team. pic.twitter.com/O56K1eSkGZ Retweeted by CSES

SCREENSHOT 1. Source: <http://cses.org/>.

<sup>1</sup> Comparative Study of Electoral Systems, “Home,” <http://cses.org/> (accessed November 20, 2014).

## Assessing the Quality of Surveys

Howell laid down some points in determining the quality of surveys by giving tips on determining the reliability and validity of these quantitative instruments. Quoting W. Phillips Shively, he reminded listeners that, on the one hand, a measure's reliability refers to its producing the same outcome after repeated measurements. On the other hand, validity refers to whether a specific instrument accurately measures the variables it intends to quantify.

In analyzing the reliability and validity of a survey, one should first look at its transparency. A study is transparent if it "provides documentation of its question wording, translations (for multi-lingual settings), sampling strategy, and data collection methods." Howell cited as an example the 1996 CSES Module 1 Election Study Archive study on the Czech Republic. It has an English translation of the survey, originally written in Czech, as well as separate macro and design reports.

One should also look for transparency in a questionnaire's design. Howell said that transparency can be looked at through pretesting of questions and their application to a similar population; the validation by previous surveys of questions to be used in it; the use of reliability tests in testing the strength of new questions; the direct relationship of items to measures; and the use of new questions in focus groups and cognitive interviews.

Howell cited the International Social Survey Programme (ISSP) and the late pollster Janet Harkness in threshing out the transparency of surveys done with the use of translated questionnaires. A critical reader of the translated questionnaires will ask how these were assessed, who did the translation, whether the translated questionnaire was pretested, and of course, whether the translation was accurate. In translating survey questions, one must keep in mind the issue of certain concepts being absent in the context of other languages and cultures. One must also take note whether there were problems in translating the questionnaire.



A critical aspect of any survey is its sampling method, on which Howell asked the following questions: whether the coverage of issues limited the extent to which the subjects or respondents had been gathered; whether the sample gathered had either probability or non-probability components; whether the study had detailed the entire sampling process; and whether the sample size had been “adequate to represent the population and sub-populations of interest.”

Users of surveys should also look at the appropriate mode of interviewing that the researcher had utilized for the survey. Examples of interview modes include online surveys, by mail to be self-completed, by telephone, face-to-face, and mixed.

For field practices, one should look at response and refusal rates, as well as whether interviewers had undergone general interviewer training “and training in the administration of the specific survey.” For refusal rates, one should also look at whether refusal conversion had occurred despite attempts at reaching a high response rate. Refusal conversion is an important aspect to consider in the conduct and retrieval of surveys because it measures the percentage of respondents who did not answer the questionnaires; they may have the characteristics needed for the study. It is also important to consider whether differential non-response had occurred.

As many studies are being conducted in two or more countries, one should also take into account certain cross-national and cultural issues. These concerns affect measurement equivalence and comparison of the results. Howell zeroed in on the kinds of methods applicable for countries considered in a study, the comparable use of questions incorporated in a study, and the constraints that may hamper the response of respondents in a specific nation. Howell also drew into focus the cultural and legal appropriateness of survey questions to be used in various countries. Specifically, he asked, “Do the questions make sense in all of the participating countries? Are the questions culturally appropriate? Were any questions sensitive or not legally allowed in certain countries?”

As an example, he said that it might be considered inappropriate to ask questions on homosexuality in the Middle East. Similarly, it may not be suitable to use the Democratic-Republican bipartisan classification in the United States in other countries. Meanwhile, from a legal point of view, laws in European countries prohibit survey takers from asking about religion. Just as important is to check “whether certain response(s) to questions are legally prohibited.”

After discussing sampling, question-formulation, and data-gathering concerns, Howell then proceeded to post-survey processing. At this stage, a careful analysis of the survey should determine whether the data had been “processed and documented.” Specifically, one should look at the patterns of skipping, absence of certain information, and the presence of outlier checks; the distribution of and relationships between variables; the identification of codes and meanings; and the determination of weights for sample and non-response. It is also crucial to verify the presence of a mechanism that will warn readers of possible “further” errors in the survey.

In light of the problems that may persist in studies, Howell advised that one should determine whether these researches document those issues. As imperfect real world efforts having real constraints, the limitations of surveys should be revealed and even highlighted. By doing so, a researcher enhances a project’s credibility while also improving the quality of analyses that will come out from the surveys.

## Datasets and Archives: Some Features

For the next part of the lecture, Howell showed several interesting features and content of the CSES modules and other survey indices. The module is a repository of 140 studies on elections in more than 50 states worldwide, the Philippines being one of them. It should be noted that CSES modules are in open access mode, which means users can readily download these datasets free of charge and devoid of embargo rules. The modules contain certain dependent variables common to

all, such as voter turnout, choice parameters, party systems and cleavages; economic voting; accountability of governments; satisfaction with democracy; and citizen engagement. There are four modules classified according to periods from the 1990s until the present decade, and also according to specific topics (i.e., distributional politics, mobilization, representation, and accountability; electoral choices; and system performance).

Apart from the CSES modules, Howell also discussed the East Asian, Asian, and Global Barometers. These are studies that dissect various facets of social and political life of people from various parts of the world. East Asian Barometer conducts socioeconomic analyses on certain Asian nations.



**East Asia Barometer**  
 Comparative Survey of Democratization and Value Changes

**Online Data Analysis (Step-1): Sample Selection**

Select the samples you want to analyze and click [Confirm selection] or else click [Select sample] to select a different sample. You can change the sample selection at any time.

Confirm selection Select Round

	Select all
Cambodia	<input type="checkbox"/> 2008
China	<input type="checkbox"/> 2008
Indonesia	<input type="checkbox"/> 2006
Mongolia	<input type="checkbox"/> 2006
Philippines	<input type="checkbox"/> 2005
Singapore	<input type="checkbox"/> 2006
Taiwan	<input type="checkbox"/> 2006
Thailand	<input type="checkbox"/> 2006
Vietnam	<input type="checkbox"/> 2005

Site Shortcuts:

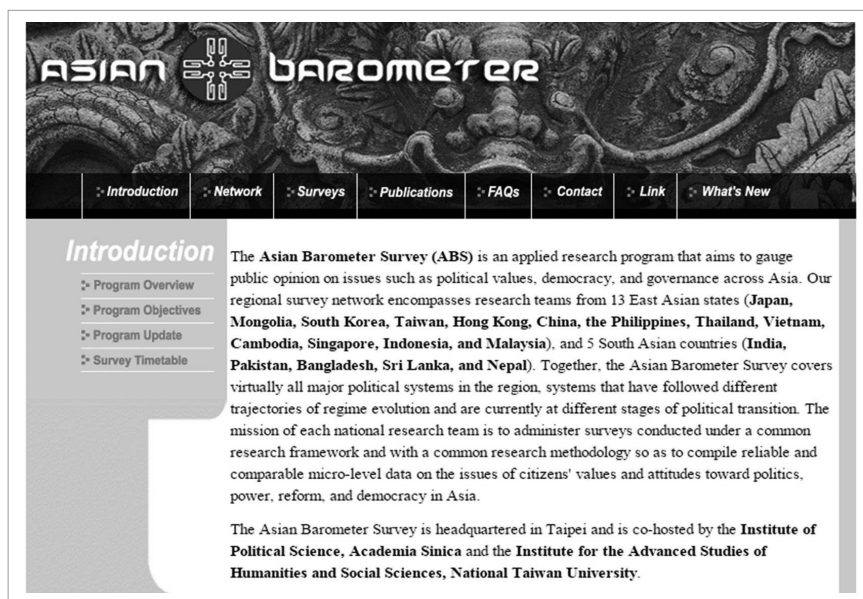
- Download Survey Data Files
- Directory of Investigators
- Online Data Analysis
- Technical information
- Documentation of Data
- Contact

Cambiar a Español

SCREENSHOT 2. Source: <http://www.jdsurvey.net/eab/AnalyzeSample.jsp>.

For the Philippine study, the analysis sought to answer questions on the country's prospects, current and past economic trends, and role in globalization. It also sought to analyze the preference of Filipino voters and their political involvement, trust among citizens of governmental institutions, socio-demographics, satisfaction by citizens with democracy and government, social capital, safety and security, relations with other countries, and the nation's "most important problems."

East Asian Barometer is related to the Asian Barometer which, since its inception in 2003, has been measuring socio-economic indices for East, Southeast, South, and Central Asia. It specifically dissects the "affective and cognitive qualities of life" of respondents; measures the physical, psychological, and sociological dimensions of ordinary Asians; "prepares their (developmental, democratic, and regionalizing) market potential," and reveals the kinds of services and goods that these Asian respondents put premium on "to improve the quality of their own lives and their country's." Here is a screenshot of the Asia Barometer:



SCREENSHOT 3. Source: <http://www.asianbarometer.org/newenglish/introduction/default.htm>.

Finally, Global Barometer’s goal is to measure the “attitudes and values toward politics, power, reform, democracy, and citizens’ political actions.” It has conducted comparative surveys on 55 countries wherein a research team for each country conducts face-to face surveys.<sup>2</sup> This worldwide Barometer is considered as “(representing) the largest, most careful and systematic comparative survey of attitudes and values toward politics, power, reform, democracy and citizens’ political actions in Africa, Asia, Latin America and the Arabic region.”<sup>3</sup>



**Global Barometer Survey**

**GLOBAL BAROMETER STUDY**

Welcome to the online data analysis website for the Global Barometer.

This project represents the largest, most careful and systematic comparative survey of attitudes and values toward politics, power, reform, democracy and citizens' political actions in Africa, Asia, Latin America and the Arabic region. It is based on a common module of questions contained in regional barometer surveys; for additional regional-specific questions, see the original surveys at [www.afrobarometer.org](http://www.afrobarometer.org), [www.arabbarometer.org](http://www.arabbarometer.org), [www.asianbarometer.org](http://www.asianbarometer.org) and [www.latinobarometro.org](http://www.latinobarometro.org).

Cross-national comparative surveys have been implemented in 55 political systems - namely Japan, Korea, Mongolia, Philippines, Taiwan, Thailand, Indonesia, Singapore, Vietnam, Argentina, Bolivia, Brazil, Colombia, Costa Rica, Chile, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Uruguay, Venezuela, Benin, Botswana, Cape Verde, Ghana, Kenya, Lesotho, Madagascar, Malawi, Mali, Mozambique, Namibia, Nigeria, Senegal, South Africa, Tanzania, Uganda, Zambia, Zimbabwe, Bangladesh, India, Nepal, Pakistan, Sri Lanka, Jordan, Palestine, Algeria, Morocco, Kuwait and Lebanon. In each of the 55 countries or regions, a national research team administers a country-wide face-to-face survey using standardized survey instruments to compile the required micro-level data under a common research framework and research methodology.

**Site Shortcuts:**

-  Download Survey Data Files
-  Directory of Investigators
-  Online Data Analysis
-  Technical information
-  Contact

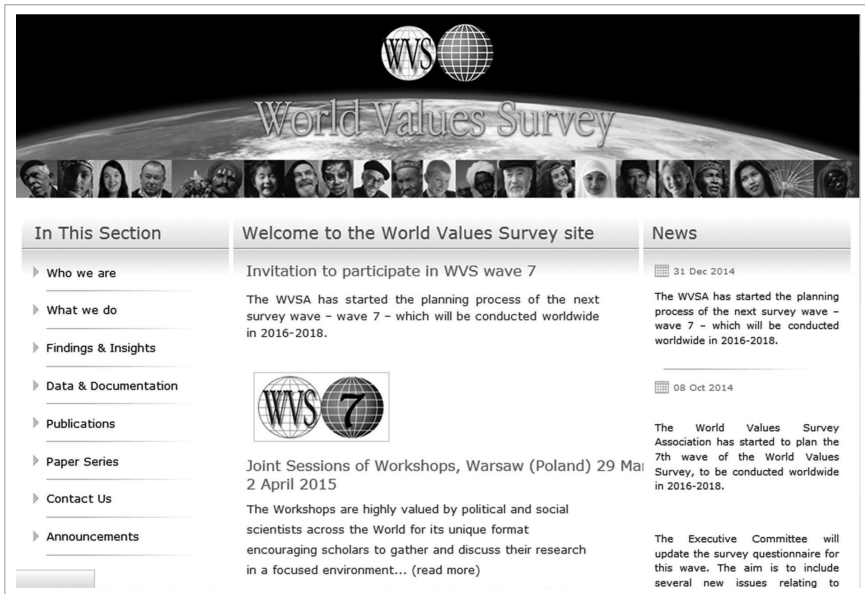
Cambiar a Español

SCREENSHOT 4. Source: <http://www.jdsurvey.net/gbs/gbs.jsp>.

<sup>2</sup> *Global Barometer Surveys*, “Global Barometer Study,” <http://www.jdsurvey.net/gbs/gbs.jsp> (accessed November 20, 2014).

<sup>3</sup> *Ibid.*

The World Values Survey (WVS) and ISSP measure socio-economic indicators for various parts of the globe. Initiated in 1981, WVS has been doing surveys on repression and legitimacy, trust and civic norms, democracy, gender, and globalization, among other topics. So far, its surveys have covered 450,000 respondents; data have been gathered from 100 countries covering nine-tenths of the world's population.



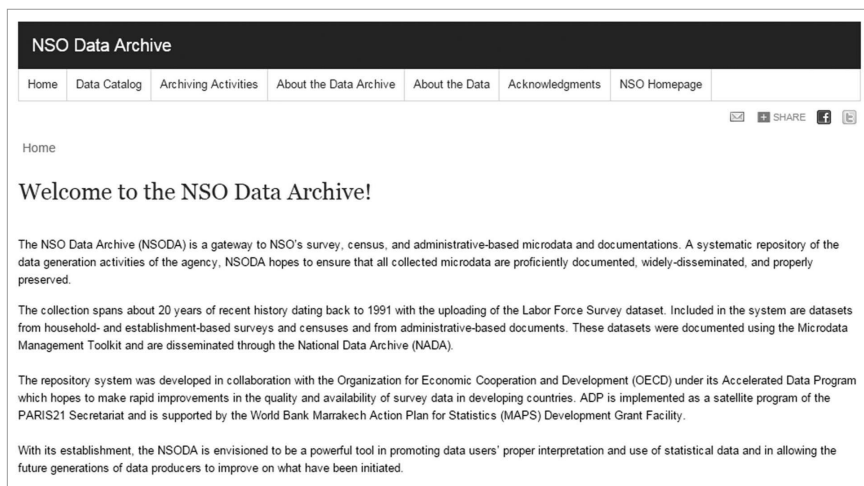
SCREENSHOT 5. Source: <http://www.worldvaluessurvey.org/WVSContents.jsp>.

Meanwhile, ISSP is conducting studies on family, work orientations, citizenship, governments' role in their societies, national identities, and other community involvement issues. Numerous researches have been done in 49 countries. The researches come in the form of modules on questions about social sciences developed to augment the regular national surveys.<sup>4</sup>

<sup>4</sup> The Inter-university Consortium for Political and Social Research, "International Social Survey Program Series," <http://www.icpsr.umich.edu/icpsrweb/ICPSR/series/124> (accessed November 21, 2014).

## Philippine and International Data Archives

Information gathered in surveys is placed in data archives. Facilities in the Philippines that maintain archives of survey information include those of the National Statistics Office (NSO) and SWS. NSO's Data Archives, which were first established in 2009, store figures on yearly surveys about overseas Filipinos, family income and spending, retail prices, family planning, quarterly surveys on the Philippine labor force, and monthly integrated surveys on specific industries. NSO Data Archives also house data sets dating back to 23 years ago, as well as the results of major surveys conducted by health and research groups. Just recently a new law has mandated the absorption of the NSO into a more-inclusive Philippine Statistics Authority.



The screenshot shows the homepage of the NSO Data Archive. At the top, there is a dark header with the text "NSO Data Archive" in white. Below the header is a navigation menu with links: Home, Data Catalog, Archiving Activities, About the Data Archive, About the Data, Acknowledgments, and NSO Homepage. To the right of the navigation menu are social media sharing icons for email, a general share icon, Facebook, and Twitter. The main content area starts with the text "Home" followed by "Welcome to the NSO Data Archive!". Below this is a paragraph describing the NSODA as a gateway to NSO's survey, census, and administrative-based microdata and documentations. It mentions that NSODA aims to ensure that all collected microdata are properly documented, widely-disseminated, and preserved. Another paragraph states that the collection spans about 20 years of recent history dating back to 1991, including the Labor Force Survey dataset, household- and establishment-based surveys, and censuses. A third paragraph notes that the repository system was developed in collaboration with the OECD under its Accelerated Data Program. The final paragraph mentions that the NSODA is envisioned to be a powerful tool for promoting data users' proper interpretation and use of statistical data.

SCREENSHOT 6. Source: <http://census.gov.ph/nsoda/index.php/home>.

Meanwhile, SWS also maintains its own data bank, which is partly fee-based, that stores surveys conducted by foreign organizations, as well as national and sub-national studies conducted since 1984.

Going beyond the Philippines, the well-known data archives in other countries include the U.K. Data Archive, University of Connecticut's Roper Center, the GESIS-Leibniz Institute for the Social Sciences, the Inter-university Consortium for Political and Social Research (ICPSR), and the ASEP/JDS data bank.

## Using and Citing Data

Data gathered from surveys can be utilized either online or offline. With offline use, one can download data into a local computer then analyze such data through fee-based software (SAS, SPSS, or STAT) or free software (R). Offline use of data is good in cases when one needs an in-depth analysis or must create derivative files or requires replication. By contrast, online data use does not require the use or downloading of local software. However, it has limited analytical capability. Online data use is good for exploration of topics by non-analysts and researchers who do not have the required software. One example would be WVS.

In addition, the Internet provides many useful resources and tutorials for data training. Data training lessons may also be obtained from learning collaboration, university courses, and summer institutes. The last two are particularly beneficial because one can develop networks with peers from other countries, even as a learner can understand new concepts about data training. Two such courses are being offered by the University of Michigan: the Summer Institute in Survey Research Techniques and the Summer Program in Quantitative Methods of Social Research. Summer institutes for data training are also to be found at the University of Essex (Summer School in Social Science Data Analysis) and at the National University of Singapore (IPSA-NUS Summer School for Social Science Research Methods).

One cannot underestimate the value of data citation. According to Howell, attributing data used in survey research is just as important as citing information culled from academic papers. Doing so would make



research projects “better known,” enables the people involved in them to “understand how their data are being used,” gives ways for their products to be improved, and also helps them find new ways of financing their products. In general, data sets maintain citation guidance links in their documentation and on their websites. An example of such citation would be: ISSP Research Group (2009): International Social Survey Programme: Leisure Time and Sports - ISSP 2007. GESIS Data Archive, Cologne. ZA4850 Data file Version 2.0.0, doi: 10.4232/1.10079.

## Why Collect and Share Data Together?

Howell then shared seven reasons why scholars, students, policymakers, governments, non-governmental organizations, journalists, data collection groups, and even the general public should exchange data on surveys with each other. The practice enhances their organizations’ credibility and visibility, and promotes scientific integrity by replication of results. Data sharing also preserves and archives data gathered, ensures that data culled “have more impact” which can satisfy funders, and helps professors in teaching data related lessons to students. Users who share data may learn new skills and methods, offer opportunities for presentation and publication, and network with colleagues who share the same interest. This comes amid the trend among funding agencies to “encourage or require the sharing of data.” However, the specifics of such data sharing differ according to data, country, and funding source. Sharing of collated data can encourage collaboration in fundraising activities and participation of stakeholders in research project governance, and also promotes the capacity of one’s organization. At the cross-national level, data sharing highlights the profile of one’s country and also encourages comparative analysis.

## Summing Up

Howell's lecture expounded on the importance of surveys and the steps to be taken to ensure they are reliable and valid. The eminent political scientist also emphasized the importance of sharing data collected through surveys and how these can be best stored, used, and cited. While survey data are numbers, if these are correctly pieced together they can tell a story of what is happening around us. After all, words do not have the monopoly of helping us make sense of the world that we live in.

## Bibliography

- Asian Barometer. "Introduction." Asian Barometer. <http://www.asianbarometer.org/newenglish/introcution/default.htm>.
- Comparative Study of Electoral Systems. "Home." Comparative Study of Electoral Systems. <http://cses.org>.
- East Asian Barometer. "Online Data Analysis (Step 1): Sample Selection." East Asian Barometer. <http://www.jdsurvey.net/eab/AnalyzeSample.jsp>.
- Global Barometer Survey. "*Global Barometer Study*." Global Barometer Survey. <http://www.jdsurvey.net/gbs/gbs.jsp>.
- National Statistics Office. "NSO Data Archives." Philippine Statistics Authority. <http://census.gov.ph/nsoda/index.php/home>.
- The Inter-university Consortium for Political and Social Research. "International Social Survey Program Series." The Inter-university Consortium for Political and Social Research. <http://www.icpsr.umich.edu/icpsrweb/ICPSR/series/124>.
- World Values Survey. "World Values Survey." World Values Survey. <http://www.worldvaluessurvey.org/WVSContents.jsp>.