

**Cognitive Economics Study:  
CogEcon 2011 Documentation**

Version 1.0  
December 2013

By: Michael Gideon, Seth Koch and Brooke Helppie McFall  
University of Michigan

## **Preface**

This documentation provides detailed information about data from the Cognitive Economics (CogEcon) 2011 Survey. Part I provides a brief description of the CogEcon Study and the CogEcon 2011 Survey development and methodology. Part II includes a detailed description of data collected and processed from CogEcon 2011.

Data from CogEcon 2011 can be merged with data from CogEcon 2008 and 2009. Please consult *Cognitive Economics Study: Development and Methodology* for additional information about the development and methodology of the CogEcon Study, specifically the CogEcon 2008 and 2009 waves of the study. Details about the data from CogEcon 2008 and 2009 are available in *Cognitive Economics Study: Data Description*.

Data from CogEcon can also be merged with rich cognitive and demographic data from CogUSA. The merged data provide a powerful tool for understanding the development of financial knowledge and financial decision-making among older adults in the United States.

CogEcon and CogUSA data must be obtained separately. There are additional details about merging these datasets in Section I.A.4. of the Data Description. For more information on the CogUSA study, please visit the website of the Unified Studies of Cognition (USC) at <http://cogusc.usc.edu/>

### **CogEcon 2011 Project Team:**

Robert J. Willis (PI); Matthew D. Shapiro; Jack McArdle; Miles Kimball; Tyler Shumway; Helen Levy; Gwenith Fisher; Pamela Giustinelli; Brooke Helppie McFall; Joanne Hsu; Cynthia Doniger; Peter Hudomiet; Michael Gideon; Nora Dillon.

## **Table of Contents**

To skip directly to a section, hold “Ctrl” and click the name of the section.

<b>Preface</b> .....	<b>2</b>
<b>I. Development and methodology</b> .....	<b>5</b>
<b>A. Background</b> .....	<b>5</b>
<b>B. Sample size</b> .....	<b>5</b>
<b>C. Mode, timing and reminders</b> .....	<b>6</b>
<b>D. Response rates</b> .....	<b>6</b>
<b>E. Survey development</b> .....	<b>7</b>
<b>II. Data Description</b> .....	<b>8</b>
<b>A. Introduction</b> .....	<b>8</b>
1. Conditions of use .....	<b>8</b>
2. Obtaining the data .....	<b>8</b>
3. Structure of data files .....	<b>9</b>
4. Merging CogEcon 2011 data with previous waves.....	<b>9</b>
5. Publications based on data .....	<b>9</b>
6. If you need to know more.....	<b>10</b>
<b>B. Summary of data content</b> .....	<b>10</b>
1. Variable naming conventions.....	<b>10</b>
2. Missing data and “don’t know” responses.....	<b>11</b>
3. Mode differences .....	<b>12</b>
4. Idiosyncrasies in CogEcon 2011 .....	<b>13</b>
<b>C. Description of variables</b> .....	<b>14</b>
1. Identification variables.....	<b>14</b>
2. Sample/system variables.....	<b>14</b>
3. Constructed variables.....	<b>16</b>
3.1 Age at date of survey completion.....	<b>16</b>
3.2 Relationship status.....	<b>17</b>
3.3 Financial respondent status .....	<b>18</b>
<b>D. CogEcon 2011 survey outline and notes</b> .....	<b>21</b>
1. Section A: Introductory questions; self-assessments.....	<b>21</b>
2. Section B: Financial sophistication .....	<b>21</b>
3. Section C: Income, employment and retirement.....	<b>21</b>
4. Section D: Household finances .....	<b>22</b>
5. Section E: Savings and consumption.....	<b>23</b>
6. Section F: Expectations .....	<b>24</b>
7. Section G: Health, taxes and risk tolerance .....	<b>25</b>
8. Section H: Closing questions .....	<b>26</b>
<b>E. Imputations and Calculated Variables</b> .....	<b>27</b>
1. Financial Sophistication .....	<b>27</b>
2. Income and wealth variables .....	<b>30</b>
3. Vehicle imputations.....	<b>31</b>

<b>Appendix 1: Range options provided throughout survey.....</b>	<b>32</b>
<b>Appendix 2: Steps to find car valuations .....</b>	<b>36</b>
<b>References:.....</b>	<b>40</b>

# I. Development and methodology

## A. Background

The Cognitive Economics Study (CogEcon) was designed by a team of economists to increase understanding of the cognitive bases of economic decision-making. This effort was led by Robert J. Willis of the University of Michigan-Ann Arbor, and made possible by a partnership with the Cognition and Aging in the USA Study (CogUSA). CogUSA investigators shared both the sample and their data.

The Health and Retirement Study<sup>1</sup> (HRS) was designed to provide academic researchers, policy analysts and program managers with reliable, current data on the economic and physical well-being of men and women 50 years of age and older in America. Co-investigators with the HRS, Robert J. Willis and Willard L. Rodgers, partnered with John J. McArdle of the University of Southern California (P.I. of the National Growth and Change Studies, NGCS<sup>2</sup>), to launch a new study called Cognition and Aging in the USA.

The CogEcon study was developed and launched as part of the Data Innovation Core of Robert J. Willis's program project (NIA P01 AG026571). The program project was renewed in 2010, with funding for two more waves of the CogEcon study—CogEcon 2011 and CogEcon 2013.

## B. Sample size

The CogEcon 2011 sample included everyone who submitted either CogEcon 2008 or CogEcon 2009 and were still eligible to participate.<sup>3</sup> In total, 951 invitation letters were mailed. The list below describes how the sample was determined:

- There were 985 submissions in 2008;
- Twenty-two web submissions were only partially complete, and these respondents were removed from the sample;
- Fourteen were deceased or had a permanent condition;<sup>4</sup>
- Two respondents did not complete the 2008 survey but were inadvertently invited to take (and subsequently submitted) the 2009 survey. These were invited again in 2011.

---

<sup>1</sup> The HRS is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan.

<sup>2</sup> US National Growth and Change Studies (NGCS) refers to the program of research started at the University of Denver in 1978 by Jack McArdle and John Horn, and now located in the CogUSC Laboratory at the University of Southern California in Los Angeles. The main goal of these studies is to use all available collections of psychological tests to better describe and understand the many changes that seem to occur to people over the adult life-span (ages 18-95). To date, CogUSC research has been funded by the National Institute on Aging (NIA) and has probed deepest into the age-related growth and declines of adult intellectual functioning.

<sup>3</sup> Respondents were considered ineligible to participate in CogEcon 2011 if they were deceased or terminally ill, cognitively or physically unable to complete a survey, no longer participating in CogUSA study, or had withdrawn from the CogEcon study.

<sup>4</sup> Six were removed prior to CogEcon 2009 and eight were removed prior to CogEcon 2011.

### **C. Mode, timing and reminders**

The survey was fielded in October 2011 using both a mail and web version of the questionnaire. Participants were invited to take the CogEcon 2011 web version of the survey if they completed either CogEcon 2008 or CogEcon 2009 on the web, or if they were not invited to complete the web survey but had previously indicated regularly using the Internet. In total, 670 (70.5%) were invited to take the web survey.<sup>5</sup> The remaining 281 (29.5%) respondents were invited to take the mail survey.

Invitation letters were mailed on October 19, 2011. All potential respondents received a \$30 check with their invitation to complete the survey. Mail invitees were provided with the questionnaire and web invitees were given instructions to log in to the Internet version of the survey. An email reminder with a link to the survey was sent to web invitees two days later.

Three reminder letters were sent to individuals who had not yet submitted a survey. First reminder letters were sent after two weeks. Second reminder letters followed two weeks after the first reminder, and third letters followed two weeks thereafter. As with the original survey, the third reminder letter invited respondents to “switch modes.”

### **D. Response rates**

Of the 951 participants invited to take CogEcon 2011, 772 returned a mail survey or submitted at least some responses to the web survey, for a response rate of 81.2%. Additionally, 27 of the 951 invitees were deceased; 10 respondents had a terminal illness or were otherwise physically or cognitively unable to complete the survey (according to an informant); 6 respondents could not be reached; 15 of the 772 submissions were only partially complete.

Excluding deceased and terminally ill respondents, and only counting completed surveys, the adjusted overall response rate is 83.1% (755 of 908). Among respondents who completed both CogEcon 2008 and 2009, the response rate was 83.3% (698 of 838).

The variable *c3\_modesubmit* gives the respondent’s submission mode. The 772 submissions included:

- 527 (68.3%) internet surveys (including 15 partially complete);
- 245 (31.7%) paper-and-pencil (mail) questionnaires;

The median internet respondent reported spending 60 minutes completing the survey, while the median mail respondent spent 70 minutes. The means were 61 minutes and 80 minutes, respectively. A timing variable embedded in the online survey closely corroborated these responses, with a median time of 68 minutes.

---

<sup>5</sup> The sample of web respondents is larger than the sample from 2008 and 2009 because the mode experiment conducted using those waves was discontinued in 2011. Hsu and McFall (2012) describe the mode experiment and analyze the findings.

Mode of invitations and submissions for CogEcon 2011

Invitation Mode	Submission Mode/Status			
	Web	Mail	No response	Total
Web	527 (78.7%)	40 (6.0%)	103 (15.4%)	670 (100%)
Mail	0 (0%)	205 (73.0%)	76 (27.1%)	281 (100%)
Total	527 (55.4%)	245 (25.8%)	179 (18.8%)	951 (100%)

### E. Survey development

The CogEcon 2011 questionnaire was developed in the spring and summer of 2011. Many questions from CogEcon 2008 and 2009 were re-asked to enable panel analyses. The project team drafted other questions from scratch, including new questions aimed at measuring health literacy and tax knowledge.

The wealth module was redesigned from prior waves. The project team reviewed wealth data from prior waves, along with methodology used in the HRS, and settled on a design that falls between the two. While it still includes detailed questions like the ones from CogEcon 2008, it also makes a clear distinction between broad categories of assets in tax-advantaged retirement accounts and assets outside of such accounts. In the end, the project team redesigned the wealth module to include fewer specific categories. It first poses broad questions, and then specifics. While this reduces direct comparability across waves, the project team believes it provides better cross-sectional data.

## II. Data Description

### A. Introduction

The data were collected as part of the 2011 wave of the Cognitive Economics Study and have been processed and stored at the Survey Research Center of the University of Michigan.

Additional documentation, including copies of the mail questionnaire, are available on the CogEcon website (<http://cogecon.isr.umich.edu>).

#### 1. Conditions of use

The CogEcon 2011 public data files contain no individual identifiers, links to individual identifiers, or secondary information that could be used to identify respondents. By removing these variables, the data are effectively anonymized; as a result, secondary data analysis may qualify for “exempt” IRB status.

By accessing the data, you agree to the conditions of use governing access to the Cognitive Economics public release data. You must agree to:

- not attempt to identify respondents;
- not transfer data to third parties except as specified;
- not share your username and password;
- provide information about publications based on CogEcon data via e-mail to [cogeconproject@umich.edu](mailto:cogeconproject@umich.edu);
- report apparent errors in the CogEcon data or documentation files via e-mail to [cogeconproject@umich.edu](mailto:cogeconproject@umich.edu).

For more information concerning privacy issues and conditions of use, please read “Conditions of Use for Public Data Files” and “Privacy and Security Notice” at the Public File Download Area of the HRS web site (URL below).

#### 2. Obtaining the data

CogEcon public data files are available free of charge to registered users. To access the data:

1. Go to the Public File Download Area of the HRS web site, at the URL: <http://hrsonline.isr.umich.edu/index.php?p=reg>
2. Register with HRS (if you have not already). You will receive a password within 24 hours.
3. Log in to the HRS data file distribution system.<sup>6</sup>
4. Once you have logged in, follow the "Datasets and Files" link, then the "CogEcon Contributed Files" link.

We encourage researchers to use the data files in conjunction with the CogUSA data (see next page for more info). A few variables associated with the CogEcon survey are

---

<sup>6</sup> They are available from the HRS data file distribution system, hosted on a secure website maintained on the premises of the Institute of Social Research (ISR) at the University of Michigan.



described in the data description but are not included in the public data files. We are open to requests for such variables and will consider their release on a case-by-case basis.

### **3. Structure of data files**

Data from a mail questionnaire are available if the questionnaire was returned to us. Data from a web survey are available if the respondent accepted the informed consent statement at the start of the survey and answered at least one question.

Variables in the data file are grouped into four broad categories:

- IDs & demographics (from CogUSA)
- Sample/system variables
- Constructed variables
- CogEcon 2011 survey content

Variables are described in this document in the same order they appear in the dataset.

### **4. Merging CogEcon 2011 data with previous waves**

Researchers with access to both the CogEcon public data and CogUSA public data can merge the files using the unique identifier, *sampid*.

For information about CogUSA, please go to the following website: <http://cogusc.usc.edu>

E-mail the Project Manager, Kelly Peters, at [peters@usc.edu](mailto:peters@usc.edu) to request access to view CogUSA data and programs.

Please note that the CogEcon sample is a subset of the CogUSA sample. Each individual-level observation from CogEcon should be matched to exactly one individual-level observation in the CogUSA file.

### **5. Publications based on data**

Please send a copy of any publications you produce based on CogEcon data, with a bibliographic reference, if appropriate, by email to [cogeconproject@umich.edu](mailto:cogeconproject@umich.edu) with “Attn: Papers and Publications” in the subject line.

Include the following citation in any research reports, papers, or publications based on Public Release data:

- In text:  
“The Cognitive Economic Study (CogEcon) is sponsored by the National Institute on Aging (grant number NIA P01 AG026571) and is conducted by the University of Michigan.”
- In references:  
“Cognitive Economics Study, ([insert Product Name]) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA P01 AG026571), Ann Arbor, MI, (year).”

Include the following citation in any research reports, papers, or publications based on Public Release data file tagged as “Early” or “Preliminary”:

“This analysis uses Early Release data from the Cognitive Economics Study, ([insert Product Name]), sponsored by the National Institute on Aging (grant number NIA P01 AG026571) and conducted by the University of Michigan. These data have not been fully cleaned and may contain errors that will be corrected in the Final Public Release version of the dataset.”

## 6. If you need to know more

This document serves as an overview of survey development and data collection for CogEcon 2011. Additional documentation about CogEcon 2008 and CogEcon 2009 are available on our website at <http://cogecon.isr.umich.edu/survey.html>.

If you have questions or concerns about the survey, public data files or the documentation, please send an e-mail to [cogeconproject@umich.edu](mailto:cogeconproject@umich.edu).

## B. Summary of data content

### 1. Variable naming conventions

Most variables from CogEcon 2011 have the prefix `c3_` to indicate that they are from the third wave of the study.

Variables from specific questions are identified first by the wave of the survey and then by their numbers on the mail questionnaire. When multiple variables are associated with the same question, suffixes distinguish between them.

Variable names have the general form:

<code>c3_S#_suffix</code>
---------------------------

With:

- S = letter of the section
- # = question number in mail questionnaire; and,
- suffix is defined according to the conventions in the table on the next page.

Suffix	Meaning
<code>_yn</code>	Yes/no part of questions that ask for yes or no and then for the value only for the yes responses.
<code>_exct</code>	Exact response, for questions with a value and range option.
<code>_rng</code>	Range response, for questions with a value and range option.
<code>_val</code>	Combined exact value and ranges. Exact value, when available. Otherwise, the selected range.
<code>_val_flag</code>	Flag about the <code>_val</code> value—whether exact value, range, implied zero, missing, etc.

_mail, _web	Mail version (only for mail respondents), web version (only for web respondents)
-------------	--

**Example:** Question D15 asks:

“Do you (or your spouse/partner) hold any tax-advantaged retirement accounts, such as IRAs (both classic/traditional and Roth IRAs), 401(k) or 403(b) plan accounts, or Keogh accounts? *If so, what is the current balance/total value of these accounts?(Or range letter if you are unsure)*”

In this case, the yes/no response variable has the suffix “\_yn.” The variable name is *c3\_D15\_yn*. The value response variable has the suffix “\_val.” The variable name is therefore *c3\_D15\_val*. Please read section E.2 for more information about how we imputed values from “range cards” responses.

## 2. Missing data and “don’t know” responses

Each variable with the suffix *\_val* has a flag variable associated with it, *c3\_S#\_val\_flag*, which contains categories for missing values as well an indication whether a respondent gave an exact or range value response. These sets of variables are available for questions about income, food consumption, financial and non-financial assets, and debts. These variables contain reported monetary values. The table below describes the flags in more detail.

<b>c3_S#_val_flag</b>		
<b>Label</b>		<b>Description</b>
ExactVal	1	Respondent reported an exact value for the amount.
Range	2	Respondent reported a range for the amount. For internet surveys, this implies that the respondent skipped the question asking for the exact value.
Missing	3	The respondent should have provided a value or range, but did not do so.
ImpliedZero	5	Respondent answered a question implying that the value is zero. This is either the Y/N part of the question, from a prior question, or because of relationship status precludes them from answering (e.g. <i>c3_C42_val</i> about spouse/partner’s earnings)
RespDontKnow	6	Respondent said he does not know the value.
RespSkip	9	We do not know whether the respondent has a value for this item because he/she skipped questions needed to know this.

### Additional notes:

- For questions without a Yes/No component—about household income (*c3\_C2*), food at home (*c3\_E19*), and food away from home (*c3\_E20*)—skipped responses are counted as missing rather than implied zero.
- *ImpliedZero*: For example, single respondents do not answer the question about spouse’s/partner’s employment, and people who indicated that they do not hold

retirement savings accounts skip past questions about assets within those accounts. Note that a respondent who does not have an asset but then writes “\$0” for the amount will be flagged, here, as an ImpliedZero even though in the data she also has a value.

### **3. Mode differences**

Most questions on the mail and web instruments were identical, but there are some differences that we have detailed below.

#### **3.1 Instructions and pre-filled information:**

Wording differed slightly in the instructions and “fills”—i.e., the Internet version automatically fills “spouse” or “partner” where relevant, whereas the mail questionnaire has “spouse/partner.”

In the web questionnaire, we were able to preload relationship status and age. We asked for confirmation/updates of this information before section D. In the mail version, the relationship update is asked in question D1. In the mail version, questions F11 and F14 are asked because age cannot be carried through for skip logic purposes in mail surveys.

Married respondents, and those who reported that they were both (a) in a “marriage-like” relationship and (b) “planning a financial future together,” were asked to include the assets and debts of their spouses/partners in questions about their finances. Singles, and respondents in “marriage-like relationships,” but who were *not* planning a financial future with their significant others, were asked only about their own assets and debts.

The wording of the questions on the mail questionnaire provided reminders of these instructions, whereas the correct wording was “piped” into the web instrument.

#### **3.2 Exact values vs. range of values**

For questions about income, assets and debts, the web instrument first asked for an exact value. If the respondent clicked “next” without responding with an exact value, the instrument then asked the respondent to either return to the previous page or choose from a list of ranges below that message.

In contrast, for most income, asset or debt questions, mail respondents were given the option of providing an exact value or choosing from a list of ranges on the same page. While not ideal, the only reasonable way to offer a range card option on the mail survey was to provide both on the same page.

Based on analysis of past waves, we decided to limit this option to certain questions on the mail survey because providing a range card option reduces missing values, but increases the likelihood of a (less exact) range response over an exact value. On the mail questionnaire, the cases in which we did **not** provide a range option include C5, C8, C11, C14, C16, C19, C21, C24, D9, D31, D34, D35, E3, E4, E7, E8, E10, E13, E15, E19, E20, E21, E22, and F12.

Because of additional measurement error that could be introduced in questions offering both a range and a value, anything with the suffix *\_val* should be used with this difference in mind.

Most questions that offered a range card included “Cannot provide a range” as an answer option. Questions C33, C42, D5, D7, D12, and D13 are the exceptions.

### **3.3 Reminder screens in web mode**

In general, the mail questionnaire doesn’t offer reminders, offers fewer opportunities to answer with ranges, and all questions that do offer ranges clearly display these on the same page as we ask for values. Also, on the web we offer “don’t know” as an answer on most reminder screens, while this option is rarely given to mail respondents.

On the web, for most questions that were not asking for values of income, assets or debt, those who skipped a question would simply see the original screen again, plus a red reminder text asking the respondent to answer the question, if willing.<sup>8</sup> Additionally, as with the original survey, “don’t know” was offered as an answer choice for multiple choice questions, but offered only the second time a respondent saw a question.

Questions offering “don’t know” on reminder screen as alternative to not answering:

C6, C9, C12, C15, C17, C20, C22, C27, C28, C29, C39, D6, D13,  
D15\_ACCT\_A-E\_R, D16\_ACCT\_R, D17, D20, D21, D22, D23, D24,  
D26, D27, D28, D32, D33, E0, E1, E2, E5, E6, E9, E12, E14, E16, E25,  
E26, E27, G13

Questions offering reminder language asking the respondent to return to previous screen and complete the question:

E24, G6, G7, G9, G10, G12, G14, G15, G17

### **3.4 Scale differences**

Small differences that are question-specific are mentioned in the description of the survey content.

## **4. Idiosyncrasies in CogEcon 2011**

### **4.1 Versions/Randomizations**

Each respondent was assigned one of four versions of the survey, which differed along two dimensions.

Section B was the True/False financial sophistication section. Each of the 16 true/false financial sophistication questions had a true version and a false version. One aspect of the randomization had to do with which versions of these questions a respondent received. The second dimension was the ordering of the questions. The 16 questions were split

---

<sup>8</sup> For programming reasons, there were a few exceptions to this. Variables for which a separate reminder screen existed can be seen in the original data. These variables are not in the public release data files.

into two parts, and one half was presented first while the other was second. The differences are explained in more detail in Section 7.1.

The variable *c3\_versassign* indicates whether respondent had version 1a, 1b, 2a or 2b.

## C. Description of variables

### 1. Identification variables

Every respondent has three identification numbers, each described in the table below. The variable *sampid* uniquely identifies respondents. It is a 10-digit string variable that has *chhid* as the first six digits, followed by a zero (“0”) and then *cpn* as the remaining three digits. We recommend keeping these three variables as string variables to prevent loss of digits.

Variable	Description
<i>sampid</i>	Unique ID for each respondent (10-digit string)
<i>chhid</i>	Household unique ID (6-digit string)
<i>cpn</i>	Person number ID (3-digit string)—identifies individuals within a household

### 2. Sample/system variables

Variable	Description
<i>c3_result</i>	Survey outcome result for those who were invited <sup>9</sup> =1 for completed submissions =2 for partially completed submissions
<i>c3_modeassign</i>	Assigned mode =1 if mail survey =0 if web survey
<i>c3_modesubmit</i>	Submission mode =1 if mail survey =0 if web survey
<i>c3_versassign</i>	Assigned version of the survey
<i>c3_verssubmit</i>	Submitted version of the survey <sup>10</sup>
<i>c3_lastQ</i>	Last question completed <sup>11</sup>

<sup>9</sup> This is the variable “datstatsubmissionstatus” from the Illume files.

<sup>10</sup> There is only one observation for which the submitted mode differs from the assigned mode.

<i>c3_compmo</i>	Month in which survey was submitted
<i>c3_compyr</i>	Year in which survey was submitted

Please note the distinction between the mode that someone was invited to, and the mode submitted. Invitees were given the option to “switch modes” with the third reminder letter.

---

<sup>11</sup> This variable is useful for respondents who submitted a partially completed survey, since it helps distinguish between people who answered most of the survey and those who completed very little of it.

### 3. Constructed variables

The dataset includes constructed variables for age at the time of the survey (*c3\_age\**), relationship status (*c3\_relstat*), an indicator for whether the individual is a household's "financial respondent" (*c3\_finr\**) as well as measures of "financial sophistication" that were derived using the set of true/false statements in section B. The financial sophistication measures are described in section E.1. All others are described below.

#### 3.1 Age at date of survey completion

The variable *c3\_age* contains respondents' age on the date they completed the survey. It was created using the same procedure CogUSA used to construct their age variables. First, calculate the number of days between the respondent's birthday and date they completed the survey. Then divide the number of days by 365.25 to convert units into years.

Birth year was available for all respondents, but seven respondents did not have birth month or day available. The variable *c3\_birth\_flag* indicates whether neither day nor month are missing (N=765), both are missing (N=3), or only day is missing (N=4). When month is known but day is not, we calculated the age as if the day is the 15<sup>th</sup>. When both day and month are missing:

- i) If survey was completed on or before July 1st, assume respondent has not yet had birthday that year (assign birthday to July 1).
- ii) If survey was completed after June 30, respondent already had their birthday that year (assign to June 30).

The variable *c3\_compdate* contains the date the survey was completed.<sup>12</sup> For web respondents (N=527), date completed is the date the survey was submitted. For mail respondents, date completed is either the self-reported date completed (N=230), *c3\_H3*, or is imputed using the date SRO received the questionnaire (N=15).<sup>13</sup> When it was missing, date completed was imputed as

$$(\text{date received}) - (\text{median delay})$$

where median delay = median difference between date received by SRO and self-reported date completed for respondents in i).<sup>14</sup>

*c3\_compdate\_flag* indicates whether completion date comes from the Illume date stamp, self-reported date (H3), or from SRO records of the date received.

*c3\_age\_flag* indicates whether calculated age uses imputed birth date (N=7), imputed completion date (N=15), or no imputed values (N=750).

---

<sup>12</sup> This variable is not available in the public release files.

<sup>13</sup> There were 13 questionnaires without responses to H3. For two others, the date was impossible, since it was from before the survey was fielded

<sup>14</sup> Some respondents did not have a "date completed" observation in the SRO files. Looking at the SRO notes, it appears that this only happened for questionnaires received on 12/22/2011. In such cases the date received is replaced with 12/22/2011, as indicated in the SRO case notes.



### Other age variables:

- *c3\_age\_y*: age in years, as an integer.
- *c3\_age\_m*: age in number of months.
- *c3\_age\_flag*: flag variable that indicates whether the birthdate or survey date was missing and needed to be imputed.

### 3.2 Relationship status

*c3\_relstat* is the respondent's relationship status. We used question D1 for mail respondents and we had web respondents verify the pre-loaded relationship status before starting section D. Note that there are 7 respondents who claimed to be married in CogEcon 2011 but later in the questionnaire said they had been widowed at some point since January 2008. Because 5 were widowed in 2010 or 2011 we suspect that these people are not referring to a remarriage in their CogEcon 2011 relationship status.

Code	Relationship status	CogEcon 2011
1	Married	524 (67.9%)
2	Partner with financial future	18 (2.3%)
3	Partner without financial future	10 (1.3%)
4	Single	220 (28.5%)
Total		772

*c3\_relstat\_chng* tells us whether relationship status changed since the previous wave. For most respondents this is the change since 2009, but for respondents who completed CogEcon 2008 but not CogEcon 2009 the change is based on 2008 to 2011.

Code	Relationship status	CogEcon 2011
0	No change	728 (94.3%)
1	Partnered to single	18 (2.3%)
2	Single to partnered	13 (1.7%)
3	Widowed	13 (1.7%)
Total		772

### 3.3 Financial respondent status

CogEcon asks all respondents about assets and debts. For household-level analysis, we recommend using values given by the designated financial respondent.

There are two approaches used to designate a financial respondent for each household in CogEcon 2011. The variables *c3\_finrA* and *c3\_finrB* equal 1 for financial respondents, and equal 0 otherwise. *c3\_finrA\_how* and *c3\_finrB\_how* tells you how the financial respondent was chosen, as described below:

#### Financial respondent method A: using CogEcon 2011 data only

- (1) If respondent was the only person in a household invited to CogEcon 2011, he/she is the de facto financial respondent (N=392).
- (2) If two household members were invited to CogEcon 2011 but only one completed it, he/she is the financial respondent (N=32).
- (3) If two household members completed CogEcon 2011 but at least one of them claims to be single or partnered without a financial future, then both are assigned to be a financial respondent (N=10).<sup>15</sup>
- (4) All other cases involve households with two respondents who agree they are married or partnered and planning a financial future together. We assign a financial respondent based on responses to questions about knowledge of their household finances (N=296).
  - a. D2: “Which member of your immediate family is most knowledgeable about your family’s assets, debts and retirement planning?” The answer options were “Me”, “My spouse/partner”, “Both me and my spouse/partner” and “Someone else in the family.”<sup>16</sup>

Respondents who answered “Both me and my spouse/partner” or “Someone else in the family” were asked a follow-up question:
  - b. D3: “Suppose you and your spouse/partner were asked to provide information about your combined assets, debts, and retirement plans. Between you and your spouse/partner, who could provide the most accurate information? Please check the box “Me” if you do not have a spouse or partner.” The answer options were “Me” and “My spouse/partner.”

If one respondent said “Me” to either D2 or D3, and the other did not disagree by also saying “Me” to either D2 or D3, then this person is considered the most knowledgeable and the designated financial respondent. If one respondent said “My spouse/partner” to either D2 or D3, and the spouse/partner did not disagree by also saying “My spouse/partner,” then this spouse/partner is considered the most knowledgeable and the designated financial respondent.
- (5) If we still have not assigned each household a financial respondent, we selected the partner/spouse who used a larger number of objective information sources to complete the questionnaire (N=18). This was based on responses to H1 at the end of the questionnaire, and counted financial software, tax returns and account statements as objective sources of information.

---

<sup>15</sup> There were a few cases in which one person claimed to be married and the other single. There were also a few cases of partners, one of whom claims they have a financial future together and the other does not.

<sup>16</sup> This question is very similar to one used in the HRS for the same purpose.

- (6) If both used the same number of information sources, we compared their responses to questions A8 and A9, about confidence listing financial assets with and without account records (N=16).
- a. A8: “How much do you agree or disagree with the following statement? Without my financial records, I could list the assets in my savings and investment accounts, along with the approximate value of each asset.”
  - b. A9: “How much do you agree or disagree with the following statement? Using my financial records, I could list the assets in my savings and investment accounts, along with the approximate value of each asset.”
- The answer options were “Strongly agree,” “Agree,” “Slightly agree,” “Slightly disagree,” “Disagree,” and “Strongly disagree.” If neither used information then we compared responses to A8 (confidence without financial records) (N=10), while if both used the same number of information sources we compared responses to A9 (confidence with financial records) (N=6). In both cases, the person who expressed a higher level of agreement with the statement was chosen as the financial respondent.
- (7) If we still had not assigned each household a financial respondent then we selected one on a case-by-case basis (N=8).<sup>17</sup> For these four remaining couples we selected based on an assessment of financial sophistication scores, cognitive ability, and responses to key questions about financial assets.

### **Financial respondent method B: using CogEcon 2008 FINR as a baseline**

For CogEcon 2011, *c3\_finrB* is based on the CogEcon 2008 variable *c1\_finr* unless the designated financial respondent did not complete the 2011 survey or relationship status changed between waves.

- (1) Repeat steps (1)-(3) used to identify financial respondent A (N=434).
- (2) All other cases involve households with two respondents who agree they are married or partnered and planning a financial future together. Assign financial respondent status from 2008 (or, 2009, if R did not complete 2008) (N=336).
- (3) If both were financial respondents in 2008 and 2009, then we use the financial respondent from method A to select among them (N=2).<sup>18</sup>

Please see Cognitive Economics Study: Data Description for documentation about how the financial respondent was selected in 2008 and 2009. The procedure using CogEcon 2011 is similar but incorporates information from a “tie-breaker” question (D3) that we included for cases in which the respondents says both partners are equally knowledgeable about their household assets, debts and retirement planning.

---

<sup>17</sup> We performed the same analysis for respondents who received FinR assignment based on A8, A9 or the number of information sources, and this reinforces our decision to use those rules before assigning on a case-by-case basis. More information about the hand chosen financial respondents is in the file “*finr\_handchoice\_9-7-2012.txt*”

<sup>18</sup> This would only be the case if they were originally assigned the same household ID, considered themselves single or partnered without a financial future in 2008, but then married or partnered with a financial future in 2011.

How financial respondent was selected in methods A and B

Method A	Method B: using 2008 Finr as baseline			Total
	Only R	2008 finr	uses c3_finrA	
Only R in HH	434	0	0	434
D2, D3- agree	0	294	2	296
More info sources	0	18	0	18
No info: A8 confidence	0	10	0	10
Same info: A9 confidence	0	6	0	6
Chosen by hand	0	8	0	8
Total	434	336	2	772

The table below gives the breakdown of the number of people who are the financial respondent according to method A or method B.

Financial respondent in 2011 based on methods A and B

Method A	Method B: using 2008 Finr as baseline		
	No	Yes	Total
No	139	30	169
Yes	30	573	603
Total	169	603	772

## **D. CogEcon 2011 survey outline and notes**

This part of the documentation contains an outline of the questions and notes of things to consider when working with these data. The questionnaire had eight sections, A thru H.

### **1. Section A: Introductory questions; self-assessments**

#### **A1-A4: E-mail and web usage**

*Note:* Variables associated with A2 are dummy variables for whether an e-mail address was provided.

#### **A5-A12: Self-assessed health, financial knowledge, and day-to-day finances**

*Note:* Respondents could select one or two answers for question A12. The variable *c3\_A12* is a categorical variable for whether the respondent's household manages their own finances (=1), hires a financial planner or advisor (=2), or both uses an advisor and manages themselves (=3). The variables *c3\_A12\_1* and *c3\_A12\_2* treat each check box as a separate Yes/No question, at equal one if the respondent selected it as an answer.

### **2. Section B: Financial sophistication**

*Please see Section E.1 for details about CogEcon 2011 financial sophistication questions and how they compare with CogEcon 2008 and CogEcon 2009.*

### **3. Section C: Income, employment and retirement**

#### **C1: Number of household members**

#### **C2-C24: Household income**

*Note:* CogEcon 2011 questions about Social Security (C3-C14) and employer-sponsored pensions (C15-C24) are asked separately for the respondent and the respondent's spouse/partner. In previous waves these questions were only asked at the household level.

#### **C25-C29: Advice about money management; tax preparation and planning**

*Note:* For questions C25, C28 and C29, respondents were directed to select all answers that apply. Text responses to C25 are not publicly available.

#### **C30-C46: Labor supply and earnings (in 2010); current employment status; retirement plans**

*Note:* The first few questions ask about the respondent and remaining ones ask about the respondent's spouse/partner. These questions are before the instructions (associated with question D1) distinguishing between partners with whom someone is planning a financial future. There might be responses about partners in this section although wealth questions do not include the partner's assets.

For questions C34 and C43, the number in the variable subscript corresponds to the number of the checkbox associated with this employment status. The ninth check box was for “other,” and provided space for respondents to answer. Responses provided in these text boxes are coded in the variables *c3\_C34\_9code* and *c3\_C43\_9code*, respectively. If working fulltime then *c3\_C34\_9code*=1, if working part time then *c3\_C34\_9code*=2, if unemployed and looking then *c3\_C34\_9code*=3, and so on.

*Comparability details:* Question 22 was asked of all respondents in the mail mode, but only of internet mode respondents who had indicated that they were retired in question 21.

C35 is web-only; it asks age of retirement, and is asked only if R checked “retired” in C34. On mail, C35\_mail is asked of all R’s, and is a yes/no question with age of retirement if “yes”. On the web, C36 is asked if C34 says “retired,” while in the mail C36 is asked if C35\_mail says “yes.” Parallel logic holds for C44/C44\_mail and C45. The variables *c3\_C35\_mail\_age* and *c3\_C44\_mail\_age* are the specified age when they retired.

C37\_1 and C46\_1 are only asked on web (and only if first age of expected retirement question, C37 for R or C46 for SPP, is skipped)

## **4. Section D: Household finances**

### **D1-D3: Relationship status, household financial decisionmaking**

*Note:* D2 and D3 are used to select a financial respondent for each household in the sample (*c3\_finrA* and *c3\_finrB*); see Section C.3.3 for more detail.

*Comparability details:* D1 was only asked on the mail survey and responses are in the variable *c3\_D1\_mail*. The question was used as a way to then give directions for questions about assets and debts. The answer to this question determined whether the asset and debt questions were asked with respect to the respondent only, or with respect to the respondent and his/her significant other.

Question D1 was not needed on the web survey because the information was already contained in previous answers and questions were tailored automatically based on this information.

Married respondents and those in “marriage-like” relationships who said they were “planning a financial future together” were asked to include the assets and debts of their spouse/partner in the questions that followed. Singles and respondents in “marriage-like relationships,” but who were not planning a financial future with their significant others, were asked only about their own assets and debts.

Question D2 on the web version only offers the middle two answer options if R has a spouse or partner.

#### **D4- D14: Housing wealth—primary residence and other property**

*Note:* Respondents who do not own their primary residence are asked whether they pay rent (D8-D9). If they neither own nor rent their primary residence they are asked what they do. Text responses to D10 are not publicly available.

#### **D15: Financial assets in tax-advantaged retirement assets**

*Comparability details:* When asking for the breakdown of the total value into asset categories, the web version pipes in totals and gives current total/ more cues to help Rs give values that add up; this is not possible in the mail version.

#### **D16: Financial assets outside of tax-advantaged retirement accounts**

*Comparability details:* When asking for the breakdown of the total value into asset categories, the web version pipes in totals and gives current total/ more cues to help Rs give values that add up; this is not possible in the mail version.

#### **D17-D24: Asset Allocation**

*Note:* D20-D23 and D28: These items ask for values of assets if the respondent holds more than \$5000 in that asset. On the mail, Rs can give values <\$5000; on web, they are constrained to give a value above \$5000. Also, range 3 on the web (the lowest possible) starts at \$5,000, while the corresponding mail category starts at \$5001 because offering a separate range card for these questions would be confusing.

#### **D25-D26: Stock trading frequency and reasons**

#### **D27-D28: Farm and business ownership; any other assets**

*Note:* Questions 27 and 28 ask about other assets, but put constraints on what constitutes “yes.” Question D27 asks about ownership when above 5% of the business or partnership, and D28 asks for the value of other assets if worth more than \$5000.

#### **D29-D37: Credit and Financial distress**

### **5. Section E: Savings and consumption**

#### **E1-E13: Active savings**

*Comparability details:* Respondents were told to skip questions E1 thru E5 if neither the respondent nor the spouse/partner had an employer in 2010. On the web there was a question E0 asking if they had an employer in 2010. E11 is always asked in the mail version and only if D15=yes in web version.

#### **E14-E18: Distributions from tax-advantaged retirement accounts**

#### **E19-E22: Household spending**

*Comparability details:* Mail questionnaires did not allow range responses. The web survey asked for a range if the respondent skipped the question asking for an exact value.

## **E23-E24: Vehicle assets and loans**

*Please see Section E.3 for details about how vehicle asset values were imputed.*

*Comparability details:* The web version of E23 explicitly asks for the number of vehicles, which is in the variable *c3\_E23\_num*. The mail version only offers four lines and does not specify that it should be the four most valuable vehicles.

## **E25-E27: Spending adjustments**

*Notes:* These questions are repeated from 2009, but with a different time horizon. In CogEcon 2009 it asks about changes since July 2008, whereas in CogEcon 2011 it asks about changes in the past 12 months.

## **6. Section F: Expectations**

### **F1-F4: Inflation, interest rates, equity premium, stock market volatility**

### **F5-F15: Subjective probability questions about housing, unemployment, Social Security and stocks:**

*Notes:* Questions F7, F11 & F17 are asked only on the mail version. F11-F17 are only asked if respondents are of a relevant age. On the mail surveys, questions F11 & F14 prompt the respondent whether they should answer F12-F13 and F15-F17, respectively. On the web version, a preloaded variable with age is used to determine whether the respondent should be asked each set of questions. F9\_p and F10\_1 are asked only on the web version. The validation check in F10 (that shows if answer sums to 100%) is only on the web version as well. F8 depends on F7 on the mail version. On the web version it depends on work status from Section C.

*Comparability details:* Probability questions on the web version ask a follow-up question to responses that are 50%. This question is meant to clarify whether the respondent meant that the probability is 50/50 or if the respondent did not know. These questions include F5\_50, F6\_50, F13\_50, F15\_50, F16\_50 & F17\_50.

### **F16-F17: Hypothetical changes in value of retirement accounts**

*Notes:* Questions F16 and F17 asked about responses to hypothetical changes in the value of retirement accounts. The mail version always asked about a \$50,000 change. On the web the questions were asked if people said they had retirement accounts (yes to D15) and specified a round number around 30% of the total value in those accounts. If respondent gave exact value for D15 then a value was used, if respondent gave a range for D15 then a range was used for F18, and if respondent has retirement accounts but did not provide a value or range then 30% was used. The flag variable *c3\_F16\_flag* equals 1 if a value was used, 2 if a range was used, and 3 if 30% was used. The amounts used are available in the variables *c3\_F16\_flag\_val* and *c3\_F16\_flag\_rng*. *c3\_F16\_hyp\_chg* indicates whether the hypothetical change was a loss or gain, and then *c3\_F16\_hyp\_loss* and *c3\_F16\_hyp\_gain* give the hypothetical loss and gain, respectively.



<b>Range selected in D15</b>	<b>Range presented in F16</b>	<b>Value assigned in c3_F16_hyp_loss &amp; c3_F16_hyp_gain</b>
\$0		0
\$1-\$2,500	\$0 - \$750	1250.5
\$2,501-\$5,000	\$750 - \$1,500	3750.5
\$5,001-\$10,000	\$1,500 - \$3,000	7500.5
\$10,001-\$25,000	\$3,000 - \$7,500	17500.5
\$25,001-\$50,000	\$7,500 - \$15,000	37500.5
\$50,001-\$100,000	\$15,000 - \$30,000	75000.5
\$100,001-\$250,000	\$30,000 - \$75,000	175000.5
\$250,001-\$500,000	\$75,000 - \$150,000	375000.5
\$500,001-\$1,000,000	\$150,000 - \$300,000	750000.5
More than \$1,000,000	More than \$300,000	1400001.4
Cannot provide a range		

**F18-F23: Subjective probability questions about housing, unemployment, Social Security and stocks:**

*Comparability details:* F18 provided a range option on the web. F19: always asked after F18 on mail; on web, ½ the time was asked before and ½ the time after (randomized—got it before if version 1a or 1b, after if version 2a or 2b)

F19\_V2\_50, F19\_V1\_50, F20\_50, F21\_50, F22\_50 and F23\_50 are asked on web only, as follow-ups to 50% answers.

**7. Section G: Health, taxes and risk tolerance**

**G1-G5: Risk tolerance**

*Notes:* Questions G1-G5 can be used to estimate risk preference parameters. The answer to G1 can be used to create the first “bound” on measures of risk aversion, and determines whether respondents should then answer questions G2-G3 or G4-G5 to determine the other bound.

*Comparability details:* On the web G4 was only asked if the respondent answered “no” to G1. Mail respondents were asked G4 if they answered “no” to G1 or if they answered “yes” to G3. Responses to G4 resulting from the latter case should be ignored when imputing risk tolerance.

**G6-G10: Federal income tax rates**

**G11-G15: Health literacy**

*Notes:* Some responses to G14 on the mail survey were coded differently because they were reasonable but not what was intended. The month was filled in, even if

R did not give date/day of month. For G14, respondents were asked to give a date (1-31). However, given the instructions, some respondents wrote a day (Monday, Tuesday, etc) or something else. The variable *c3\_G14\_alt* codes the alternative response to the part about date. *c3\_G14\_alt\_text* is a description of the answer when neither a day of week nor a date.

Questions G12, G14, and G15 were “multi-control,” which were programmed differently from others. In these cases, respondents were prompted to answer the question by selecting an option from a drop-down menu. In these cases, when someone skipped the question the default choice (e.g., which said “--hour--” when the menu provided hours). These have been recoded as missing. For mail respondents we do not know if they skipped the question or if they provided an answer that was not hourly and therefore could not be entered into the Illume console. This was quite uncommon and is only relevant to the variables:

*c3\_G17\_mo, c3\_G17\_yr, c3\_G12\_hour, c3\_G12\_ampm, c3\_G14\_mo, c3\_G14\_day, c3\_G15\_hour, c3\_G15\_ampm.*

### **G16-G17: Widowed since January 2008**

*Notes:* Some respondents answered “yes” to G16 even though they were widowed prior to January 2008.

### **G18: Political affiliation**

### **G19: Lump-sum tax rebate**

## **8. Section H: Closing questions**

### **H1: Use of financial records to answer questions**

### **H2: Personal assistance to answer questions**

*Comparability details:* Slightly different wording between mail and web versions. On the web respondents who said someone helped with filling out the survey were given space to explain. This is relevant to few people and the data is not publicly available.

### **H3-H5: Date and time completing survey**

*Comparability details:* H3 was only asked in the mail survey because the information was saved automatically for web respondents. The variable *c3\_H3\_\** combines the data from the mail and web. The web variable is *c3\_datewebcompleted*. The variables are combined into *c3\_time\_hour* (hour and minutes) and *c3\_time\_ampm* (AM or PM). These date variables are not publicly available.

## E. Imputations and Calculated Variables

### 1. Financial Sophistication

There are two versions of each True/False statement. We standardized the responses in order to compare them across respondents and questions. We also constructed summary measures of “financial sophistication” using the entire set of true/false statements. The randomization of questions, the procedure for constructing variables from the responses, and the summary measures are described below in greater detail.

#### Randomization

Respondents were randomly assigned to receive one of four versions of the survey: 1A, 2A, 1B, or 2B. Versions 1A and 1B had the same set of questions, but were ordered differently. Relative to version 1A, in version 1B questions B8-B14 were placed in front of questions B1-B7, and the order of questions B15 and B16 were swapped. Relative to version 2A, in version 2B questions B8-B14 were placed in front of questions B1-B7, and the order of questions B15 and B16 were swapped.

Additionally, every question that was the “True” version in set 1 was the “False” version in set 2, and vice-versa. For example, the variable “*c3\_verssubmit*” is equal to 1A, 2A, 1B or 2B; corresponding to the version of the questionnaire they submitted.<sup>19</sup> The public release files include version 1A of the questionnaire. The appendix to this questionnaire contains the alternate true/false versions of the questions.

Constructed variable name	Version 1A question # and T/F status	Version 2A question # and T/F status	Version 1B question # and T/F status	Version 2B question # and T/F status	Corresponding CogEcon 2008 (v1) Question #
fs1	B1-FALSE	B1-TRUE	B8-FALSE	B8-TRUE	18
fs2	B2-TRUE	B2-FALSE	B9-TRUE	B9-FALSE	19
fs3	B3-TRUE	B3-FALSE	B10-FALSE	B10-TRUE	20
fs4	B4-TRUE	B4-FALSE	B11-FALSE	B11-TRUE	23
fs5	B5-FALSE	B5-TRUE	B12-FALSE	B12-TRUE	25
fs6	B6-FALSE	B6-TRUE	B13-FALSE	B13-TRUE	26
fs7	B7-TRUE	B7-FALSE	B14-TRUE	B14-FALSE	27
fs8	B8-FALSE	B8-TRUE	B1-FALSE	B1-TRUE	28
fs9	B9-TRUE	B9-FALSE	B2-TRUE	B2-FALSE	29
fs10	B10-TRUE	B10-FALSE	B3-TRUE	B3-FALSE	31
fs11	B11-FALSE	B11-TRUE	B4-TRUE	B4-FALSE	33
fs12	B12-FALSE	B12-TRUE	B5-FALSE	B5-TRUE	34
fs13	B13-FALSE	B13-TRUE	B6-FALSE	B6-TRUE	40

<sup>19</sup> The version submitted is the same as the version assigned (*c3\_versassign*) in all but one case.

fs14	B14-TRUE	B14-FALSE	B7-TRUE	B7-FALSE	41
fs15	B15-TRUE	B15-FALSE	B16-TRUE	B16-FALSE	none
fs16	B16-TRUE	B16-FALSE	B15-TRUE	B15-FALSE	none

### Construction of the variables

Respondents are asked whether a statement is true or false on a 12 point scale based on their degree of certainty. True/false measures of financial sophistication are on a scale ranging from 100% to 50% confidence that the statement is “false,” and 50% to 100% confidence that the statement is “true.”

Each statement has two versions—one which is “true” and one which is “false.” Respondents receive one version of the question, and are therefore answering about whether that particular version of the statement is true or false.

For example, the “true” version of question fs1 reads: “Financially, investing in the stock market is *better* than buying lottery tickets.” The “false” version reads: “Financially, investing in the stock market is *no better* than buying lottery tickets.” The italics have been added to indicate the parts of these questions that differ. In either case, the respondent is instructed to decide whether the statement is “true” or “false,” and to indicate their confidence in this answer.

#### Version 1A:

**B1** Financially, investing in the stock market is no better than buying lottery tickets.

Most Likely False						Most Likely True					
Surely False						Guess False		Guess True		Surely True	
100%	90%	80%	70%	60%	50%	50%	60%	70%	80%	90%	100%

←Please Circle One Number→

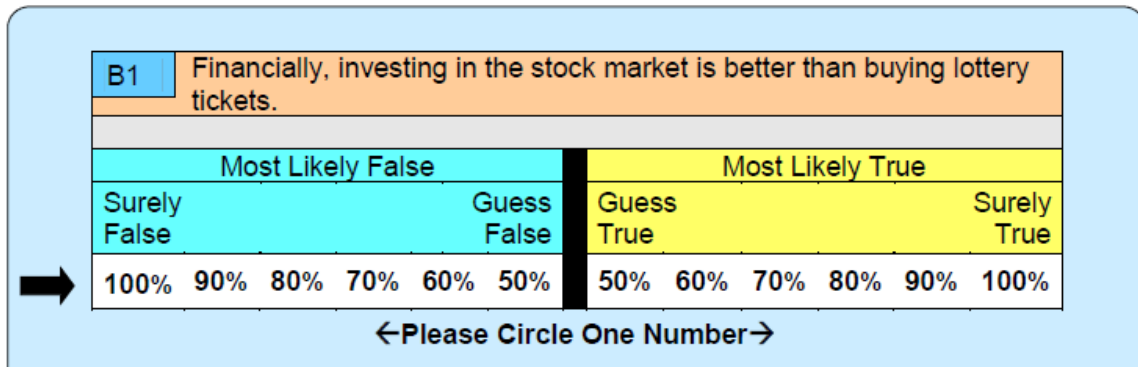
#### Version 1B:

**B8** Financially, investing in stocks is no better than buying lottery tickets.

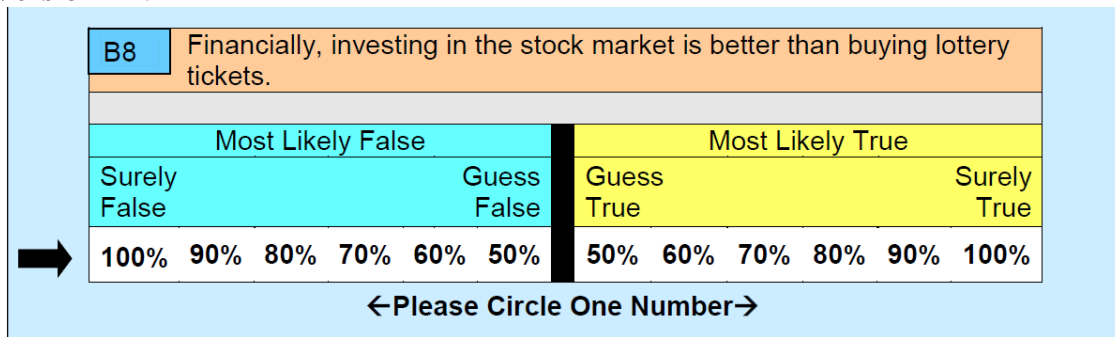
Most Likely False						Most Likely True					
Surely False						Guess False		Guess True		Surely True	
100%	90%	80%	70%	60%	50%	50%	60%	70%	80%	90%	100%

←Please Circle One Number→

**Version 2A:**



**Version 2B:**



We first convert all questions to the same scale as the “true” version. 100% Surely False become 0% Surely True. That is, the new scale can be interpreted as "0% Surely False" to "100% Surely True."

The table below shows how we converted the categorical responses to the assigned scores. In Column A are the raw responses for a question that is true. For example, 100% Surely False is complete confidence in the statement being incorrect. 100% Surely True is complete confidence in the statement being correct.

The scores are based on the assumption that respondents have a probability that the statement in the question is true in their mind, and they select their answer choice by rounding off their probability to the nearest choice on our 12-point scale (Column B). We can then construct intervals within which a respondent would round to each answer choice (Column C), and the point-value we assign is the midpoint of this interval (Column D).

They indicate their level of certainty by circling a percentage.

A. Categorical response		B. Respondent's belief that the statement is true (percent)		C. Implied probability range		D. Assigned score (midpoint of the implied probability range)
surely false	100%	0%	5%	0.00	0.05	0.025
	90%	5%	15%	0.05	0.15	0.1
	80%	15%	25%	0.15	0.25	0.2
	70%	25%	35%	0.25	0.35	0.3
	60%	35%	45%	0.35	0.45	0.4
guess false	50%	45%	50%	0.45	0.50	0.475
guess true	50%	50%	55%	0.50	0.55	0.525
	60%	55%	65%	0.55	0.65	0.6
	70%	65%	75%	0.65	0.75	0.7
	80%	75%	85%	0.75	0.85	0.8
	90%	85%	95%	0.85	0.95	0.9
surely true	100%	95%	100%	0.95	1.00	0.975

### Summary measures of financial sophistication

The following variables are constructed using these assigned midpoint scores:

- *c3\_mean\_score*: mean score on each question
- *c3\_fs\_score*: (sum of individual assigned scores for each question) / (number of questions answered)
- *c3\_num\_ans*: number of financial sophistication questions answered

## 2. Income and wealth variables

Respondents had two ways to answer questions about the value of income, assets and debts—either by providing an exact value or by selecting a range of values from the list on the “range card.” In the case of the mail questionnaire, respondents saw the “range card” option at the time they were asked for the exact value; web respondents were asked for a range only if they skipped the question asking for an exact value.

## **Total value: exact values and range**

Mail respondents were asked to either provide an exact value or to select a range of values from a list of ranges. In contrast, web respondents were first asked to provide an exact value. If they skipped the question they were subsequently asked to select a range of values from the ranges listed. In some instances, a range option was not given on the mail questionnaire but was used if the respondent did not provide a value in the web survey. These instances are marked in the range option tables as (web).

An exact value was imputed from each range using the mid-point of the range. For the uppermost bracket, 1.4 times the lower bound was used.<sup>20</sup>

## **Value assigned for ranges chosen**

All of the variables with the suffix `_val` are given in terms of monetary value. See Appendix 1 to see the different range options offered in the survey and the questions for each range was given.

## **3. Vehicle imputations**

Vehicle asset values were imputed as estimates of the resale value based on information about year, make and model. Respondents provided the following information about up to four cars. When possible, the values were found using Kelley Blue Book ([www.kbb.com](http://www.kbb.com)). The vehicles were looked up using respondent information about year, make and model, while assumptions were made about the car's trim and mileage, described below.

*Notes:* The public data files have the following variables about vehicle ownership and values:

- `c3_cars_imp`: sum of the imputed values for all cars (up to four)
- `c3_cars_numimp`: number of cars with imputed values
- `c3_cars_num` is the number of cars based on response to E23 & E24.

---

<sup>20</sup> Except the top bracket (for example, "More than \$1,000,000", which was imputed to be 1.4 times the lower bound of the bracket.

## Appendix 1: Range options provided throughout survey

Nine different range options were provided on the survey, some on both versions and others only on the web. This was done to more accurately capture the reasonable range of data values for each question.

Range options 1	Range selected	Value assigned
C2, C33, C42, D5, D7, D12, D14, D15, D15A-E, D16, D16A-E, D20, D21, D22, D23, D27, D28, D34(web), E10(web), E13(web), E15(web), F12(web)	\$0	0
	\$1-\$2,500	1250.5
	\$2,501-\$5,000	3750.5
	\$5,001-\$10,000	7500.5
	\$10,001-\$25,000	17500.5
	\$25,001-\$50,000	37500.5
	\$50,001-\$100,000	75000.5
	\$100,001-\$250,000	175000.5
	\$250,001-\$500,000	375000.5
	\$500,001-\$1,000,000	750000.5
	More than \$1,000,000	1400001.4
	Cannot provide a range	

Range options 2	Range selected	Value assigned
E19(web), E20(web)	\$0	0
	\$1-\$25	12.5
	\$26-\$50	38
	\$51-\$100	75.5
	\$101-\$150	125.5
	\$151-\$250	200.5
	\$251-\$350	300.5
	\$351-\$500	425.5
	\$501-\$750	625.5
	\$751-\$1,000	875.5
	\$1,001-\$1,500	1250.5
	More than \$1,500	2101.4
	Cannot provide a range	

Range options 3	Range selected	Value assigned
C16(web), C19(web), C21(web), C24(web),	\$0	0
	\$1-\$250	125.5
	\$251-\$500	375.5
	\$501-\$1,000	750.5
	\$1,001-\$1,500	1250.5



	\$1,501-\$2,500	2000.5
	\$2,501-\$3,500	3000.5
	\$3,501-\$5,000	4250.5
	\$5,001-\$7,500	6250.5
	\$7,501-\$10,000	8750.5
	\$10,001-\$15,000	1250.5
	More than \$15,000	21001.4
	Cannot provide a range	

<b>Range options 4</b>	<b>Range selected</b>	<b>Value assigned</b>
C5(web), C8(web), C11(web), C14(web),	\$0	0
	\$1-\$250	125.5
	\$251-\$500	375.5
	\$501-\$1000	750.5
	\$1001-\$1500	1250.5
	\$1501-\$2000	1750.5
	\$2001-\$2500	2250.5
	\$2501-\$3000	2750.5
	\$3001-\$3500	3250.5
	More than \$3,500	4901.4
	Cannot provide a range	

<b>Range options 5</b>	<b>Range selected</b>	<b>Value assigned</b>
D9(web)	\$0	0
	\$1-\$250	125.5
	\$251-\$500	375.5
	\$501-\$1000	750.5
	\$1001-\$1500	1250.5
	\$1501-\$2000	1750.5
	\$2001-\$3000	2500.5
	\$3001-\$5000	4000.5
	\$5001-\$10,000	7500.5
	More than \$10,000	14001.4
	Cannot provide a range	

<b>Range options 6</b>	<b>Range selected</b>	<b>Value assigned</b>
D31(web)	\$0	0
	\$1-\$500	250.5

	\$501-\$1000	750.5
	\$1001-\$2500	1750.5
	\$2501-\$5000	3750.5
	\$5001-\$10,000	7500.5
	\$10,001-\$25,000	17,500.5
	\$25,001-\$50,000	37,500.5
	\$50,001-\$100,000	75,000.5
	\$100,001-\$250,000	175,000.5
	More than \$250,000	350,001.4
	Cannot provide a range	

<b>Range options 7</b>	<b>Range selected</b>	<b>Value assigned</b>
E3(web), E4(web), E21(web),	\$0	0
	\$1-\$1000	500.5
	\$1001-\$2500	1750.5
	\$2501-\$5000	3750.5
	\$5001-\$10,000	7500.5
	\$10,001-\$20,000	15,000.5
	\$20,001-\$30,000	25,000.5
	\$30,001-\$40,000	35,000.5
	\$40,001-\$50,000	45,000.5
	More than \$50,000	70,001.4
	Cannot provide a range	

<b>Range options 8</b>	<b>Range selected</b>	<b>Value assigned</b>
E7(web), E8(web)	\$0	0
	\$1-\$1000	500.5
	\$1001-\$2500	1750.5
	\$2501-\$5000	3750.5
	\$5001-\$7500	6250.5
	\$7501-\$10,000	8250.5
	\$10,001-\$12,500	11250.5
	More than \$12,500	17501.4
	Cannot provide a range	

<b>Range options 9</b>	<b>Range selected</b>	<b>Value assigned</b>
E22(web)	\$0	0
	\$1-\$1000	500.5
	\$1001-\$2500	1750.5
	\$2501-\$5000	3750.5
	\$5001-\$10,000	7500.5

	\$10,001-\$25,000	17,500.5
	\$25,001-\$50,000	37,500.5
	\$50,001-\$100,000	75,000.5
	\$100,001-\$250,000	175,000.5
	More than \$250,000	350,001.4
	Cannot provide a range	

## Appendix 2: Steps to find car valuations

### For used cars between the years of 1992-2011:

1. Go to kbb.com
2. Hover over the “What’s my current car worth” and select “I plan to sell it myself”

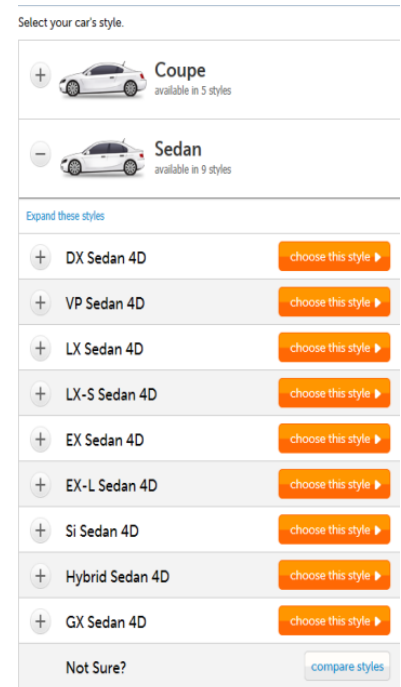


3. Enter the car’s *year*, *make*, *model* and *mileage* into the appropriate text boxes. Ex: 2007, Honda, Civic

A screenshot of the 'Get Your Blue Book Value' form. The form is titled 'Tell Us Which Car You Own' and includes a link for 'Valores en Español'. It contains four input fields: 'Year' (a dropdown menu), 'Make' (a text box with 'e.g. Honda' below it), 'Model' (a dropdown menu with 'e.g. Accord' below it), and 'Mileage' (a text box). An orange 'next' button is located at the bottom of the form.

- a. If make is missing, Google search can be used to find the correct make. Ex: “2009 F-150” will bring up “2009 Ford F-150”
  - b. If model is missing, we fill in the most common model by looking at the other cars. If this is inconclusive, Google search was used to find a popular model. Ex: typing in “2009 Ford” comes up with the first three results being “Ford F-150”
  - c. If year is missing, we use the year 2002 as an average estimate
  - d. Any missing data is labeled on the data “1=missing make, 2=missing model, 3=missing year, 13=missing make and year, 23=missing model and year)
  - e. No cars will be missing the make and model
4. The mileage was totaled using the following method: (2011-year of car)\*12,000
    - a. For 2011 and 2012 models, new prices are from autos.yahoo.com (see below)

5. The zip code used was 48109, in Ann Arbor, MI. This zip code was found to have neutral prices compared to other U.S. cities.<sup>22</sup>
6. The car's style is selected. If unspecified, the default model is used. Ex: Honda Civic default model is listed first, the DX Sedan 4D
  - a. The default trim includes standard features, such as automatic windows, air conditioning, and cruise control
  - b. We selected this trim because it includes the standard equipment. For example, if the EX-Sedan 4D were used instead, we would be assuming they paid extra for heated front seats, Bluetooth capabilities, and leather seats.
  - c. Because we know in fact that all models include all features included in the base model, we are making a safe estimate of the car's value rather than a guess at what they may have additionally purchased.



<sup>22</sup> See document comparing KBB values across different cities for additional details.

7. In the next menu, we choose to see value with standard equipment unless otherwise noted
  - a. Some vehicles' engine or drivetrain were modified, for example a few were listed as 4WD but default was FWD

### Get Your Blue Book Value

2009 Honda Civic  
DX Sedan 4D | Mileage: 24,000

Tell us your car's options or [see value with standard equipment](#)

Standard equipment pre-selected below

<p><b>Engine</b></p> <p><input checked="" type="checkbox"/> 4-Cyl. VTEC, 1.8 Liter</p>	<p><b>Transmission</b></p> <p><input type="checkbox"/> Automatic, 5-Spd w/Overdrive</p> <p><input checked="" type="checkbox"/> Manual, 5-Spd w/Overdrive</p>
<p><b>Drivetrain</b></p> <p><input checked="" type="checkbox"/> FWD</p>	<p><b>Comfort and Convenience</b></p> <p><input type="checkbox"/> Air Conditioning</p> <p><input checked="" type="checkbox"/> Power Windows</p> <p><input type="checkbox"/> Power Door Locks</p> <p><input type="checkbox"/> Cruise Control</p>
<p><b>Braking and Traction</b></p> <p><input checked="" type="checkbox"/> ABS (4-Wheel)</p>	<p><b>Safety and Security</b></p> <p><input checked="" type="checkbox"/> Dual Air Bags</p> <p><input checked="" type="checkbox"/> Side Air Bags</p> <p><input checked="" type="checkbox"/> FR Head Curtain Air Bags</p>
<p><b>Steering</b></p> <p><input checked="" type="checkbox"/> Power Steering</p> <p><input checked="" type="checkbox"/> Tilt &amp; Telescoping Wheel</p>	<p><b>Seats</b></p> <p><input type="checkbox"/> Power Seat</p> <p><input type="checkbox"/> Dual Power Seats</p> <p><input type="checkbox"/> Leather</p>
<p><b>Entertainment and Instrumentation</b></p> <p><input type="checkbox"/> AM/FM Stereo</p> <p><input type="checkbox"/> Cassette</p> <p><input type="checkbox"/> CD (Single Disc)</p> <p><input type="checkbox"/> CD (Multi Disc)</p> <p><input type="checkbox"/> MP3 (Single Disc)</p> <p><input type="checkbox"/> MP3 (Multi Disc)</p> <p><input type="checkbox"/> Premium Sound</p> <p><input type="checkbox"/> Navigation System</p> <p><input type="checkbox"/> DVD System</p>	<p><b>Roof and Glass</b></p> <p><input type="checkbox"/> Sun Roof</p> <p><input type="checkbox"/> Moon Roof</p>
<p><b>Lighting</b></p> <p><input checked="" type="checkbox"/> Daytime Running Lights</p>	<p><b>Exterior</b></p> <p><input type="checkbox"/> Rear Spoiler</p>
<p><b>Wheels and Tires</b></p> <p><input checked="" type="checkbox"/> Steel Wheels</p> <p><input type="checkbox"/> Alloy Wheels</p> <p><input type="checkbox"/> Premium Wheels</p> <p><input type="checkbox"/> Premium Wheels 15"+</p>	

8. We then note the fair, good, and excellent prices of the car using the information outlined above.

**2009 Honda Civic**

Style: DX Sedan 4D

edit options | change style
Mileage: 24000 change
Like this car

**Trade-In Value**  
when trading in at a dealership

Excellent ●  
**\$11,377**

Very Good ●  
**\$10,927**

Good ●  
**\$10,627**

Fair ●  
**\$9,452**

[Verify Condition](#)

**Private Party Value**  
when selling the car yourself

Shop for your next car [price a new car](#)

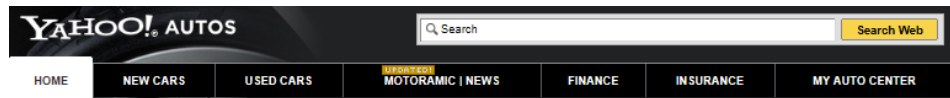
Instant Trade-In Offer  
powered by AutoTrader [get the offer](#)

**8.7**  
Out of 10 Own it? Love it? [review this car](#)

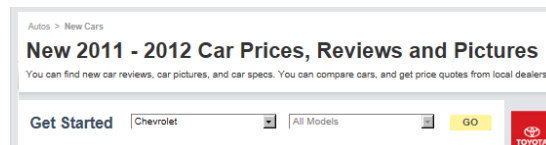
VALUES FOR MAR 16 2012 - MAR 22 2012 (updated weekly)

## For new cars (2011 & 2012):

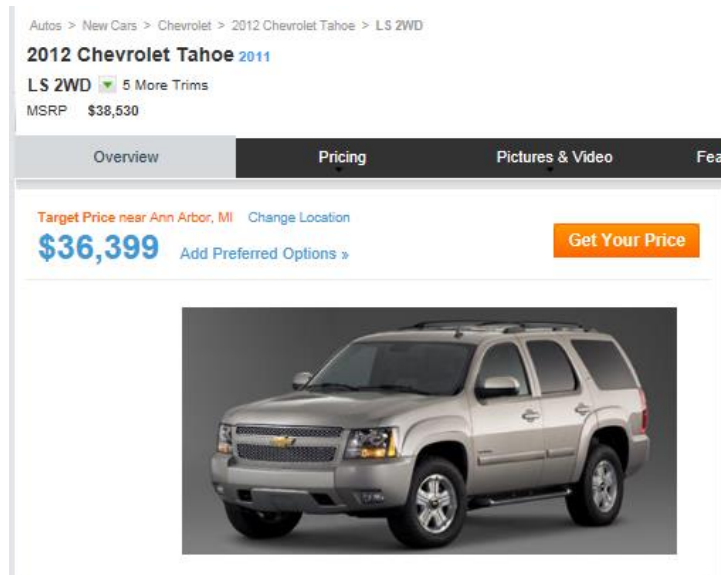
1. Go to autos.yahoo.com and click on the “New Cars” tab



2. Enter the make and model in the form boxes, ex: Chevrolet, Tahoe



3. Unless otherwise noted, use default MSRP price listed for the model equivalent to the kbb.com default model. If the model is 2011 instead of 2012, click on the blue “2011” link next to the vehicle name and record that price. Ex: MSRP price of 2012 Chevrolet Tahoe is \$38,530.



## **References:**

Fisher, G.G., Gideon, M., Hsu, J. & McFall, B.H. "Cognitive Economics Study: Development and Methodology," November 2011. University of Michigan.

Fisher, G.G., Gideon, M., Hsu, J., & McFall, B.H. "Cognitive Economics Study: Data Description," January 2012. University of Michigan.

Hsu, J., Fisher, G.G., & Willis, R.J. "Internet access and cognitive ability: Analysis of selectivity of internet interviews in the Cognitive Economics Survey," August 2008. Conference on Measurement and Experimentation with Internet Panels: Innovative features of Internet Interviewing. Zeist, Netherlands.

Hsu, J. & McFall, B.H. "Mode effects in mixed-mode surveys: Insight from the Cognitive Economics Study," September 2011. RAND/University of Michigan Workshop on Internet Interviewing and the Health and Retirement Study. Aspen, Colorado.