

**Cognitive Economics Study:  
Data Description**

Version 1.0  
January 2012

By: Gwenith Fisher, Mike Gideon, Joanne W. Hsu and Brooke Helppie McFall  
University of Michigan

## **Preface**

This document provides researchers with detailed information about using data from the CogEcon Public Core (Version 1.0) data release. General information about sample sizes and response rates, survey methodology and content development is available in the document *Cognitive Economics Study: Development and Methodology*.

This document combines content from the wave-specific documentation:

- Gwenith Fisher and Brooke Helppie McFall, “Cognitive Economics Study: 2008 Survey Data: Background, data description and usage.” September 30, 2009.
- Joanne W. Hsu and Brooke Helppie McFall, “Cognitive Economics Post-Crash Survey 2009: Survey and Data Documentation.” Last updated September 30, 2010.

Mike Gideon synthesized the material for the public use data documentation.

### CogEcon 2008 Project Team:

Robert J. Willis (PI); Miles Kimball; Tyler Shumway; Dan Benjamin; Matthew Shapiro; Gwenith Fisher; Brooke Helppie McFall; Joanne Hsu; Miles Putnam; Claudia Sahm; Kathy Ladronka; Andrew Hupp; Hueichun Peng; ZoAnne Blackburn

### CogEcon 2009 Project Team:

Robert J. Willis (PI); Miles Kimball; Tyler Shumway; Dan Benjamin; Matthew Shapiro; Gwenith Fisher; Brooke Helppie McFall; Joanne Hsu; Cynthia Doniger; Peter Hudomiet; Michael Gideon

## **Table of Contents**

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

Preface.....	2
I. Overview .....	4
A. Introduction .....	4
1. Conditions of use.....	4
2. Obtaining the data .....	4
3. Structure of data files .....	5
4. Merging CogEcon data with CogUSA data .....	5
4. Publications based on data .....	6
5. If you need to know more.....	6
B. Summary of data content .....	6
1. Variable naming conventions.....	6
2. Missing data and “don’t know” responses .....	8
3. Mode differences.....	10
4. Idiosyncrasies in CogEcon 2008 .....	12
5. Idiosyncrasies in CogEcon 2009 .....	13
II. Description of variables .....	15
A. Identification variables .....	15
B. Demographic variables from CogUSA.....	15
C. Sample/system variables.....	15
D. Constructed variables .....	16
1. Age at date of survey completion.....	16
2. Relationship status.....	17
3. Financial respondent status .....	18
E. CogEcon 2008 survey content .....	19
F. CogEcon 2009 survey content .....	25
III: Imputations and Calculated Variables .....	29
A. Financial Sophistication .....	29
1. Randomization.....	29
2. Construction of the variables.....	31
3. Summary measures of financial sophistication .....	32
B. Income and wealth variables .....	32
1. Total value: exact values and range .....	32
2. Magnitude of changes .....	34
C. Mapping CogEcon 2009 to CogEcon 2008.....	36

# I. Overview

## A. Introduction

This document provides detailed information about the CogEcon Public Core (Version 1.0) data release. The data were collected as part of the first two waves of the Cognitive Economics Study, CogEcon 2008 and CogEcon 2009, and have been processed and stored at the Survey Research Center of the University of Michigan.<sup>1</sup>

Additional documentation, including copies of the mail questionnaire, are available on our website (<http://cogecon.isr.umich.edu>). Please consult *Cognitive Economics Study: Development and Methodology* for details about methodology, sample sizes and response rates, and development of the two survey instruments.

### 1. Conditions of use

The CogEcon Public Core 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.<sup>2</sup>

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 to us about publications based on CogEcon data
- report apparent errors in the CogEcon data or documentation files via e-mail to [cogecon@isr.umich.edu](mailto:cogecon@isr.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.

### 2. Obtaining the data

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

1. Go to 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.

---

<sup>1</sup> CogEcon 2009 is also sometimes referred to as the CogEcon Post-Crash Survey, since it was fielded shortly after, and in response to, the start of the financial crisis in fall 2008.

<sup>2</sup> Respondents’ date of birth, date of survey completion, and geographic location were removed from both this and the CogUSA public use data. The date of survey completion was asked about in CogEcon 2008 question 160 and CogEcon 2009 question 125. Finally, the open-ended “text” components of the following questions have been removed:

CogEcon 2008: questions 104, 123, 129, 132, 138, 140, 147, 157;

CogEcon 2009: questions 21, 30, 62, 80, 81.

3. Log in to the HRS data file distribution system.<sup>3</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. A few variables associated with the CogEcon survey are 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

CogEcon Public Core (Version 1.0) data is in Stata format (.dta). There are 987 unstacked individual observations with 582 variables each. Researchers wanting to analyze the data as a panel will need to rename variables and reshape the data.

As explained in the *Development and Methodology*, 987 respondents submitted CogEcon 2008 or CogEcon 2009. Of these, 985 respondents submitted CogEcon 2008 and two submitted CogEcon 2009 but not CogEcon 2008.<sup>4</sup>

A mail questionnaire was “submitted” if it was returned to us. A web survey was “submitted” 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 six broad categories:

1. IDs & demographics (from CogUSA)
2. Sample/system variables
3. Constructed variables
4. CogEcon 2008 survey content
5. CogEcon 2009 survey content
6. Imputations and calculated variables

Variables are described in this document in the same order they appear in the dataset. Please make note of the differences between waves and across versions within the same wave.

### 4. Merging CogEcon data with CogUSA data

E-mail [mcardlelab@gmail.com](mailto:mcardlelab@gmail.com) or call 877-846-9551 to request access to CogUSA public use data. For information about CogUSA, please go their website:

[http://kiptron.usc.edu/studies/cogusa\\_login.php](http://kiptron.usc.edu/studies/cogusa_login.php)

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

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.

---

<sup>3</sup> The HRS data file distribution system is a secure website maintained on the premises of the Institute of Social Research (ISR) at the University of Michigan.

<sup>4</sup> Respondents are said to have submitted the survey if they give implied consent and answer at least one question.

## 4. Publications based on data

As part of the data registration process, you agree to include specified citations in and inform CogEcon of any papers, publications, or presentations based on CogEcon data. Please send a copy of any publications you produce based on CogEcon data, with a bibliographic reference, if appropriate, by email to [cogecon@isr.umich.edu](mailto:cogecon@isr.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.”

## 5. If you need to know more

This document serves as an overview of the publicly available data from CogEcon 2008 and CogEcon 2009. We will update the data and documentation when necessary. These changes will also be documented 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 [cogecon@isr.umich.edu](mailto:cogecon@isr.umich.edu). We will do our best to provide a timely response.

## B. Summary of data content

### 1. Variable naming conventions

Most variables have the prefix c1\_ or c2\_ to indicate association with CogEcon 2008 or CogEcon 2009, respectively. Identifiers and demographic variables do not have prefixes because they do not change across waves.

Variables from specific questions are identified first by the wave of the survey and then by their numbers on the mail questionnaire.<sup>5</sup> When multiple variables are associated with the same question, suffixes distinguish between them.

The general form of these variables is

cW_q#_suffix
--------------

with

7. W=1 for CogEcon 2008 and W=2 for CogEcon 2009;
8. # = question number in mail questionnaire; and,
9. suffix is defined according to the conventions in the table on the next page.

**Example 1:** Responses to question 17 in the CogEcon 2009 mail survey are in the variable *c2\_q17*.

**Example 2:** Question 94 from CogEcon 2008 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 *c1\_q94\_yn*. The value response variable has the suffix “\_val.” The variable name is therefore *c1\_q94\_val*. Please read section 10.1 for more information about how responses using “range cards” were converted into values.

**Example 3:** Questions 107-122 on CogEcon 2009 had different response scales in the mail questionnaire and the web survey. For question 107, *c2\_q107\_mail* is the variable for the mail version and *c2\_q107\_web* for the web version.

---

<sup>5</sup> Question numbers from the mail questionnaire are used because the questions on the web survey were not numbered

<b>Suffix</b>	<b>Meaning</b>
_yn	Yes/no
_val	Value
_i	Interest rate
_val_flag	Flag about the _val value—whether exact value, range, implied zero, missing, etc.
_mail, _web	Mail version (only for mail respondents), web version (only for web respondents)
_p, _p_	For questions specific to unmarried partners
_np	For questions specific to those not in unmarried partnerships (either single or married)
_s, _s_	For questions specific to spouses
_dir	Direction of change
_path	Direction of change and method of reporting magnitude of change
_inc_pct, _dec_pct	Percent increase, percent decrease
_inc_val, _dec_val	Value of increase, value of decrease
_in_, _out_	Transfer money into stocks, out of stocks
_week, _2weeks, _month, _year	The respondent can choose the length of time under consideration
_buy_, _sell_	Bought stocks, sold stocks

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

### **2.1 Missing response codes within variable**

For most variables, this information is recorded using different “missing” response codes within each variable. These response codes can be seen by tabulating a variable, while including missing values; for example, by typing “tab *varname*, mi” in Stata.



Code	Reason for missing
.	Survey not submitted
.d	Don't know
.e	Respondent was skipped because of a logic error in survey instrument
.r	Refused
.x	Does not apply (specifics depend on variable)
.u	Reference person has no partner (for partner variables)
.z	Quit survey (mid survey) before responding to this question

**NOTE:**

- There are slight differences between the missing values for CogEcon 2008 and CogEcon 2009. In particular, “.u” is not used for the CogEcon 2008 variables, but this information is captured by “.x.”
- The missing values ".z" and ".r" both indicate that a respondent refused to answer the question. However, “.z” is used when that the respondent had altogether stopped completing the survey prior to this question.
- For questions 50 and 59, responses were posed in a "check all that apply" format. For these cases a refusal cannot be distinguished from "does not apply to me." To avoid confusion, these have been coded as "."

**2.2 Flag indicating type of response or missing**

Variables with the suffix *\_val* are dealt with differently. Each has a flag variable associated with it, *cW\_q#\_val\_flag*, which contains categories for missing values as well as both range vs. exact values. This set of variables includes questions about income, food consumption, financial and non-financial assets and debts. These variables contain reported monetary values.

Respondents known to have a positive amount/value for a particular item are flagged with 1, 2 or 3, with 1 indicating that the respondent gave an exact value, 2 indicating that the respondent gave a range, and 3 indicating that the respondent should have provided a value or range, but left it blank. Respondents who we know have zero for a particular item, but who might not have explicitly written this, are flagged with 5. And when we do not know whether the respondent has a positive value we flag with 8 or 9, depending on the reason this is unknown. The table below describes the flags in more detail.

cW_q#_val_flag		
Label		Description
ExactVal	1	Respondent reported an exact value for the amount.

Range	2	Respondent reported a range of values for the amount. For internet surveys this implies that the respondent skipped the question asking for the exact value.
Missing	3	The respondent (i) should have provided a value or range, (ii) did not provide value or range.
ImpliedZero	5	Respondent answered a question implying that the value is zero. This is either the Y/N part of the question or from the skip logic.
NotApplicable	7	Not asked to certain individuals either because single (e.g. <i>c1_q70_val</i> about spouse/partner's earnings) or because not assigned the questions because of the randomization ( <i>c1_q154_min</i> or <i>c1_q154_max</i> ).
ErrorSkip	8	We do not know whether the respondent has a positive value for this item because of a problem with the survey instrument.
RespSkip	9	We do not know whether the respondent has a positive 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 (*c1\_q42*), food at home (*c1\_q52*), and food away from home (*c1\_q53*)—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.
- *ErrorSkip*:
  - Question 89: mail respondents were inadvertently instructed to "skip to Q91" if they answered "no" on Q86.
  - Questions 104, 123 and 129: added to net revision 4, so respondents with earlier versions were not asked these questions.

**3. Mode differences**

Most questions on the mail and web instruments were identical, but there are some important differences.

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

*Relevant for both CogEcon 2008 and CogEcon 2009.*

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."

Married respondents, and those who reported that they were both (a) in “marriage-like” relationships 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.

Reminders of these instructions were placed at the top of pages in the mail questionnaire, whereas the correct wording was “piped” into the web instrument.

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

*Relevant for both CogEcon 2008 and CogEcon 2009.*

For several questions about income, assets and debts, mail respondents had the option of providing an exact value or choosing a range of values from a list of ranges. In contrast, the web instrument first asked for an exact value. A list of ranges was provided if the respondent did not respond with an exact value.

Not surprisingly, mail respondents were much more likely to choose a range rather than provide an exact value. Because of additional measurement error that could be introduced, questions about income, assets and debts—anything with the suffix *\_val*—should be used with this difference in mind.

Section 8.1 for additional details about how we dealt with values and ranges.

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

*Relevant for both CogEcon 2008 and CogEcon 2009.*

For questions that did not require respondents to type a value, 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>6</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.

### **3.4 Scale differences**

There were also several questions in CogEcon 2008 and CogEcon 2009 that had different response scales in the mail and web versions:

- *CogEcon 2008*: Credit card interest rate question had different response categories in the web and mail.
- *CogEcon 2009*: Questions 107-122 offered a 5-point agree-disagree scale in the mail version of the survey, while offering a 6-point scale in the web version. These variables are stored separately in the dataset. For example, responses to question 107 are stored in *c2\_q107\_mail* and *c2\_q107\_web*.

---

<sup>6</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.

## 4. Idiosyncrasies in CogEcon 2008

### 4.1 Versions/Randomizations

Each respondent was randomly assigned one of two versions of the survey. The two versions were exactly the same except for the two sets of questions that were randomized across respondents.

1. Questions 17-41: True/False financial sophistication section.

Each of the 25 true/false financial sophistication questions had a true version and a false version. Respondents received one of the two versions for each question. These differences are explained in Section 7.1 about the financial sophistication measures and in their respective survey content sections. The variable *c1\_fs\_random* indicates whether respondent had version 1 or 2.

2. Question 154's min/max wording. Version 1 respondents were asked to provide a maximum value and version 2 respondents were asked to provide a minimum value. The variable *c1\_q154\_random* indicates whether respondent had version 1 or 2.

### 4.2 Revisions

There were two substantive revisions of the web survey while it was in the field, and the mail questionnaire was altered for Release 2. The changes and comparability implications are detailed below.

*Mail Revisions:* See *c1\_mail\_revision* variable.

*Mail Revision 0:* Instrument approved by the IRB on March 5, 2008, fielded immediately thereafter. Release 1 mail respondents used this questionnaire.

*Mail Revision 1:* Amended instrument approved July 7, 2008, used for Release 2. Changes:

- 1) Fixed skip pattern errors in questions 76 and 86. The first of these originally skipped some respondents to Q78, and was amended to skip to Q81. This should not have resulted in any loss of data. The second of these erroneously instructed respondents who said they did not "own any other property, land or real estate" (most respondents fell into this category) on question 86 to skip to 91. They should have only been skipped to question 89. This correction was made for Release 2. *Questions 89 and 90 were erroneously skipped by most mail respondents in Release 1. Use caution in analysis of these variables.*
- 2) Added questions 158-160. These asked for the date, time, and how long the questionnaire took to complete.
- 3) Added a date indicator to the front cover (in the box covered by the label) to differentiate between this revision and the prior instrument.

*Net Revisions:* See *c1\_net\_revision* variable. A respondent who had started but not completed a survey before a revision was posted was given two choices: either to start

over or to resume from where s/he left off. If that respondent resumed from where s/he left off, the old instrument was used. If s/he opted to start over, the revised instrument was presented. The *cl\_net\_revision* variable indicates which version was used.

*Net Revision 2:* 101 cases. Instrument approved end of January and fielded in February, 2008.

*Net Revision 3:* 287 cases. Instrument posted February 26, 2008. Missing reminder screens were added, and minor skip logic errors were amended in this revision.

*Net Revision 4:* 119 cases. Changes posted March 14, 2008, including:

- Q52, there was a small change in wording
- Q104 and Q104\_text were added
- Q123 and Q123\_text were added
- In Q124, there was a slight wording change to include bonds
- Q129 and Q129\_text were added
- Changes to Q130 and addition of Q130\_ret
- In Q131 and Q132, there was a small change in wording to be more specific about including “other” assets.
- Q132\_text was added
- Q136-140 were re-worded to ask about both respondents *and* their spouses or partners.

*Net Revision 5:* 6 cases. Posted April 21, 2008. Small logic fix to show reminder screen where none had been shown.

*Net Revision 6:* 1 case. Posted May 7, 2008. Slight change to question 66. Logic streamlining—two questions were combined into one.

## **5. Idiosyncrasies in CogEcon 2009**

### **5.1 Versions/Randomizations**

Participants in CogEcon 2009 were assigned the same version number of the survey as they had completed in CogEcon 2008. The version number affected only which version of the 12-point scale financial sophistication questions (questions 2-14) respondents were asked.

### **5.2 Revisions and errors**

- *Logic error:* Question 26 in the mail questionnaire was originally fielded with faulty logic: a respondent who indicated that his or her significant other was employed in 2008 was falsely instructed to skip the questions asking for details about this employment. The error was fixed in subsequent mailings, but Survey Research Operations telephoned 64 respondents who had returned the first version of the survey and indicated that their significant other was working, but skipped

question 27-29. Questions 27-29 were asked of them. Three respondents who would have been called were not eligible for call-back, because of an agreement with the CogUSA study that we would not contact respondents within 30 days of a study participant's scheduled CogUSA study contact date. The web version of the survey did not contain this error.

- *Questions 42, 46, and 67:* in the mail mode, a fourth answer choice, “N/A, have not had” was given, so respondents who did not have debt could check that instead of “No change.” In the web instrument, logic skipped respondents who did not have debt past the questions about changes in these debts.
- *Question 59 (Percent change in stock):* Until June 9, 2009, web respondents who reported buying stock were given the opportunity to report the value of purchases as a percent. Mail respondents were not given this option. This option was removed on June 9 for web respondents.

## II. Description of variables

### A. Identification variables

Every respondent has three identification numbers. Each of these are 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 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

### B. Demographic variables from CogUSA

Several demographic variables from CogUSA are included in the CogEcon public release data. These include: *birthyr*; *gender*; *race*; *racem1*; *racem2*; *female*; *hispanic*; *white*; *black*; *educ*; *hs\_deg*; *coll\_deg*.

*race1m* and *race2m* distinguish between the first and second “race” categories mentioned by the respondent. The other variable labels should be self-explanatory.

### C. Sample/system variables

Variable	Description
<i>wave_ce</i>	Indicates which waves of CogEcon the respondent completed
<i>cW_invite</i>	Indicator for whether or not a respondent was invited for that wave (W=1 or 2) of the CogEcon survey
<i>cW_result</i>	Survey outcome result for those who were invited =1 for completed submissions =2 for partially completed submissions
<i>cW_modeassign</i>	Assigned mode =0 for web, =1 for mail
<i>cW_modesubmit</i>	Submission mode

	=1 if mail survey =0 if web survey Two individuals requested to be interviewed by phone
cW_lastQ	Last question completed <sup>7</sup>
cW_compmo;	Month in which survey was submitted
cW_compyr	Year in which survey was submitted
cW_fs_random	Gives the version of financial sophistication questions, see Section 8.1 for more details about these questions

Please note the distinction between the mode that someone was invited to, and the mode submitted. In both CogEcon 2008 and CogEcon 2009, web invitees were given the option to “switch modes” with the third reminder letter.

There are also variables associated with one wave but not the other:

- *c1\_release*: CogEcon 2008 variable, equals 1 if respondent was in release 1 and 2 if respondent was in release 2.
- *c1\_mail\_revision*: CogEcon 2008 variable, indicates revision for mail respondents.
- *c1\_net\_revision*: CogEcon 2008 variable, indicates revision for web respondents.
- *c2\_invite\_flag*: gives the reason not invited to CogEcon 2009 (missing if invited). Don't have this for CogEcon 2008 because every respondent in the sample was invited.

## D. Constructed variables

The dataset includes constructed variables for age at the time of the survey (*cW\_age\**), relationship status (*cW\_relstat*), an indicator for whether the individual is a household's “financial respondent” (*cW\_finr\**) as well as measures of “financial sophistication” that were derived using the set of true/false statements. The financial sophistication measures are described in section 7.1. All others are described below.

### 1. Age at date of survey completion

The variable *cW\_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. When month is known but day is not, we calculated the age as if the day is the 15<sup>th</sup>. Otherwise,

<sup>7</sup> This variable is most useful for respondents who submitted a partially completed survey, since it could help distinguish respondents who answered most of the survey and those who completed very little of it.



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

For web respondents, date completed is the date the survey was submitted. For mail respondents, date completed is either the self-reported date completed or is imputed using the date SRO received the questionnaire. CogEcon 2008 only asked Release 2 mail respondents for the date completed (question 160), whereas in CogEcon 2009 all respondents were asked for the date (question 125). 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). This uses *c1\_q160* and *c2\_q125* for CogEcon 2008 and CogEcon 2009, respectively.

### Other age-related variables:

- *cW\_age\_y*: age in years, as an integer.
- *cW\_age\_m*: age in number of months.
- *cW\_age\_flag*: flag variable that indicates whether the birthdate or survey date was missing and needed to be imputed. The table below gives the distribution for this variable. The large number of imputations for CogEcon 2008 date completed is because the release 1 mail respondents were not explicitly asked when they completed the survey.

<i>cW_age_flag</i>	What was imputed?	CogEcon 2008 (W=1)	CogEcon 2009 (W=2)
1	Neither	696 (70.7%)	834 (98.5%)
2	Birth date	6 (0.6%)	7 (0.8%)
3	Date completed	282 (28.6%)	6 (0.7%)
4	Both dates imputed	1 (0.1%)	0 (0%)

## 2. Relationship status

The categorical variable for relationship status, *c1\_relstat*, is constructed using

- i. Question 75 for mail respondents (who answered the question).
- ii. Questions 43-49 for web respondents (who were not asked question 75) and mail respondents who skipped it.
- iii. Missing if neither (i.) nor (ii.) is satisfied. (30 obs)

For CogEcon 2009, the variable *c2\_relstat* is equal to *c1\_relstat* if there is no indication that relationship status changed.

Code	Relationship status	CogEcon 2008	CogEcon 2009
1	Married	662 (69.3%)	568 (67.1%)
2	Partner with financial future	25 (2.6%)	24 (2.8%)
3	Partner without financial future	16 (1.7%)	12 (1.4%)
4	Single	252 (26.4%)	243 (28.7%)
Total		955	847

### 3. Financial respondent status

CogEcon asks asset and debt questions of all respondents. For household-level analysis, we recommend using the values given by the household's designated financial respondent.

The variables *cW\_finr* and *cW\_finr\_alt* are set equal to 1 if the respondent is designated as the household's financial respondent. *cW\_finr\_how* and *cW\_finr\_alt\_how* tell you how the financial respondent was chosen, as is described below:

- (1) If a respondent was single, partnered but not planning a financial future with his/her partner, or was the only person in the family to complete a questionnaire, that respondent is automatically designated as the financial respondent.
- (2) If two respondents from a household submitted questionnaires, however, we first tried to assign a financial respondent based on a question very similar to one used in the HRS for the same purpose: "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." If there was strong agreement that one partner/spouse knew the most -- that is, if one said "me" and the other said "my spouse/partner"-- the financial respondent is the respondent who was said to be more knowledgeable. If there was weak agreement, in that one partner indicated one person and the other partner's answer did not conflict (for example, if one said "me" and the other said "both"), the partner who was weakly indicated was chosen as the financial respondent.
- (3) In cases where a financial respondent was still not clearly indicated, the answer to "Who makes the decisions about how to save for retirement and other large expenses?" was used in the same manner.
- (4) In cases where a financial respondent was still not clearly indicated, we chose the respondent who reported using more objective data sources (tax returns, account statements, or financial software) while completing the survey.
- (5) Lastly, if there was still no financial respondent, *cW\_finr\_alt* randomly assigned one from the members of the couple. By contrast, *cW\_finr* incorporates two more steps to try to choose a financial respondent before finally randomly choosing for the remaining couples: first, the respondent who claims to have a greater number of the following assets: primary home (q76), tax-advantaged retirement account

(q94), or checking account (q113) was chosen to be the financial respondent; second, if the first was a tie, the respondent who refused to answer fewer questions of the set: total hh income (q42), primary home value (q77\_val), retirement assets (q94\_val), and checking/CDs accounts (q113\_val) was chosen to be the financial respondent.

For CogEcon 2009, *c2\_finr* is based on the CogEcon 2008 variable *c1\_finr* unless the designated financial respondent did not complete the 2009 survey. In that case, the person who did complete CogEcon 2009 was chosen as the financial respondent.

For participants who only completed the CogEcon 2009 survey, we followed the same procedure for selecting the financial respondent that we used to select the CogEcon 2008 financial respondent.

## **E. CogEcon 2008 survey content**

Below is a list of topics covered by the survey, including question numbers. For the most part, the topics are listed in the same order as the question numbers.

### **E-mail usage: q1-3**

Variables associated with questions 1 and 2 (*c1\_q1* and *c1\_q2*) are dummy variables for whether the e-mail address was provided. Missing values for *c1\_q2* and *c1\_q3* have been recoded to “.x” if in question 1 the respondent indicated he/she “do not have” a primary e-mail address.

### **Risk tolerance: q4**

This question asks for a self-assessment of risk tolerance.

### **How much of current financial \_ would a financial advisor recommend changing: q5**

### **Hypothetical annuity choice: q6-q7**

### **Expected investment returns for oneself vs. a typical individual: q8-12**

### **Self-assessed knowledge of stock market, math, day-to-day finances: q10-12**

### **Financial sophistication: q13-16**

Questions 17-41 make up the financial sophistication battery, which is used to construct an index of financial sophistication. Several questions in the survey are designed to measure behaviors related to these topics.

*Please see Section 7.1 for details about both the CogEcon 2008 and CogEcon 2009 financial sophistication questions.*

### **Household income: q42**

### **Relationship status, # of children: q43-50**

*Note:* The variable *c1\_relstat* is a categorical variable that provides a simple measure of relationship status.

### **Food expenditure: q52-53**

*Comparability details:* On question 52, the range option labeled “None” in the mail instrument actually read “Don’t spend money on food” prior to net\_revision 4 in the web instrument, at which point wording was changed to “None.”

### **Household financial decision making: q51, q54-55**

These questions were included for comparability with HRS. These questions are used to determine who in a couple is the financial respondent (*c1\_finr* and *c1\_finr\_alt*); see Section 6.3 for more detail.

### **Economics background: q56-58**

#### **Social Security and pension income: q59-63**

*Note:* The wording of these questions asks about the respondent and the respondent’s spouse/partner in the same question; they do not provide information about who receives payments. For example, question 59 asks “Do you or your spouse/partner currently receive monthly Social Security checks?” and we do not know which member is receiving these checks.

### **Investment choice in defined benefit retirement plan: q64**

#### **Employer-sponsored retirement planning seminars: q65-66**

*Comparability details:* Question 65 was the same for all respondents. Question 66 was slightly different across the mail and web survey. More specifically, web respondents who were partnered (but not married) were asked two separate questions instead of the one question asked of everyone else. Responses have been combined to create the variable *c1\_q66*, with the categories:

- 1: both have attended
- 2: I have attended
- 3: spouse has attended
- 4: partner has attended
- 5: neither of us has attended

There are 12 observations for which answers to two questions are combined in this variable. Note that there is one observation with inconsistent answers across the two questions —one saying no, the other saying yes for both R and R’s partner.

### **Employment: q67-74**

*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 75) 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.

## **Marital and partnership status: q75**

*Comparability details:* This question was only asked on the mail survey and responses are in the variable *c1\_q75\_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 (76 through 140) were asked with respect to the respondent only, or with respect to the respondent and his/her significant other.

Question 75 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.

## **Real estate assets and debts: q76-88**

*Comparability details:*

1. Web respondents were asked a series of two separate questions instead of the one question asked in the mail survey. Question 76 is a yes/no question asking whether the respondent (or spouse/partner) owns their primary home. Responses to this general question are contained in the variable *c1\_q76*.

Twelve web respondents who were partnered (but not married) received questions with different wording and skip patterns. They were asked: “Do you or your partner own your own home?” This case is “*c1\_q76\_p*” in the data. Answer options are:

1. We own a home jointly
2. Only my partner owns a home
3. Only I own a home
4. We each own a separate home
5. Neither of us owns a home

Answer options 1, 3, or 4 lead to the same question series as in the mail questionnaire. Option 2 leads to a question series starting from “What would your partner’s home be worth if sold today” and through the loans and interest rate on this home. Option 4 asks the same series as in the mail questionnaire, *plus* the questions on the partner’s home, as in option 2.

Questions about the partner’s home are named *c1\_q77\_p\** - *c1\_q80\_p\**.

In other words, the variable *c1\_q77\_p\_val* is only available if

- the mode of participation was web,
- the respondent had a partner with whom he/she was planning a financial future,

- this partner owned a home, and
- this home was either in addition to the respondent's home (as in *c1\_q76\_p=4*), or in a case in which the respondent does not own a home (as in *c1\_q76\_p=2*).

There were only two responses to question 76 which led to these partner home questions being asked, so this is a very minor comparability issue.

2. There were 13 people who entered both a value and a range for the value of their home. We think these respondents reported an exact value in thousands of dollars (\$100 for \$100,000), later realized we were asking for thousands, and hit back until they got to the range card. Twelve of those entered ranges that included their original value\*\$1000, and were re-coded as having entered their original value\*\$1000. One person entered a range that included his or her value, so we kept the original value (didn't multiply by \$1000). Additionally, one person reported \$0 for the exact value then reported a range. This person's value for *c1\_q77\_val* comes from the reported range.
3. The yes/no portion of question 80 was not asked in the web or data entry programs. Instead "do not have..." was an answer option for the interest rate, and contained in variable *c1\_q80\_i*.

### **Vehicle assets and loans: q89-91**

*Comparability details:* Questions 89 and 90 were skipped by most Release 1 mail respondents due to incorrect skip logic in the questionnaire (respondents were instructed to "skip to Q91" if "no" on question 86). The skip logic was removed from the mail data entry program so all answers (even though they shouldn't have been answered, in most cases) would be coded. Note that presence of a loan on a vehicle indicates ownership of a vehicle, but "no" on question 91 doesn't indicate whether someone owns a vehicle or not.

### **Credit card and other debts: q92-93**

*Comparability details:* Question 92 asked for the interest rate on credit cards. For question 92, the web version of the survey and the mail survey presented different categories for the interest rate. If using these data, consider both *c1\_q92\_i* (for web respondents) and *c1\_q92\_i\_mail* (for mail respondents), because the answer options do not match between the mail and web versions of the question.

### **Tax-advantaged retirement assets: q94-104**

*Comparability details:* Question 104 was first added to the internet instrument at net revision 4, but was present in all mail surveys.

Initial internet data inspection in February suggested that questions 95-103 were not an exhaustive list for these types of assets. Specifically, some respondents who appeared to be answering many of these questions very carefully indicate total values of retirement assets (94) greater than the sum of 95-103. We added a "catch-all" question, 104 (with a text box), to find the value of any other

retirement assets not listed, and also an open-ended text box with a prompt “Please describe” to understand which other types of assets are common.

### **Tax-advantaged educational savings accounts: q105-106**

*Comparability details:* Skip logic in mail version was missing after a “no” to question 105. “No” to question 106 should be recoded as missing if 105 was also answered “no.”

### **Wills and life insurance: q107-112**

*Comparability details:* Question 110 was slightly different in the mail and web versions life insurance for a spouse or partner’s death. Web respondents with a partner were asked specifically about benefits from their partner’s death. Married web respondents were asked specifically about benefits from their spouse’s death. These have been combined to form a single variable, *c1\_q110\_yn\_web*. All mail respondents (including singles) were asked about “Your spouse’s/partner’s death” This is in *c1\_q110\_yn\_mail*.

### **Liquid assets, farm and business ownership: q113-115**

#### **Assets outside of retirement accounts: q116-129**

*Comparability details:* Question 123 was first added to the web instrument at net revision 4 but was present in all mail surveys. Initial data inspection in February suggested that the table did not comprise an exhaustive list for these types of assets. We added a “catch-all” question, 123, to find the value of any mutual fund assets not listed, and also an open-ended text box with a prompt “Please describe” to understand which other types are common. The words “or bonds” were first added to question 124, and question 129 was added to the instrument to capture assets held in bonds of individual companies, at net\_revision 4.

### **Diversification: q130**

*Comparability details:* There question(s) on the web version was changed while the survey was in the field, and neither version is exactly comparable to the mail version.

Mail:

- All respondents were asked question 130: “In all, in *how many different companies* do you (or your spouse/partner) hold stock, outside of mutual funds?”

Web:

- Net revisions 2 and 3: same question as in the mail survey, but only asked to respondents who answered “Yes” to question 124: “Not including what is in retirement accounts or educational savings accounts do you [and your spouse/partner] have any stocks of individual companies outside of mutual funds?”
- Net revision 4: Another question (variable *c1\_q130\_ret*) was added; this was asked about stock holdings within tax-advantaged retirement

accounts. Both questions were only asked of respondents who said “Yes” to 124.

### **Other assets: q131-132**

*Comparability details:* For net revision 4, the wording of these questions was altered, and a text box was provided in question 132. Question 131 was altered from “and collectibles?” to “, collectibles...” to make clear that other assets should be included here. Question 132 was similarly changed.

Respondents who answered ‘no’ to question 124 about ownership of stocks or bonds of individual companies should have been directed to 130. A skip logic error on the web survey directed these respondents to 133. This caused 310 (60.31%) of the 514 web respondents to inadvertently skip questions 130-132.

### **Diversification: q133-135**

*Comparability details:* Questions 134 and 135 were asked of all mail respondents who were currently working, or whose significant others were currently working. For web respondents, question 134 was only asked if they answered “yes” to 133. Similarly, question 135 was only asked if respondents answered “yes” to 134.

### **Stock trading frequency and reasons: q136-140**

*Comparability details:* Mail respondents who answered “Never” to question 136 or 138 were asked to skip to 141.

Question 137: Web respondents were offered “never” as an answer option, and respondents who chose this option were skipped ahead to Q141. For mail respondents “never” can be inferred from their response to Q136.

net\_revision 4 made two changes to match the mail questionnaire:

- (i.) For Questions 136-140, added “or your spouse” /“or your partner” to the question roots.
- (ii.) Added an answer option to question 136: “I don’t know. My spouse/partner handles these transactions.”

### **Asset Allocation q141-143**

### **Credit Distress: q144-146**

### **Use of financial records to complete questionnaire: q147**

### **Risk tolerance: q148-152**

Questions 148-152 can be used to estimate risk preference parameters. The answer to 148 can be used to create the first “bound” on measures of risk aversion, and determines whether respondents should then answer questions 149-150 or 151-152 to determine the other bound.

*Note:* Mail respondents who answered YES to question 150 were inadvertently asked 151-152 while they should have been directed to skip to 153. All other



respondents had the correct skip logic. Any such responses have been ignored and anyone presented questions 149-150 is assumed to not have been asked 151-152.

### **Decisions about long-term care: q153-155**

*Details:* Respondents were randomly assigned one of two versions of question 154. The variable *c1\_q154\_random* indicates whether the respondent was asked about the *minimum* bequest she would be willing to leave (*c1\_q154\_random*=2), or for the *maximum* willingness to pay for care (when *c1\_q154\_random*=1). The variable *c1\_q154\_max\_val* contains responses from the *maximum* formulation of the question, while *c1\_q154\_min\_val* contains responses from the *minimum* formulation.<sup>8</sup>

*Comparability details:* For question 154 web respondents who did not answer the question were given a reminder screen and asked to select a range of values from a list of ranges. There was also an option for “Don’t know.” This was selected by 20 respondents, and the value was recoded as missing.

### **Help completing questionnaire, time and day completed: q156-160**

Questions 156 and 157 ask whether respondents’ received help completing the questionnaire. Questions 158-160 ask about the date and time of completion, and are only asked to Release 2 mail respondents.

### **Section H (Release 1 web mode only)**

This section was not administered in the mail survey. Data derived from these questions are not yet publicly available.

## **F. CogEcon 2009 survey content**

### **Risk tolerance: q1, q102-106, q107-111**

Question 1 and questions 102-106 were asked in CogEcon 2008. Questions 102-106 have also been asked in the Health and Retirement Study.

### **Financial sophistication and confidence in knowledge: q2-14**

*Please see Section 7.1 for details about both the CogEcon 2008 and CogEcon 2009 financial sophistication questions.*

---

<sup>8</sup> *Note about Q154:* There was an error in the program used to enter data from mail questionnaires such that the responses were entered incorrectly for the twelve respondents who (a.) submitted a different version from what had been assigned and pre-loaded in the system, and (b.) provided a range rather than an exact value. These responses have been recoded and we do not think this programming error affected data entry for any other questions.

### **Income, labor supply and earnings: q15-37**

Questions 15-20 and 26-29 are repeated from CogEcon 2008, and cover household income and the labor supply of respondents and their significant others.

Questions 21-25 and 30-37 were added to the survey to gather information about changes in retirement status, employment status, and expected age of retirement.

*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.

These questions ask about “during 2008” whereas in the 2008 survey the questions ask about “last year.”

### **Real estate ownership, value, and changes: q38-47**

This section repeats questions from CogEcon 2008 about real estate assets and debts. New questions were added about changes in these assets and debts.

### **Tax-advantaged retirement assets and changes: q48-54**

This section repeats questions about tax-advantaged retirement assets, and asks about stock trades, account balances, contributions and withdrawals had changed since July 2008.

*Comparability details:* For question 50, mail respondents were asked a Yes/No question, while web respondents were asked to “check all that apply.” Therefore, web respondents who would answer “no” (if it were offered as an option) cannot be distinguished from those who refuse to answer the question .”

The categories for c2\_q50\_in\_yn and c2\_q50\_out\_yn are in the table below. For consistency, the refusals for mail respondents have been coded “no/DK (mail).” The category 3 for web respondents is comparable to the sum of 2 and 4 for mail respondents.

<b>c2_q50_*_yn</b>	<b>label</b>
1	yes
2	no (mail)
3	no/DK (web)
4	no/DK (mail)

## **Non-tax-advantaged assets, holdings and changes: q55-60**

This section repeats questions about financial assets held outside of tax-advantaged retirement accounts, and asks about changes in the value of these assets. Additionally, questions 59 and 60 ask whether respondents have bought or sold stocks or mutual funds.

*Comparability details for question 59:*

- *Percent change in stock:* Until June 9, 2009, web respondents who reported buying stock were given the opportunity to report the value of purchases as a percent. Mail respondents were not given this option. This option was removed on June 9 for the web respondents.
- Mail respondents were asked a Yes/No question, while web respondents were asked to “check all that apply.” Therefore, skips cannot be distinguished from “no.” In this question the refusals for mail respondents are also considered “no/DK.” The categories for c2\_q59\_buy\_yn and c2\_q59\_sell\_yn are the same as for c2\_q50\_in\_yn and c2\_q50\_out\_yn in the section above.

## **Stock trading behavior: q61-62 (also see 50 and 59-60)**

Questions 61 and 62 ask about changes in investment allocation in retirement accounts and are repeated from CogEcon 2008.

## **Portfolio allocation: q63-65**

These three questions give a rough measure of portfolio allocation, and are repeated from CogEcon 2008.

## **Credit and financial distress: q66-74**

Questions 66 and 68-70 are repeated from CogEcon 2008.

## **Spending adjustments: q75-79**

These questions are new. CogEcon 2008 asked for weekly spending on food, whereas these questions ask how household spending on food, non-durables, and vehicles has changed since July 2008.

## **Information sources and advice: q80-85**

Questions 80 and 81 are repeated from CogEcon 2008.

Questions 82-85 are new to this survey, and ask about a respondent’s attention to financial matters and financial learning since July 2008.

## **Inflation, interest rates, equity premium, stock market volatility: q86-92**

These questions measure respondents’ knowledge and expectations about inflation, interest rates, the returns on stocks versus bonds, and stock market volatility.

**Subjective probability questions about housing, unemployment, Social Security and stocks: q93-101**

This battery of questions measures respondents' expectations about the future of Social Security, housing values, unemployment, and the stock market. Like similar question in the HRS, these questions solicit a probability between 0% and 100% that a particular situation will occur.

**Attitudes, optimism and trust: q112-122**

This battery of questions was added to measure participants' feelings of well-being and trust.

**Time and day completed: q123-125**

These questions are repeated from CogEcon 2008.

### III: Imputations and Calculated Variables

#### A. Financial Sophistication

There are two versions of each True/False statement about financial matters, so we standardized the responses in order to compare them across respondents and questions. We also constructed summary measures of “financial sophistication” using the set of true/false statements in CogEcon 2008 questions 17-41 and CogEcon 2009 questions 4-14. These questions from CogEcon 2009 were asked in both waves.

The randomization of questions, the procedure for constructing variables from the responses, and the summary measures are described below in greater detail.

##### 1. Randomization

Respondents were randomly assigned to receive one of two versions of these questions in CogEcon 2008. If the pre-assigned group variable “*c1\_fs\_random*” (or *c2\_fs\_random*) is equal to 1, a participant received “version 1” of the questionnaire, and *cW\_fs\_random* = 2 indicates “version 2” of the questionnaire.<sup>9</sup> Respondents were asked the same version of these questions in CogEcon 2009 .

CogEcon 2008 Question #	CogEcon 2009 Question #	Version 1 survey	Version 2 survey
17	***	FALSE	TRUE
18	4	FALSE	TRUE
19	***	TRUE	FALSE
20	***	TRUE	FALSE
21	5	TRUE	FALSE
22	***	FALSE	TRUE
23	6	TRUE	FALSE
24	***	TRUE	FALSE
25	7	FALSE	TRUE
26	***	FALSE	TRUE
27	***	TRUE	FALSE
28	8	FALSE	TRUE
29	***	TRUE	FALSE
30	***	FALSE	TRUE
31	9	TRUE	FALSE
32	***	TRUE	FALSE
33	10	FALSE	TRUE

<sup>9</sup> Although version 2 is not available in its entirety in the public release files, the set of financial sophistication questions are appended to the version 1 questionnaire.

34	11	FALSE	TRUE
35	***	FALSE	TRUE
36	***	TRUE	FALSE
37	***	FALSE	TRUE
38	12	TRUE	FALSE
39	13	TRUE	FALSE
40	***	FALSE	TRUE
41	14	TRUE	FALSE

## 2. 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 17 reads: “An investment advisor tells a 30-year-old couple that \$1,000 in an investment that pays a certain, constant interest rate would double in value to \$2,000 after 20 years. If so, that investment *would* be worth \$4,000 *in less than* 45 years.” The “false” version reads: “An investment advisor tells a 30-year-old couple that \$1,000 in an investment that pays a certain, constant interest rate would double in value to \$2,000 after 20 years. If so, that investment *would not* be worth \$4,000 *for at least* 45 years.” 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.

Because some respondents face a true version of a question while others face a false version, we first convert all questions to the true version so that the scale can be interpreted as “0% surely (correct answer)” to “100% surely (correct answer).” 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. We can then construct intervals within which a respondent would round to each answer choice, and the point-value we assign is the midpoint of this interval:

Cog Econ response		Interval of underlying probabilities that statement is true		Assigned score
surely false	100%	0%	5%	0.025
	90%	5%	15%	0.1
	80%	15%	25%	0.2
	70%	25%	35%	0.3
	60%	35%	45%	0.4
guess false	50%	45%	50%	0.475
guess true	50%	50%	55%	0.525
	60%	55%	65%	0.6

	70%	65%	75%	0.7
	80%	75%	85%	0.8
	90%	85%	95%	0.9
surely true	100%	95%	100%	0.975

### 3. Summary measures of financial sophistication

The following variables are constructed using these assigned midpoint scores:

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

## B. 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.

CogEcon 2009 had additional questions about whether and to what extent the current value or account balance had changed since July 2008. These “magnitude of change” questions gave respondents the option of using a percentage or an exact value.

### 1. 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.

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>10</sup>

### 1.2 Value assigned for ranges chosen

All of the variables with the suffix *\_val* are given in terms of monetary value.

<sup>10</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.



<b>Range options 1</b>	<b>Range selected</b>	<b>Value assigned</b>
CogEcon 2008: questions 42, 70, 74, 77-80, 82-85, 87- 88, 90-105, 109-115, 117- 123, 125-129, 132  CogEcon 2009: questions 16, 20, 29, 39, 41, 43, 45, 48, 55, 57, 66	\$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

<b>Range options 2</b>	<b>Range selected</b>	<b>Value assigned</b>
CogEcon 2008: questions 52-53	\$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
	None**	

<b>Range options 3</b>	<b>Range selected</b>	<b>Value assigned</b>
CogEcon 2008: questions 60, 63	\$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

Range options 4	Range selected	Value assigned
CogEcon 2008: questions 154_min, 154_max	\$0	0
	\$1-\$2500	1125.5
	\$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-\$75,000	62500.5
	\$75,001-\$100,000	87500.5
	\$100,001-\$125,000	112500.5
	\$125,001-\$150,000	137500.5
	\$150,001-\$175,000	162500.5
	\$175,001-\$200,000	187500.5
	Don't know	.

## 2. Magnitude of changes

CogEcon 2009 asked additional questions about the direction and magnitude of changes in values/balances since July 2008. The respondent could report the magnitude either as a percentage change or as a value of change. (An exception is question 40: because we ask about values of homes in the respondents' neighborhood, only a percentage change is relevant here.) Since we also ask for the current value, we imputed the value of change for those who reported a percentage, and vice versa, according to the following:

$$\% \text{ increase} = \frac{\text{amount increase}}{\text{July 2008 balance}} * 100 = \frac{\text{amount increase}}{\text{current balance} - \text{amount increase}} * 100$$

$$\% \text{ decrease} = \frac{\text{amount decrease}}{\text{July 2008 balance}} * 100 = \frac{\text{amount decrease}}{\text{current balance} + \text{amount decrease}} * 100$$

Current balances are found in the questions preceding the questions about magnitudes of change.

No imputations are made for those who report neither a percent nor a value change, or for those who report both. For web surveys, the latter might occur if the respondent used back/forward buttons repeatedly while completing this section of the survey.

Questions about the magnitude of changes were used to construct the following variables:

Variable name	Variable description
c2_q*_dir	Categorical variable for direction of change (increase or decrease) reported by the respondent
c2_q*_dec_pct (c2_q*_inc_pct)	Percentage change for those reporting a decrease (increase); this percentage is either reported by the respondent, or imputed if the respondent reported a change in value as well as a current value
c2_q*_dec_val (c2_q*_inc_val)	Value change for those reporting a decrease (increase); this value of change is either reported by the respondent, or imputed if the respondent reported a percentage change as well as a current value
c2_q*_path	Categorical variable for direction of change as well as whether the respondent originally reported a value of change, percentage of change, both, or neither. This can be used to determine whether the value or percentage was imputed.
c2_q*_dk	(web only) For questions 50 and 59 only: indicator variable for respondents who checked the “don’t know” option on the question asking whether or not respondents bought or sold stock
c2_q59_dk_ref	(mail only) For question 59 only: indicator variable for respondents who checked none of the yes/no boxes for selling or buying stock

Unless otherwise stated in the value labels, the values of *c2\_q\*\_path* variables are as follows:

- 1 increase, magnitude reported as value
- 2 increase, magnitude reported as percent
- 3 increase, magnitude reported both as a value and percent
- 4 increase, no magnitude reported
- 5 decrease, magnitude reported as value
- 6 decrease, magnitude reported as percent
- 7 decrease, magnitude reported both as a value and percent
- 8 decrease, no magnitude reported
- 9 no change
- 10 N/A/ have not had
- .d don't know
- .y did not report direction

NOTE: Questions 50 and 59 are slightly different than the others, since they ask about buying and selling of stock within retirement accounts (question 50, worded as “transferred money out of / into stock”) and outside of retirement accounts (question 59). The mail survey offers yes/no checkboxes for each action (transfer out or transfer in for question 50, and sell or buy for question 59).

## C. Mapping CogEcon 2009 to CogEcon 2008

Many questions were asked in both CogEcon 2008 and CogEcon 2009. Please compare wording carefully for slight differences, however. An important example is that CogEcon 2009 wealth questions asked about broader categories than—and may reflect only an imprecise correspondence to—the sum of several CogEcon 2008 responses..

Table 1: Mapping CogEcon 2009 to CogEcon 2008

<b>CogEcon 2009 =</b>	<b>CogEcon 2008</b>	<b>CogEcon 2009 =</b>	<b>CogEcon 2008</b>
1	4	between 37 & 38	75
2	10	38	76
3	13	39	77
4	18	41	78-80
5	21	43	81-82, 86-87
6	23	45	83-85, 88
7	25	48	94
8	28	55	113
9	31	57	116, 118-123, 124-127
10	33	61	137
11	34	62	138
12	38	63	141
13	39	64	142
14	41	65	143
16	42	66	92
17	67	68	144
18	68	69	145
19	69	70	146
20	70	80	147
26	71	102-106	148-152
27	72	123	158
28	73	124	159
29	74	125	160