USER'S GUIDE FOR RISK PREFERENCE PARAMETERS

by

Miles S. Kimball, Claudia R. Sahm, and Matthew D. Shapiro

July 26, 2007

Online documentation and data for

Miles S. Kimball, Claudia R. Sahm, and Matthew D. Shapiro
"Imputing Risk Tolerance from Survey Responses" *JASA* (forthcoming).

ABSTRACT

This document explains how to use the values of individual risk preference based on the responses to the hypothetical gambles over lifetime income. The accompanying spreadsheet gives numerical values of the imputations. This document and the spreadsheet are archived at http://www.amstat.org/publications/jasa/supplemental_materials.

This document explains how to use the values of individual risk preference based on the responses to the hypothetical gambles over lifetime income. Refer to "Imputing Risk Tolerance from Survey Responses" (Kimball, Sahm, and Shapiro, 2007) for details on the survey question and the estimation. Refer to Barsky, Juster, Kimball, and Shapiro (1997) for the original implementation of this approach.

Risk Preference Imputations

The file r**iskprefs_impute.xls** provides two sets of imputations using the techniques in Kimball, Sahm, and Shapiro (2007).

1) HRS Respondent Imputations (**hrsimpute**): Proxy values for each of the original HRS respondents using the individual identifier of hhidpn (hhid + pn*1000)
2) Response Category Imputations **(rcimpute**): Proxy values for each possible set of gamble response categories. These imputations can be merged into other data sets.

HRS Respondent Imputations (**hrsimpute**)

The first set of imputations is for analysis with the original respondents in the Health and Retirement Study. Merge the desired proxy for log risk tolerance (*log_rtol*), risk tolerance (*rtol*) or risk aversion (*ravers*) to your data set by the individual identifier.

Response Category Imputations **(rcimpute**)

The second set of imputations is for analysis with gamble responses on data sets other than the HRS, such as the 1996 Panel Study of Income Dynamics. This data set translates a set of gamble responses to an individual measure of risk tolerance.

When using other data sets, there are four additional steps:

1) Determine the version of the gamble questions (original or status quo bias free)
2) Calculate an individual's gamble response category c={1-2,1, 2, 3, 4, 5, or 5-6}in each wave
3) Merge the imputation data set with the gamble response categories

Identifying the Question Version

The HRS has fielded two versions, the first or "original" version and the revised or "status-quo bias free" version. The original questions "o" begin:

> "Suppose that you are the only income earner in the family, and you have a good job guaranteed to give you your current (family) income every year for life. You are given the opportunity to take a new and equally good job, with a 50-50 chance it will double your (family) income and a 50-50 chance that it will cut your (family) income by a third. *Would you take the new job?"*

The revised questions "f" begin:

> "Suppose that you are the only income earner in the family. Your doctor recommends that you move because of allergies, and you have to choose between two possible jobs. The first would guarantee your current total family income for life. The second is possibly better paying, but the income is also less certain. There is a 50-50 chance the second job would double your total lifetime income and a 50-50 chance that it would cut it by a third. *Which job would you take --- the first job or the second job?"*

Assignment of Response Categories

If the question series asks each individual two gambles, then there are four possible response categories (as in the 1992 HRS). If instead an individual answers up to three gambles, then there are six possible response categories (as in the 1994 HRS and later waves). Table 1 relates the response category to the largest downside risk accepted and smallest risk rejected across the gambles

Table 1. Gamble Response Categories

4-way (*rc4_Qw*)			6-way (*rc6_Qw*)		
Cateogory	Accept	Reject	Category	Accept	Reject
1-2	None	1/5	1	None	1/10
3	1/5	1/3	2	1/10	1/5
4	1/3	1/2	3	1/5	1/3
5-6	1/2	None	4	1/3	1/2
			5	1/2	3/4
			6	3/4	None

Note that the four-way categories of "1-2" and "5-6" are both composites of two 6-way categories. These variables are recorded in the imputation file in character not numeric format. A code of "0" is denotes no response to a particular question.

The response category of individuals on your data set should be recorded in a character variable *rcC_Qw*, where *C* is the number of categories (4 or 6), *Q* is the question version ("o" or "f"), and *w* is the wave number ("1", "2", or "3"). This is the structure of the category variable on the imputation set and will be the merging variable.

General Remarks on Using the Imputations

Each set of imputations includes the conditional expectation of log risk tolerance (*log_rtol*), risk tolerance (*rtol*), and risk aversion (*ravers*). It is not correct to take a function of these variables, for example $E(\theta^2 \mid c) \neq [E(\theta \mid c)]^2$. Higher order imputations are possible, but they require the use of the distributional parameters and the formulas in the paper.

When using the individual measure of risk preference in a linear regression, the GMM estimator and R-squared formula with the true-to-proxy variance ratio λ should be used. This procedure allows for proper inference on both the estimated effect of risk preferences and of the other covariates. Since the variability in the proxy can differ across samples, the ratio λ should be re-computed in your particular application. For log risk tolerance, the numerator for the ratio λ is 0.526 (the estimated variance of risk tolerance) in each application, but the denominator should be the variance of the proxy values for the

individuals in the particular application. Table 2 contains the λ ratio associated with each of the three proxy forms for the HRS respondents.

Table 2. True-to-Proxy Variances Ratio λ for HRS

Proxy	True Variance	Variance of Proxy	True-to-Proxy Variances Ratio λ
Log-Risk Tolerance	0.526	0.091	5.76
Risk Tolerance	0.030	0.005	6.32
Risk Aversion	46.57	5.27	8.84

NOTE: The sample includes 11,616 original HRS respondents who answer an income gamble in 1992 or 1994.

To illustrate the importance of calculating the true-to-proxy variance ratio λ for each sample, Table 3 repeats the calculations for individuals who answered the original version of the income gambles in the 1996 PSID. Tabulations of the gamble response categories in the PSID 1996 are provided online (http://psidonline.isr.umich.edu/Data/Documentation/Cbks/Supp/rt.html). For the calculations in Table 3, the imputed levels of risk tolerance from **rcimpute** are assigned to the PSID respondents.

Table 3. True-to-Proxy Variances Ratio λ for PSID

Proxy	True Variance	Variance of Proxy	True-to-Proxy Variances Ratio λ
Log-Risk Tolerance	0.526	0.081	6.53
Risk Tolerance	0.030	0.004	6.81
Risk Aversion	46.57	4.386	10.62

NOTE: Proxy values from 5,565 working household heads in the 1996 PSID.

The estimated distribution of true risk preference is the value from the HRS sample. (Note that the PSID only asks the income gambles in one survey wave, so it is not possible to re-estimate the maximum-likelihood that accounts for survey response error.) With only one response from each individual and six possible response categories, the proxy values are also less variable in the PSID than in the HRS. Thus the true-to-proxy variance ratio λ is higher for the PSID respondents.

An alternative to the GMM estimator is to re-estimate the maximum-likelihood model with the gamble responses and the covariates in your particular application and form new imputations.

Index of Worksheets on the Imputation File (**riskprefs_impute.xls**)

mle	Provides maximum-likelihood estimates from Kimball, Sahm, and Shapiro
hrsimpute	Provides the proxy values for the original HRS respondents
rcimpute	Provides the proxy values based on survey response categories

Variable Definitions on Imputation File

hrsimpute	*hhidpn*	HRS identifier (hhid x 1000 + pn)
	rc4_1992	4-way response category in 1992 HRS
	rc6_1994	6-way response category in 1994 HRS
	rc6_1998	6-way response category in 1998 HRS
	rc6_2000	6-way response category in 2000 HRS

	rc6_2002	6-way response category in 2002 HRS
	log_rtol	Imputed value of log risk tolerance
	rtol	Imputed value of risk tolerance
	ravers	Imputed value of risk aversion
rcimpute	*rc4_o1*	4-way response category from original question
	rc6_o1	6-way response category from original question
	rc6_f1	6-way response category from first SQB-free question
	rc6_f2	6-way response category from second SQB-free question
	rc6_f3	6-way response category from third SQB-free question
	log_rtol	Imputed value of log risk tolerance
	rtol	Imputed value of risk tolerance
	ravers	Imputed value of risk aversion

The imputed values of risk preference (*log_rtol*, *rtol*, and *ravers)* are conditional on the individual's gamble response categories. The unconditional mean for log risk tolerance, risk tolerance, and risk aversion is -1.84, 0.206, and 8.2 respectively.

Health and Retirement Study Data and Documentation

The source data from the Health and Retirement Study can be found online (http://hrsonline.isr.umich.edu/). The end of this document contains a summary of the wording and sampling of the job gamble questions in the 1992, 1994, 1998, 2000, and 2002 HRS.

For further information and potential updates of these data, refer to the Researcher Contribution section of the online HRS site (http://hrsonline.isr.umich.edu/data/avail.php#rescon) or to http://www-personal.umich.edu/~shapiro/data/risk_preference/.

References

Barsky, Robert B., F. Thomas Juster, Miles S. Kimball, and Matthew D. Shapiro, "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study," *The Quarterly Journal of Economics*, 1997, *112*(2), 537-579.

Luoh, Ming-Ching and Frank Stafford, "Estimating Risk Tolerance from the PSID," Online Documentation, http://psidonline.isr.umich.edu/Data/Documentation/Cbks/Supp/rt.html.

Kimball, Miles S., Claudia R. Sahm, and Matthew Shapiro, "Imputing Risk Tolerance from Survey Responses," *Journal of the American Statistical Association* (forthcoming).

1992 HRS
Section L: Cognition
Variables: v5122, v5123
Sample: All non-proxy respondents

v5122
L14. Now I have another kind of question. Suppose that you are the only income earner in the family, and you have a good job guaranteed to give you your current (family) income every year for life. You are given the opportunity to take a new and equally good job, with a 50-50 chance it will double your (family) income and a 50-50 chance that it will cut your (family) income by a third. Would you take the new job?

1. YES	5. NO	8. DK	9. NA
→L14a	→ L14b	→ End	→ End

v5123
L14a. Suppose the chances were 50-50 that it would double your (family) income, and 50-50 that it would cut it in half. Would you still take the new job?

1. YES	5. NO	8. DK	9. NA
→End	→ End	→ End	→ End

L14b. Suppose the chances were 50-50 that it would double your (family) income and 50-50 that it would cut it by 20 percent. Would you then take the new job?

1. YES	5. NO	8. DK	9. NA
→End	→ End	→ End	→ End

1994 HRS
Module 5
Variables: W9344, W9345, W9346, W9347, W9348
Sample: 10% sub-sample of respondents

W9344
5-1 Now I have another kind of question. Suppose that you are the only income earner in the family, and you have a good job guaranteed to give you your current (family) income every year for life. You are given the opportunity to take a new and equally good job, with a 50-50 chance it will double your (family) income and a 50-50 chance that it will cut your (family) income by a third. Would you take the new job?

1. YES	5. NO	8. DK	9. NA
→5-1a	→ 5-1b	→ End	→ End

W9345
5-1a Suppose the chances were 50-50 that it would double your (family) income, and 50-50 that it would cut it in half. Would you still take the new job?

1. YES	5. NO	8. DK	9. NA
→5-1d	→ End	→ End	→ End

W9346
5-1b Suppose the chances were 50-50 that it would double your (family) income and 50-50 that it would cut it by 20 percent. Would you then take the new job?

1. YES	5. NO	8. DK	9. NA
→End	→ 5-1c	→ End	→ End

W9347
5-1c Suppose the chances were 50-50 that it would double your (family) income and 50-50 that it would cut it by 10 percent. Would you then take the new job?

1. YES	5. NO	8. DK	9. NA
→End	→ End	→ End	→ End

W9348
5-1d Suppose the chances were 50-50 that it would double your family income and 50-50 that it would cut it by 75 percent, would you still take the new job?

1. YES	5. NO	8. DK	9. NA
→End	→ End	→ End	→ End

1998 HRS

Section H: Expectations
Variables: F4614, F4615, F4616, F4617, F4618
Sample: All new, non-proxy respondents, 10% random sample of returning HRS respondents

F4614
H18. Now I have another kind of question. Suppose that you are the only income earner in the family. Your doctor recommends that you move because of allergies, and you have to choose between two possible jobs. The first would guarantee your current total family income for life. The second is possibly better paying, but the income is also less certain. There is a 50-50 chance the second job would double your total lifetime income and a 50-50 chance that it would cut it by a third. Which job would you take -- the first job or the second job?

1. FIRST JOB	2. SECOND JOB	8. DK/NA	9. RF
→H18.c	→ H18.a	→ End	→ End

F4615
H18.a Suppose the chances were 50-50 that the second job would double your lifetime income, and 50-50 that it would cut it in half. Would you take the first job or the second job?

1. FIRST JOB	2. SECOND JOB	8. DK/NA	9. RF
→End	→ H1.b	→ End	→ End

F4616
H18.b Suppose the chances were 50-50 that the second job would double your lifetime income and 50-50 that it would cut it by seventy-five percent. Would you take the first job or the second job?

1. FIRST JOB	2. SECOND JOB	8. DK/NA	9. RF
→End	→ End	→ End	→ End

F4617
H18.c Suppose the chances were 50-50 that the second job would double your lifetime income and 50-50 that it would cut it by twenty percent. Would you take the first job or the second job?

1. FIRST JOB	2. SECOND JOB	8. DK/NA	9. RF
→H1.d	→ End	→ End	→ End

F4618
H18.d Suppose the chances were 50-50 that the second job would double your lifetime income and 50-50 that it would cut it by 10 percent. Would you take the first job or the second job?

1. FIRST JOB	2. SECOND JOB	8. DK/NA	9. RF
→End	→ End	→ End	→ End

2000 HRS

Section H: Expectations

Variables: G5027, G5033, G5034, G5035, G5036

Sample: 10% random sample of respondents under age 65; Also a rule to include respondents who were asked income gambles in 1998 and exclude respondents who were asked altruism module in 1996 with a random portion of ties also included.

G5027
H18. Now I have another kind of question. Suppose that you are the only income earner in the family. Your doctor recommends that you move because of allergies, and you have to choose between two possible jobs. The first would guarantee your current total family income for life. The second is possibly better paying, but the income is also less certain. There is a 50-50 chance the second job would double your total lifetime income and a 50-50 chance that it would cut it by a third. Which job would you take -- the first job or the second job?

1. FIRST JOB	2. SECOND JOB	8. DK/NA	9. RF
→H18.c	→ H18.a	→ End	→ End

G5033
H18.a Suppose the chances were 50-50 that the second job would double your lifetime income, and 50-50 that it would cut it in half. Would you take the first job or the second job?

1. FIRST JOB	2. SECOND JOB	8. DK/NA	9. RF
→End	→ H1.b	→ End	→ End

G5034
H18.b Suppose the chances were 50-50 that the second job would double your lifetime income and 50-50 that it would cut it by seventy-five percent. Would you take the first job or the second job?

1. FIRST JOB	2. SECOND JOB	8. DK/NA	9. RF
→End	→ End	→ End	→ End

G5035
H18.c Suppose the chances were 50-50 that the second job would double your lifetime income and 50-50 that it would cut it by twenty percent. Would you take the first job or the second job?

1. FIRST JOB	2. SECOND JOB	8. DK/NA	9. RF
→H1.d	→ End	→ End	→ End

G5036
H18.d Suppose the chances were 50-50 that the second job would double your lifetime income and 50-50 that it would cut it by 10 percent. Would you take the first job or the second job?

1. FIRST JOB	2. SECOND JOB	8. DK/NA	9. RF
→End	→ End	→ End	→ End

2002 HRS

Section P: Expectations
Variables: HP036, HP037, HP038, HP039, HP040
Sample: All non-proxy respondents under age 65

HP036

Now I have another kind of question. Suppose that you are the only income earner in the family. Your doctor recommends that you move because of allergies, and you have to choose between two possible jobs. The first would guarantee your current total family income for life. The second is possibly better paying, but the income is also less certain. There is a 50-50 chance the second job would double your total lifetime income and a 50-50 chance that it would cut it by a third. Which job would you take -- the first job or the second job?

1. FIRST JOB	2. SECOND JOB	8. DK/NA	9. RF
→HP039	→ HP037	→ End	→ End

HP037

Suppose the chances were 50-50 that the second job would double your lifetime income, and 50-50 that it would cut it in half. Would you take the first job or the second job?

1. FIRST JOB	2. SECOND JOB	8. DK/NA	9. RF
→End	→ HP038	→ End	→ End

HP038

Suppose the chances were 50-50 that the second job would double your lifetime income and 50-50 that it would cut it by seventy-five percent. Would you take the first job or the second job?

1. FIRST JOB	2. SECOND JOB	8. DK/NA	9. RF
→End	→ End	→ End	→ End

HP039

Suppose the chances were 50-50 that the second job would double your lifetime income and 50-50 that it would cut it by twenty percent. Would you take the first job or the second job?

1. FIRST JOB	2. SECOND JOB	8. DK/NA	9. RF
→HP040	→ End	→ End	→ End

HP040

Suppose the chances were 50-50 that the second job would double your lifetime income and 50-50 that it would cut it by 10 percent. Would you take the first job or the second job?

1. FIRST JOB	2. SECOND JOB	8. DK/NA	9. RF
→End	→ End	→ End	→ End