February 18, 2009

"Risk Preferences in the PSID: Individual Imputations and Family Covariation"
by Miles S. Kimball, Claudia R. Sahm, and Matthew D. Shapiro
American Economic Review Papers and Proceedings (May 2009)

APPENDIX

This appendix provides additional details on our maximum-likelihood estimates of risk tolerance in the PSID. For a more thorough discussion of the general approach, see Kimball et al., "Imputing Risk Tolerance from Survey Responses" (2008) *Journal of the American Statistical Association*, 103(483) 1028-38. Further information on using the imputations is provided at http://www.umich.edu/~shapiro/data/risk_preference

A-I. Interpreting Gamble Responses

The 1996 PSID poses up to three hypothetical gambles to family respondents who are working at the time of the survey. The gambles differ only by the downside risk associated with the risky job. Specifically, individuals choose between a job that guarantees their current lifetime income and one that offers a 50-50 chance of doubling their lifetime income and a 50-50 chance of cutting it by a fraction π. We assume that an individual accepts the risky job only if its expected utility exceeds that of the certain job, thus individuals with higher risk tolerance θ are willing to accept jobs with higher downside risk π. With constant relative risk aversion, $U(C) = (1 - C^{1-1/\theta})/(1-1/\theta)$, gamble responses further imply an upper and lower bound on an individual's risk tolerance *in the absence of response error*. Table A-1 defines the gamble response categories in terms of the smallest downside risk rejected and the highest downside risk accepted. The last two columns provide the bounds on relative risk tolerance consistent with these categories.

Table A-1: Risk Tolerance Response Categories

Response	Downside Risk of Risky Job		Bounds on Risk Tolerance	
Category	Accepted	Rejected	Lower	Upper
1	None	1/10	0	0.13
2	1/10	1/5	0.13	0.27
3	1/5	1/3	0.27	0.50
4	1/3	1/2	0.50	1.00
5	1/2	3/4	1.00	3.27
6	3/4	None	3.27	∞

The response category is our summary statistic of an individual's sequence of gamble responses and we use the implied bounds on risk tolerance in the maximum-likelihood estimation as the known cut points for an ordered probit (interval regression).

A-II. Statistical Model and Estimation of Risk Tolerance

To translate the bounds on response categories to a parameter estimate, we first assume that risk tolerance is log-normally distributed,

$$\log\theta = x \sim N\left(\mu, \sigma_x^2\right), \tag{1}$$

which corresponds well with the fact that the modal gamble response implies low risk tolerance, but there is substantial heterogeneity across individuals. Previous analysis of individuals who answered the gambles repeatedly over several waves of the HRS suggests that the survey responses provide a noisy signal of risk tolerance. Therefore, we estimate a model of noisy log risk tolerance from the gamble responses:[1]

$$\xi = \log\theta + \varepsilon = \log\theta + b + e \tag{2}$$

The survey response error ε includes both a time-constant status quo bias term b and a transitory classical measurement error term $e \sim N\left(0, \sigma_e^2\right)$. For individuals in the PSID, the probability of being in response category j is

$$P(c=j) = P\left(\log\underline{\theta}_j < \xi < \log\overline{\theta}_j\right) = \Phi\left(\frac{\log\overline{\theta}_j - \mu - b}{\sqrt{\sigma_x^2 + \sigma_e^2}}\right) - \Phi\left(\frac{\log\underline{\theta}_j - \mu - b}{\sqrt{\sigma_x^2 + \sigma_e^2}}\right) \tag{3}$$

where Φ is the cumulative normal distribution function. However, with only a single response from each PSID respondent, it is not possible to separately identify the variance of true log risk tolerance σ_x^2 from the variance of the response error σ_e^2 or to estimate the status quo bias induced by the new/risky job wording. Thus with the PSID gamble responses, we first estimate the mean $\mu_\xi = \mu + b$ and variance $\sigma_\xi^2 = \sigma_x^2 + \sigma_e^2$ of the noisy signal ξ. The first column of Table A-2 provides the estimates from the PSID that ignore survey response error. The second column provides the estimates from the HRS panel that account for survey response error.[2] Some HRS respondents answer the gambles in more than one wave, so the variance of true risk tolerance is identified by the covariance in an individual's gamble responses at two points in time. This identification requires

[1] To simplify notation, the model equations in the text did not explicitly include the status quo bias term b.

[2] The HRS sample for the estimation in this paper includes original HRS respondents who were between ages 20-69 in 1992. To be consistent with the PSID data where gamble responses are only obtained from working respondents, gamble responses from the 1992, 1994, 1998, 2000, and 2002 HRS are included in the dataset for this paper only if the respondent is working at the time of the interview. In addition, we require a valid gamble response in 1992 to be in the HRS sample. Hence, the sample from the HRS used in this paper differs from that used in Kimball, Sahm, and Shapiro (2008). That study does not impose age limits and it includes all respondents, not just those working for pay when they answered the gamble question.

that the measurement error in the gamble responses is transitory and that preferences are the only source of persistence in the gamble responses. The status quo bias is identified in the HRS, since there are two versions of the question—one in which only the risky job is new and one in which both the certain and the risky jobs are new. The final column of Table A-2 shows the estimated distribution of risk tolerance in the PSID that incorporates the estimated variance of true risk tolerance and status quo bias from the HRS. The mean of true log risk tolerance is higher and the variance is considerably smaller than in the first column where there is no correction for measurement error.

Table A-2: Maximum-Likelihood Estimates

Parameter	PSID: Ignoring Response Error	HRS: Modeling Response Error	PSID: Calibrating Response Error
Log of risk tolerance			
Mean	-1.26	-1.77	-1.05
	(0.02)	(0.04)	(0.02)
Variance	2.46	0.76	0.76
	(0.07)	(0.07)	-
Status-quo bias		-0.21	-0.21
		(0.04)	-
Transitory response error			
Variance		2.03	1.69
		(0.07)	(0.07)

Note: Asymptotic standard errors are in paretheses. The PSID sample includes 5,466 respondents in the 1996 PSID. The HRS sample includes 7,648 respondents and 10,502 gambles responses in the 1992, 1994, 1998, 2000, and 2002 HRS.

A-III. Individual Imputations

Table 1 in the text provides imputations of risk preferences for each of the gamble response categories. These imputations are the expected value of an individual's risk preference conditional on his or her gamble response category. We use the parameter estimates from the final column of Table A-2 and the following formulas to compute the conditional expectations, which rely on the log-normality of risk tolerance.

The conditional expectation of log risk tolerance for individuals in response category *c* is

$$E(\log\theta \mid c) = \mu + \left(\sigma_x^2 / \sigma_\xi\right) \frac{\phi\left((\log\underline{\theta}_j - \mu - b)/\sigma_\xi\right) - \phi\left((\log\overline{\theta}_j - \mu - b)/\sigma_\xi\right)}{\Phi\left((\log\overline{\theta}_j - \mu - b)/\sigma_\xi\right) - \Phi\left((\log\underline{\theta}_j - \mu - b)/\sigma_\xi\right)} \quad (4)$$

where ϕ is the standard normal density function.

Using the moment-generating function, the conditional expectation of risk tolerance is

$$E(\theta \mid c) = \exp(\mu + \sigma_x^2 / 2) \frac{\Phi\left((\log \underline{\theta}_j - \mu - b - \sigma_x^2) / \sigma_\xi\right) - \Phi\left((\log \overline{\theta}_j - \mu - b - \sigma_x^2) / \sigma_\xi\right)}{\Phi\left((\log \overline{\theta}_j - \mu - b) / \sigma_\xi\right) - \Phi\left((\log \underline{\theta}_j - \mu - b) / \sigma_\xi\right)} \quad (5)$$

Finally, the conditional expectation of risk aversion $\gamma = 1/\theta$ is

$$E(\gamma \mid c) = \exp(-\mu + \sigma_x^2 / 2) \frac{\Phi\left((\log \underline{\theta}_j - \mu - b + \sigma_x^2) / \sigma_\xi\right) - \Phi\left((\log \overline{\theta}_j - \mu - b + \sigma_x^2) / \sigma_\xi\right)}{\Phi\left((\log \overline{\theta}_j - \mu - b) / \sigma_\xi\right) - \Phi\left((\log \underline{\theta}_j - \mu - b) / \sigma_\xi\right)} \quad (6)$$

Again, the formulas for the conditional expectations make clear that the imputation of risk aversion is not simply the reciprocal of the imputation of risk tolerance.

A-IV. Family Estimation

We use the unique intergenerational structure of the PSID to examine the covariation of risk preferences within families. We use pair-wise comparisons of gamble responses from two different types of family members, such as adult children and their fathers. Similar to the statistical model of Section A-II, we model the noisy signal of risk tolerance from the first family member's gamble response as

$$\xi_1 = \log \theta_1 + \varepsilon_1 \sim N(\mu_{\xi 1}, \sigma_{\xi 1}^2) \quad (7)$$

We allow the second family member to have a different mean and variance:

$$\xi_2 = \log \theta_2 + \varepsilon_2 \sim N(\mu_{\xi 2}, \sigma_{\xi 2}^2) \quad (8)$$

The main parameter of interest (reported in Table 3 in the text) is the covariance between family members $\text{Cov}(\xi_1, \xi_2) = \sigma_{12}^2$. We assume that the response errors are uncorrelated across family members, so the covariance term reflects the covariation in true risk preferences. There are two gamble responses observed for each family, so the likelihood of family member 1 being in gamble response category j and family member 2 being in response category k is calculated as

$$P(c_1 = j, c_2 = k) = \vec{\Phi}(\overline{N}_{j1}, \overline{N}_{k2}, \rho) + \vec{\Phi}(\underline{N}_{j1}, \underline{N}_{k2}, \rho) - \vec{\Phi}(\overline{N}_{j1}, \underline{N}_{k2}, \rho) - \vec{\Phi}(\underline{N}_{j1}, \overline{N}_{k2}, \rho), \quad (9)$$

where $\vec{\Phi}$ is the bivariate normal cumulative distribution function, ρ is the correlation between the two family members, $\overline{N}_{j1} = (\log \overline{\theta}_j - \mu_1)/\sigma_{\xi 1})$ $\underline{N}_{j1} = (\log \underline{\theta}_j - \mu_1)/\sigma_{\xi 1})$, $\overline{N}_{k2} = (\log \overline{\theta}_k - \mu_2)/\sigma_{\xi 2})$ and $\underline{N}_{k2} = (\log \underline{\theta}_k - \mu_2)/\sigma_{\xi 2})$.[3] With two gamble responses from the same family we can identify the family covariance term; however, with only one response from each individual we cannot separate the idiosyncratic variance of true risk tolerance from the variance of response errors. Likewise, with the family pairs, we cannot estimate the status quo bias, since we only have responses to the "new job" version of the question.

Table A-3 provides the estimated distribution of risk tolerance for the various family members. As in Table A-2, we adjust the estimates from the family member pairs with the variance of true log risk tolerance and the status quo bias estimated in the HRS. We assume that the values of these two calibrated parameters are the same for all family members.

The row labeled "Pair-Specific" Variance provides the estimates that are reported (below the diagonal) in Table 3 in the text. In line with the age effects discussed in the text, the mean risk tolerance of the older family member is lower and the variance of the response error is higher.

[3] The estimator uses Gaussian quadrature to approximate the probability in equation (9).

Table A-3: Maximum-Likelihood Estimates for Family Members

	Family Member Pairs			
Parameter	Child-Father	Child-Mother	Older Sibling	Wife-Husband
Log of risk tolerance				
Mean, 1st in Pair	-0.71	-0.76	-0.91	-1.64
	(0.07)	(0.06)	(0.04)	(0.09)
Mean, 2st in Pair	-1.60	-1.79	-1.05	-1.65
	(0.10)	(0.08)	(0.04)	(0.10)
Variance	0.76	0.76	0.76	0.76
	-	-	-	-
Idiosyncratic	0.65	0.59	0.39	0.45
	-	-	-	-
Pair-Specific	0.11	0.18	0.37	0.32
	(0.13)	(0.11)	(0.06)	(0.13)
Status Quo Bias	-0.21	-0.21	-0.21	-0.21
	-	-	-	-
Transitory Response Error				
Variance, 1st in Pair	1.38	1.31	1.83	1.36
	(0.20)	(0.17)	(0.11)	(0.28)
Variance, 2st in Pair	2.91	2.08	1.83	1.91
	(0.38)	(0.25)	(0.11)	(0.36)
Number of Pairs	557	757	2,300	710
Mean age difference	26	23	5	7
(Standard deviation)	(5)	(5)	(4)	(5)

Note: The estimates above use the total variance of true log risk tolerance (equal to 0.76) and the status quo bias from the HRS and assume that these two parameters are the same for all family members. The gamble responses of parents, adult children, and adult siblings are from the 1996 PSID. The gamble responses of spouses (not interviewed together) are from the 1992 HRS. Each column is a separate estimation.

A-V. Survey Questions in PSID and HRS

Both the PSID and HRS pose the gamble as a choice between two jobs. The wording of the PSID question is similar to the original version of the HRS question.

Specifically, the PSID asks:

> Suppose you had a job that guaranteed you income for life equal to your current, total income. And that job was (your/your family's) only source of income. Then you are given the opportunity to take a new, and equally good, job with a 50-50 chance that it will double your income and spending power. But there is a 50-50 chance that it will cut your income and spending power by a third. *Would you take the new job*?

Similarly, the 1992 and 1994 HRS asks:

> 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?*

Starting in 1998, the HRS modified the frame of the question to avoid the potential for status quo bias:

> 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?*

The italics (added here) highlight the main difference in the questions. Status quo bias is identified as the average difference in the gamble responses across the two versions in the HRS.