RISK PREFERENCES IN THE PSID:

INDIVIDUAL IMPUTATIONS AND FAMILY COVARIATION

January 1, 2009

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

Prepared for the American Economic Association Annual Meeting, San Francisco, January 2009

Session on "New Empirical Approaches to Decision Making under Uncertainty"

* Miles S. Kimball and Matthew D. Shapiro Department of Economics, University of Michigan, Ann Arbor, MI 48109 and National Bureau of Economic Research, Cambridge, MA 02138. Claudia R. Sahm, Federal Reserve Board, Washington DC 20551. The views presented are solely those of the authors and do not necessarily represent those of the Federal Reserve Board or its staff. Kimball and Shapiro gratefully acknowledge the support of National Institute on Aging grant 2-P01-AG10179

corresponding author:
Matthew D. Shapiro
Department of Economics
University of Michigan
611 Tappan St
Ann Arbor MI 48109-1220
tel. 734 764-5419
shapiro@umich.edu

Risk Preferences in the PSID: Individual Imputations and Family Covariation

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

Unobserved heterogeneity greatly complicates empirical analysis in economics. Unobserved heterogeneity in preferences is particularly troublesome because there are so few theoretical restrictions on the distribution of preference parameters in the population. Therefore, despite potential pitfalls, we have developed direct survey measures of preference parameters (based on hypothetical choices) and appropriate econometric techniques for dealing with the inevitable measurement error in any such measures. Our work on survey measures of preference parameters focuses on risk tolerance (Robert B. Barsky, F. Thomas Juster, Miles S. Kimball, and Matthew D. Shapiro 1997 [BJKS hereafter] and Kimball, Claudia R. Sahm, and Shapiro 2008 [KSS hereafter]), time preference and the elasticity of intertemporal substitution (BJKS), and labor supply elasticities (Kimball and Shapiro 2008).

Risk tolerance is central to portfolio choice—which is relatively independent of most other preference parameters—and to many other economic decisions, such as choices about insurance and career choices. In this paper, we discuss how to go from categorical survey responses to imputed values of preference parameters. The procedure takes into account measurement error from survey response, and has implications for the appropriate use of imputed preference parameters in econometrics analysis. We present the risk tolerance imputations for the survey responses in the Panel Study of Income Dynamics (PSID). We then apply these imputations to present quantitative evidence on the covariation in preferences between parent and child, and between sibling and sibling

I. Survey Measures of Risk Preference

Numerous surveys have fielded measures of an individual's willingness to take risk, including the Health and Retirement Study (HRS), which pioneered the use of hypothetical gambles in a large survey to measure the economic preference parameter of risk tolerance (BJKS 1997). In this paper, we analyze the gambles fielded in the 1996 PSID that ask respondents the following[1]:

> 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?

Individuals who would take this risky job were then asked about a riskier job:

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

In contrast, individuals who would not take the initial risky job were asked about a less risky job:

> Now, suppose the chances were 50-50 that the new job would double your (family) income, and 50-50 that it would cut it by 20 percent. Then, would you take the new job?

Given their first two responses, some individuals are asked to consider a risky job with either a 75 percent downside risk or a 10 percent downside risk. These responses allow us to order individuals into six categories. Unlike the HRS, the PSID only asked these questions of working respondents and did not ask the questions of other members of the household. The targeting of the questions to workers in the PSID particularly affects the selection of the youngest and oldest respondents, so we have limited our analysis to respondents between the ages of 20 and 69.

After collecting over ten years of gamble responses in the HRS and similar questions in surveys like the PSID and the National Longitudinal Study, a number of lessons on measuring risk preferences have emerged. First, the gamble responses are subject to considerable

[1] This question was included in the 1996 PSID (both original respondents and offsprings in split-off households).

measurement error. In the first two waves of the HRS, KSS report a rank correlation in the gamble response categories of less than 0.3. In addition, Sahm (2008) shows that much of the transitory variation in the gamble responses remains unexplained even after including a rich set of individual and household covariates. In response, Barksy et al. (1997) and subsequent analyses use multiple responses from some individuals to isolate the variance owing to measurement error from the variance in true risk preference.

Second, extraneous details in the description of the gambles can affect the measurement of risk preferences. For example, in the original HRS version and the PSID version of the question, the risky job is described as a new job. This frame has the potential to induce status quo bias in which individuals are averse to taking the new job independent from its income risk. Starting in the 1998 wave, the HRS addressed this potential problem by using a scenario in which the individual has to move for health reasons and is given a choice between two new jobs. This variation in the question wording in the HRS also allows us to estimate the degree of status quo bias in the original version and to correct the estimates of risk tolerance from the PSID. Finally, the interpretation of a job-related gamble may vary across workers who are at different stages of their career. In designing the question, a choice of jobs was used to create a large shock to lifetime resources; however, the fraction of lifetime income associated with labor income likely declines with age. The job gamble may be particularly hard for retirees and other non-workers to interpret. The HRS now uses an investment gamble related to an unexpected inheritance for respondents age 65 and older, and gives the job-gamble question only to those below 65 years old. Similarly, the PSID targeted its job-related question only to workers.

II. Individual Imputations

The responses to hypothetical gambles in the PSID suggest that most individuals have a low tolerance for risk, though there is substantial heterogeneity. The first column of Table 1 shows that as the modal response, 31 percent of the respondents rejected all of the risky jobs, but almost 7 percent accepted all the risky jobs. One advantage of these choices about hypothetical gambles relative to qualitative measures of risk tolerance is that one can use them to quantify the degree of risk tolerance and its dispersion across individuals. As in BJKS and KSS, we assume that individuals have constant relative risk aversion utility and will reject the risky job when its expected utility is less than that of the certain job. Along with the risks specified in the questions, these assumptions allow us to assign a range for the coefficient of relative risk tolerance to each gamble response category. Previous analysis from the panel of gambles in the HRS suggests that these questions provide a noisy signal of risk tolerance reflecting both status quo bias and classical measurement error. Therefore, as in KSS, we estimate a model of noisy log risk tolerance, $\xi = \log\theta + e$, where risk tolerance log θ is distributed $N\left(\mu, \sigma_x^2\right)$ and the e is classical measurement error distributed $N\left(0, \sigma_e^2\right)$. With only a single response from each PSID respondent, it is not possible to identify separately the variance of true log risk tolerance and the variance of the response error. The PSID responses identify the mean and the total variance of the noisy signal ξ. We impose KSS's estimate of the variance of true log tolerance σ_x^2 to divide the PSID total variance into variance of true preferences and variance of error. For the PSID, the estimates of the parameters are μ=-1.05, σ_x^2=0.87, and σ_e^2=1.3.[2] Using these distributional parameters, we can impute individual-level estimates of preference parameters based on the conditional expectation of the true parameter given the individuals' survey responses. Table 1 provides the individual imputations for the PSID. The conditional

[2] The PSID responses are also adjusted by -0.21 for status quo bias (again using the parameter estimated by KSS in the HRS).

expectation of each preference parameter is computed using the moment generating function, so as shown in the last two columns, the reciprocal of imputed risk tolerance is not equal to imputed risk aversion.

Table 1. Risk Tolerance in the PSID

		Imputatations		
Response Category	Percent of Respondents	Log Risk Tolerance	Risk Tolerance	Risk Aversion
1	30.9	-1.60	0.27	6.7
2	18.2	-1.18	0.40	4.2
3	15.6	-0.98	0.49	3.5
4	15.0	-0.77	0.60	2.8
5	13.7	-0.50	0.79	2.2
6	6.6	-0.08	1.22	1.4

Note: The 1996 PSID has 5,466 gamble respondents who are working and are between ages 20 and 69 at the time of the survey. Tabulations and regressions are unweighted. Imputations use MLE estimates from the PSID gambles that are adjusted for response error and status quo bias using estimates from the HRS.

Researchers interested in studying differences in risky behavior can use these individual imputations as an observable attribute. The imputations offer advantages relative to categorical controls for gamble responses. First, the imputations summarize the sequence of gamble responses in a single cardinal measure of preferences that can be used to assess the quantitative predictions of behavioral models. Second, our estimation procedure allows accounting for the measurement error in the survey responses, so the imputations are the conditional expectations of the individual's true preferences. The use of the imputed values in regression analysis substantially reduces the attenuation bias arising from survey response error. Nonetheless, these imputations that condition only on individual's gamble responses understate the true variation in preferences, so they may not capture all of the relevant differences in risk attitudes across individuals.

The application in KSS that uses individual imputations (as in Table 1) to study stock ownership makes these points more concrete. Using categorical controls for the gamble response category or imputations that do not account for response error leads to an attenuation bias that can substantially understate the responsiveness of behavior to risk tolerance. Even with imputations that address response error, standard multivariate estimators may not be consistent due to a nonstandard errors-in-variables problem.[3] The main issue is that the imputation based on gamble responses does not capture all the differences in true risk tolerance. To the extent that other covariates are correlated with the unmeasured part of risk tolerance, they will be correlated with the error term in the OLS regression that includes the imputations. Thus, the estimated coefficients on the other covariates would also include the indirect effects of risk tolerance. To address this issue, KSS provide a consistent GMM estimator using the imputations that scales up the covariance between imputed risk tolerance and other covariates—by a factor of 4.6 for the PSID.[4] As an example of the difference this correction makes, compared to OLS, the estimated difference in stock ownership rates between men and women is 40 percent lower with the GMM estimator and is no longer statistically different from zero at the 5% level.

The PSID illustrates how preference parameters can differ according to values of covariates. In particular, there are important differences in measured risk preference by age. For example, 61 percent of the individuals in their sixties reject all of the risky jobs versus only 23 percent of individuals in their twenties. As Table 2 shows, this pattern holds across all six gamble response categories with older individuals more concentrated in lower, less risk tolerant categories. The interpretation of such age effects remains open.[5] One possibility is that risk tolerance, in terms

[3] In univariate analysis, the estimated effect of risk tolerance on the behavior of interest would be consistent, but the R-squared would be underestimated.
[4] The scaling factor is the variance of true risk tolerance divided by the variance of imputed risk tolerance. This true-to-proxy variance ratio is 6.3 in the HRS and 4.6 in the PSID. See KSS for further details.
[5] With a cross-section of responses, the distribution of gamble responses by age may also incorporate differences in risk tolerance across birth cohorts. Malmendier and Nagel (2008) find an association between individuals' current willingness to take financial risks and the path of aggregate stock market returns experienced over their lifetime.

of the curvature of the utility function, diminishes with age. Alternatively, consumption commitments or habits may increase with age and make individuals less willing to risk a loss in income. Finally, the interpretation of the job-related gamble may simply vary with age in a way that is unrelated to true risk preferences.

Table 2. Distribution of Gamble Responses by Age

Response	Percent by Age Group				
Category	20-29	30-39	40-49	50-59	60-69
1	22.7	27.8	30.5	44.6	60.6
2	18.7	18.5	18.8	16.9	13.4
3	15.9	16.1	16.5	13.3	9.3
4	17.8	16.3	15.5	8.0	6.5
5	17.3	13.9	13.0	11.6	4.9
6	7.6	7.4	5.6	5.5	5.3

Note: Unweighted tabulations of PSID gamble respondents.

In any case, if the individual imputations also conditioned on age the differences would be sizeable. For example, if the imputation were conditioned on age in addition to the gamble response category, a 30 year old in the least risk tolerant category would have an imputed risk tolerance of 0.25, whereas a 50-year old with the same gamble responses would have an imputed risk tolerance of 0.16. One option would be to impute risk tolerance to individuals based on both their gamble response category and their age; however, this would constrain researchers who want to use the imputation as a covariate in behavioral studies. The specification of age effects in the behavioral model would have to match those in the risk tolerance estimation or a spurious correlation between imputed risk tolerance and the behavior under study could arise. For example, to use those in Table 2, one would have to use *exactly* the set of age dummies implicit in Table 2.

As mentioned above, the alternative is to use the imputations in Table 1 along with the GMM correction for errors-in-variables described in KSS. This alternative allows one to specify age effects in any way desirable for the analysis. More importantly, using the GMM correction allows one to add other covariates and get consistent estimates without redoing the imputations. This is an important advantage because generating the imputations requires a two-sample approach (using both the HRS and the PSID). The intuition for what the GMM correction in KSS does is as follows. Without the correction, other variables correlated with risk tolerance "steal some of the thunder" from risk tolerance because risk tolerance is proxied by an imperfect measure. This leads to biased coefficient estimates for both risk tolerance and these other variables. The GMM correction allocates credit to risk tolerance and the other variables appropriately.

In summary, Table 1 provides impute values of risk preference parameters based on responses to a hypothetical gamble about lifetime income in the PSID. The imputations control for survey response error. Neglecting this response error will substantially understate the correlation of survey measures of risk preferences with other variables. Kimball, Sahm, and Shapiro (2008) show how use such imputed values in multiple regressions—either my imputing the preference parameters based on multiple covariates or by using a GMM procedure that adjusts for the fact that the imputed values do not capture all the cross-sectional variation in the true preferences.

III. Family Covariation

We now imply the methodology sketched in Section II to study the covariation in preferences among family members. We exploit that the PSID has risk preference responses from members of different generations of families and that the HRS has responses from spouses.

We use our maximum-likelihood approach to quantify the covariation in family members' preferences. Consider the correlation in risk tolerance between a father f and his adult child c.

Because of the differences across age documented in Table 2, we allow the mean and variance of noisy log risk tolerance ξ to differ across fathers and children, such that $\xi_f \sim N\left(\mu_f, \sigma_f^{\,2}\right)$ and $\xi_c \sim N\left(\mu_c, \sigma_c^{\,2}\right)$ where, as above, the variances are sums of the variances of the true parameter and the response errors *e*. Because the response errors *e* are uncorrelated across family members, we can estimate the $\mathrm{Cov}\left(\xi_f, \xi_c\right) = \sigma_{fc}^2$.

The lower diagonal of Table 3 presents the variance-covariance matrix of log risk tolerance for various family members. We find a positive association between parents and their adult

Table 3. Family Covariation in Log Risk Tolerance

	Variance-Covariance \ Correlation			
	Father	Mother	Child 1	Child 2
Father	0.76	**0.41**	**0.14**	**0.14**
	(0.07)			
Mother	0.32	0.76	**0.23**	**0.23**
	(0.13)	(0.07)		
Child 1	0.11	0.18	0.76	**0.48**
	(0.13)	(0.11)	(0.03)	
Child 2	0.11	0.18	0.37	0.76
	(0.13)	(0.11)	(0.06)	(0.03)

Note: Observations are unweighted. The variance-covariance matrix with asymptotic standard errors is in the lower diagonal. The correlation coefficients in bold are above the main diagonal.
Source: PSID, HRS.

children. The correspondence between fathers and their children is relatively weak, though positive. The mother-child covariance is over 60 percent larger than the father-child covariance and is statistically different from zero at the 10 percent level.[6] The mother-child covariance is over one-fifth of the within-person variance. We do not find a stronger correlation between parents and children of the same gender. The correlations are noteworthy given the fact that

[6] The standard deviation of the common mother-child component in log risk tolerance is statistically different from zero at the 5 percent level. (The 95% confidence interval of the standard deviation and the 95% confidence interval of the variance cover non-equivalent regions of the parameter space.)

parents and children with an average age difference of over 20 years are at very different life stages and in most cases have not resided together for some time. The role of the family in shaping risk preferences is even more apparent in the gamble responses of siblings. Again, each adult sibling in the pair is either the head or spouse in an independent family when answering the gambles. The covariance in risk tolerance among siblings is more than twice the size of the mother-child covariance and is almost 50 percent of the within-person variance. The average age difference between the siblings is only 5 years, which likely makes their interpretation of the gambles more comparable. Clearly, there are a number of factors which could lead siblings to form similar risk preferences: transmission from common parents, shared experiences within the family, and similar peer and social environments. Looking for some evidence on these factors, we tested for a difference in the covariance for siblings who share both parents as opposed to those who share only one parent. The difference was statistically insignificant.

The HRS offers one more dimension of within family variation, since it asks the gamble responses to both the husband and wife in a household.[7] We find a covariance between spouses that is similar to the covariance between siblings and is about 40 percent of the within person variation. Both assortative mating and common experiences in the marriage could help account for the correlation.

The substantial covariance within families is also important for interpreting the variance of risk tolerance on the main diagonal in Table 3. The estimated variance from the HRS uses the persistent component of individuals' gamble responses over time to identify risk preference. It is possible that a repeated misinterpretation of the question could lead to persistent measurement

[7] One complication is that a spouse is sometimes present during the HRS interview which might bias an individual's response and lead to a spurious correlation in gamble responses. We limit our analysis to pairs of responses that were given in separate interviews in 1992. This may understate the true correlation if spouses with similar preferences choose to be together during interviews more than those with dissimilar preferences.

error that then would bias upward the estimated variance of true risk tolerance. Nonetheless, the size of the sibling and spousal covariances makes it unlikely that the true variance of risk tolerance on the main diagonal is much smaller. In other words, the size of the sibling and spousal covariances leaves little room for a large variance of persistent idiosyncratic response error. The correlation of response to the risk tolerance questions across family members is therefore important because there are few other ways to get a handle on the variance of persistent idiosyncratic response error.

Our results that show a correlation in risk preferences among family members are largely consistent with related studies. Using a subset of the PSID gamble responses in their study on the intergenerational transmission of wealth, Charles and Hurst (2003) find a strong correspondence between parent and child risk tolerance—particularly at the tails of the distribution. They make sample restrictions that result in a more homogeneous group of parent and child households. This leads to a stronger parent-child correlation than we find in the full sample; however, the basic finding of intergenerational transmission in risk preferences is similar.[8] Dohmen et al. (2008) use experimentally validated, qualitative measures of willingness to take risk in the German Socioeconomic Panel to also show that parents and children, as well as married couples, have similar attitudes toward risk.

IV. Conclusion

In this paper, we apply a survey-based method for imputing individual risk preferences to responses in the PSID. These procedures draw on estimates and previous lessons from analysis of the HRS gamble responses. We provide individual estimates of risk preferences based on the

[8] Charles and Hurst (2003) use a different method for assessing the covariation in preferences across parents and children. Applying our maximum likelihood procedure to their restricted sample of parent-child pairs yields a covariance of 0.25 (standard error of 0.12). The point estimate from their parent-child sample is higher than our parent-child covariance estimate of 0.15 (standard error of 0.08), though the difference is not statistically significant.

gamble response categories that can be used in other behavioral studies—both to study the effects of risk tolerance and to control for risk tolerance when looking at other effects. We use the gamble responses to document a substantial covariance in risk preferences among family members. In addition to its intrinsic interest, this covariance in risk preferences across family members helps validate these risk tolerance measures by putting an upper bound on the variance of idiosyncratic response error.

References

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

Charles, Kerwin. K. and Erik Hurst. 2003. "Intergenerational Wealth Correlations." *Journal of Political Economy* 111(6):1155-1182.

Dohmen, Thomas J., Armin Falk, David Huffman, and Uwe Sunde. 2008. "The Intergenerational Transmission of Risk and Trust Attitudes" IZA Discussion Paper No. 2380.

Kimball, Miles S., Claudia R. Sahm, and Matthew D. Shapiro. 2008. "Imputing Risk Tolerance from Survey Responses" Journal *of the American Statistical Association* 103(483):1028-1038.

Kimball, Miles S. and Matthew D. Shapiro. 2008. Labor Supply: Are Income and Substitution Effects Both Large or Both Small? NBER Working Paper No. 14208.

Malmendier, Ulrike and Stefan Nagel. 2008. "Depression Babies: Do Macroeconomic Experiences Affect Risk-Taking?" Unpublished paper. Stanford University.