Identification of risk preference

Chau Pham

August 15, 2023

1   Questions on risk preference

The method was first applied to the Health and Retirement Study (HRS) in Barsky et al. (1997) and Kimball et al. (2008) then to the Panel Study of Income Dynamics (PSID) (Kimball et al., 2009). For the HRS, questions on risk preference are presented to participants in 1992, 1994, 1998, 2000 and 2002. In the first two years, the following version of question was used:

> "...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 that 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?"

Depending on the responses the this question, participants are further asked a similar question with varying portion of income cut. If participants responded "The first job" in this question, they are presented with another question wherein the income cut of the risky job reduces to 20% whereas the income cut is increased to 50% should they choose "The second job." After 1994, the following modified version of question was administered to participants:

> "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 difference in the two versions lies in the framing of situations. As noted by Kimball et al. (2008), respondents to the original version might be affected by the status-quo bias and, at the same time, responses to both versions are contaminated by measurement errors. Having multiple responses across waves and versions allow them to separate variation in measurement error from variation in cross-sectional true risk preference. The original version of question is the same as one used in National Longitudinal of Youth 1979 (NLSY79) before 2010 in exact wordings. In 2010, 2012 and 2014, participants of NLSY79 were also asked a version of the status-quo bias free question as follows:

> "Suppose that you are the only income earner in the family, and that you have to choose between two new jobs. The first job would guarantee your current total family income for life. The second job is possibly better paying, but the income is also less certain. There is a 50-50 chance the second job would increase your total lifetime income by 20 percent and a 50-50 change that is would cut it by 10 percent."

The similarity in question framing and design allows the method applied to HRS to also be applied to NLSY79.

2 Risk imputation

The method of creating a cardinal measure of risk follows Barsky et al. (1997) and Kimball et al. (2008, 2009). The advantages of this method are two-fold. First, it provides a unidimensional, quantitative measure of risk tolerance/aversion that allows meaningful interpersonal comparison. Second, in many settings, such as demand for risky assets, economic theory makes predictions that link risk preference parameters quantitatively to economic decisions. Third, having a quantitative measure makes the correction of measurement error associated with survey responses more transparent.

2.1 Accounting for status-quo bias

The imputation procedure is as follows. Respondents are categorized into different groups by their sequence of responses. Under the assumption of **CRRA utility** (that is, an individual accepts the risky job when its expected utility exceeds the utility from the certain job) and **constant true risk preference**, we calculate the bounds for the coefficient of relative risk tolerance. For individuals who respond in only one wave, the likelihood of being in category j is

$$P(c = j) = P(\log \underline{\theta}_j < x < \log \bar{\theta}_j) = \Phi\Big(\frac{\log(\bar{\theta}_j) - \mu - b}{\sqrt{\sigma_x^2 + \sigma_{eq}^2 + \sigma_{\kappa q}^2}}\Big) - \Phi\Big(\frac{\log(\underline{\theta}_j) - \mu - b}{\sqrt{\sigma_x^2 + \sigma_{eq}^2 + \sigma_{\kappa q}^2}}\Big) \tag{1}$$

where μ and σ_x are respectively the mean and standard deviation of true log risk tolerance. σ_{eq}^2 is variance in individual's transitory response error for a particular wave and question type q while $\sigma_{\kappa q}^2$ is variance in individual's persistent response error for question type q. $\bar{\theta}_j$ and $\underline{\theta}_j$ are upper and lower bounds on the coefficient of risk tolerance. b denotes status quo bias (common bias across individuals) and σ_e standard deviation of transitory response error, assumed to be purely random and independent of an individual's true risk tolerance or any other attributes. If an individuals responded to the questions more than one time, the likelihood function employs a more complicated form of multivariate normal cumulative distribution function.

All parameters (except for bounds on risk tolerance which are calculated from respondents' choices) can be estimated using maximum likelihood. The estimated values are then used to impute the coefficient of risk aversion for individuals in response category c via conditional expectation,

$$\mathbb{E}[\theta|c] = \exp\Big(-\mu + \frac{\sigma_x^2}{2}\Big) \frac{\Phi\Big(\frac{\log(\bar{\theta}_j) - \mu - b + \sigma_x^2}{\sqrt{\sigma_x^2 + \sigma_{eq}^2 + \sigma_{\kappa q}^2}}\Big) - \Phi\Big(\frac{\log(\underline{\theta}_j) - \mu - b + \sigma_x^2}{\sqrt{\sigma_x^2 + \sigma_{eq}^2 + \sigma_{\kappa q}^2}}\Big)}{\Phi\Big(\frac{\log(\bar{\theta}_j) - \mu - b}{\sqrt{\sigma_x^2 + \sigma_{eq}^2 + \sigma_{\kappa q}^2}}\Big) - \Phi\Big(\frac{\log(\underline{\theta}_j) - \mu - b}{\sqrt{\sigma_x^2 + \sigma_{eq}^2 + \sigma_{\kappa q}^2}}\Big)} \tag{2}$$

It is further shown that the use of such a proxy in a univariate OLS would provide consistent estimates. However, in the multivariate setting, OLS estimator using the proxy is unlikely to provide consistent estimates of the population parameters as regressors that correlate with true risk tolerance (risk aversion) would also correlate with expectation error, defined as $\theta - \mathbb{E}[\theta|c]$. To obtain consistent estimates of parameters in the multivariate settings, Kimball et al. (2008) further assume linearity of other covariates in risk tolerance to derive two sets of orthogonality moment conditions that identify the model. Estimates of parameters can then be obtained via Generalised Methods of Moments (GMM).

2.2 Allowing for time-varying risk preference

Ahn (2010) adopted the approach to analyse risk preference and selection into self-employment in the NLSY79 but modified to allow for dependence of risk preference on other determinants. More specifically, she took survey responses in two years, 1993 and 2002 to calculate the true risk preference. The bounds of risk tolerance are modelled as

$$\log \theta_{it} = \beta X_{it} + \delta_i + \nu_{it} \tag{3}$$

where X_{it} is a vector of variables that contains age, personal net assets, demographic characteristics, job characteristics, and environmental factors. δ_i represents unobserved, time-invariant individual effect, assumed to be distributed according to $\delta_i \sim N(\mu, \sigma_\delta^2)$ while $\nu_{it} \sim (0, \sigma_\nu^2)$ reflects idiosyncratic measurement error. This specification (3) allows relative risk tolerance to vary with the covariates that potentially affect self-employment decision. The corresponding likelihood function for an individual i belonging to category j that responded to the survey in only one of the two years is,

$$P(c_{it} = j | X_{it}) = P(\log \underline{\theta}_j < \log \theta_{it} < \log \bar{\theta}_j) = \Phi\Big(\frac{\log \bar{\theta}_j - \mu - \beta X_{it}}{\sqrt{\sigma_\delta^2 + \sigma_\nu^2}}\Big) - \Phi\Big(\frac{\log \underline{\theta}_j - \mu - \beta X_{it}}{\sqrt{\sigma_\delta^2 + \sigma_\nu^2}}\Big) \tag{4}$$

When imputed values of risk tolerance (risk aversion) are used as proxy for risk preference, separating the effect of pure risk preference by accounting for correlation between risk tolerance (risk aversion) and other regressors are necessary to assess the effect of risk tolerance (risk aversion). Maximum likelihood is then used to compute estimates of $\mu, \sigma_\delta, \sigma_\nu$, and β. Having obtained estimates for these parameters, the predicted relative risk tolerance $\hat{\theta}_{it}$ is calculated for each individual using the category of risk she is in and covariates used in the ML estimation.

2.3 Tentative application: combining both methods

As noted by Sahm (2012), while systematic differences in risk tolerance across individuals over time exist and stem from time-invariant characteristics such as gender or ethnicity, more than 20% in the variation of risk tolerance originates from aging and changes in macroeconomic conditions (albeit short-lived). It is thus reasonable to abstract from the assumption of constant true risk preference that underlies Kimball et al. (2008, 2009) method. Their method is clearly still valuable for the separation of measurement error and status quo bias from survey data. Ahn (2010), on the other hand, abstracts from the assumption of constant true risk preference and allows risk preference to vary based on several characteristics conditioned on being in a category j. Her method, however, fails to account for the existence of status-quo bias present which could potentially affect the results found.

Since the NLSY79 has multiple waves of responses as well as different versions of questions, it is possible to separate true risk preference from measurement error, status-quo bias as well as the effects of aging and other determinants thereof. This can potentially achieved by augmenting (3) to allow for status-quo bias:

$$\log \theta_{it} = \beta X_{it} + \delta_i + b + \nu_{it} \tag{5}$$

I may then follow Ahn (2010) and estimates parameters for the series of responses contaminated by status-quo bias. The difference between the predicted values obtained by extrapolating estimates from the original

series and those obtained by estimating directly on the new series highlights status-quo bias.

References

Ahn, T. (2010). Attitudes toward risk and self-employment of young workers. *Labour Economics*, 17(2):434–442.

Barsky, R. B., Juster, F. T., Kimball, M. S., and Shapiro, M. D. (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *The Quarterly Journal of Economics*, 112(2):537–579.

Kimball, M. S., Sahm, C. R., and Shapiro, M. D. (2008). Imputing risk tolerance from survey responses. *Journal of the American statistical Association*, 103(483):1028–1038.

Kimball, M. S., Sahm, C. R., and Shapiro, M. D. (2009). Risk preferences in the psid: Individual imputations and family covariation. *American Economic Review*, 99(2):363–368.

Sahm, C. R. (2012). How much does risk tolerance change? *The quarterly journal of finance*, 2(04):1250020.