Countercyclical Risk Aversion—Evidence from 10,000,000 Auto Insurance Policies in China

Huafeng (Jason) Chen, Yinglu Deng, Pengfei Li, Hao Zhou*

Current Version: April 5, 2019

One leading explanation of various asset pricing puzzles is time varying risk aversion (Campbell and Cochrane (1999)). In bad times with high marginal utility, investors tend to become more risk averse. This countercyclical nature of risk aversion can simultaneously help to explain a large number of important yet puzzling facts: e.g., return predictability, high equity premium, and excess volatility, among many others. Such a feature also plays an important role in explaining the cross-section of stock returns (see, e.g., Zhang (2005)).

While this feature is clearly of theoretical importance, direct empirical evidence on whether risk aversion indeed behaves this way is scant and based on experimental or survey data. Cohn, Engelmann, Fehr, and Marechal (2015) study this question by priming financial professioals with either a boom or a bust scenario. Recently Guiso, Sapienza, and Zingales (2018) examine this question by using survey data on clients from one Italian bank.

Although evidence from experiments and survey data is helpful, such evidence is not a substitute for direct evidence from real transactions. In this paper, we examine this issue by estimating a large panel of individuals' risk aversion using data that are from real world, non-stock market transactions. In particular, utilizing China's comprehensive reporting of auto insurance policy data, we estimate risk aversion using auto insurance data following a seminal approach in Cohen and Einav (2007).[1]

*Preliminary and incomplete. Please do not circulate without permission. Comments solicited. Chen is at FISF, Fudan University. Deng, Li, and Zhou are from PBCSF, Tsinghua University. Please contact authors at chenh@fudan.edu.cn, dengyl@pbcsf.tsinghua.edu.cn, lipf@pbcsf.tsinghua.edu.cn, and zhouh@pbcsf.tsinghua.edu.cn. We thank China Insurance Information Technology Management (CIITC) for kindly providing us with the data. We are grateful for helpful comments and suggestions from Debbie Lucas and Liyan Yang.

[1]Cohen and Einav (2007) use one cross-section of information on auto insurance deductible choices and thus do not examine the dynamic, time-series properties of risk aversion.

We collect our data set from China Insurance Information Technology Management (CIITC), a government agency that aims to construct and operate a centralized and unified platform for sharing data and information in insurance in China. The data from the auto insurance platform of CIITC is comprehensive. We collect about 10 million auto insurance policies from 2011 to 2017, representing the entire population of 200 million auto insurance policies in China.

We follow a 3-step procedure for calculating the absolute risk aversion measure from the auto insurance data:

Step 1: Download and clean the original data, and then merge the multi-forms by distributed processing under the Hadoop MapReduce framework. The original data comprises seven distinct forms and over 260 million of items. In this paper, we process the data by five distributed computing servers with Apache Hadoop.

Step 2: After converting the data into policy-level feature data with 10 million of records, we randomly select 1% of them as the sample data for calculating the parameter matrix. The parameter matrix includes the limit distribution and the posterior distribution of claims and choices. In practice, we use the Markov Chain Monte Carlo (MCMC) algorithm, specifically the Gibbs sampling, to estimate the parameter matrix after 100,000 iterations in the sample data with the programming language of Julia.

Step 3: Calculate the absolute risk aversion by importing the full feature data into the imputed parameter matrix.

In our preliminary analysis, we find that the risk aversion imputed from the auto insurance data appear to be reasonable: males have higher imputed risk aversions than females; the imputed risk aversion varies nonmonotonically with individuals' age; the imputed risk aversion is negatively related to lottery activity in the region; and the imputed risk aversion is positively related to companies' book-to-market ratios.

More importantly, we find that risk aversion is indeed countercyclical. In our baseline specification with person fixed effects, a one-standard deviation decrease in GDP growth rate is associated with an increase in the relative risk aversion (RRA) coefficient of 1.31 (in the level regression), or a 10.84% increase in the RRA coefficient (in the log regression). To put these numbers into perspective, note that the mean and median RRA coefficients in our data sample are 11.062 and 6.305, respectively.

We believe our study utilize big data and represents a major step towards understanding better risk aversion of economic agents in the real world.