Investment Trios Are Less Prone to the Hot Hand and Gambler's Fallacies and Make Better Investment Strategies

Pamela Denice Arao

Danyel Brendan Arizabal

Seanne Veniene Esguerra

Cris Jon Quinalayo

Abstract

An experimental study was conducted to determine the minimum group size for which the mitigating effect for the hot hand and gambler's fallacies can be felt. This is quantified by looking if groups are as prone to the hot hand and gambler's fallacies in making decisions as their individual counter parts. Results suggest that groups maximize their investment returns better than individuals as the former choose to decide on their own more and rely on the experts' opinions less. Triads are the least biased with the hot hand and gambler's fallacies and thus are able to make more rational decisions and consequently maximize their investments better than the other treatments.These data allowed us to recognize the benefits of forming investment clubs consisting of three members since their decisions are more likely in line with the profit maximizing strategy in comparison with the decisions made by pairs and individuals.

Keywords: Investment, hot hand, gambler, profit

Introduction

The hot hand fallacy and the gambler's fallacy are two of the most common behavioral errors in the financial market. These biases cause people to misinterpret random sequences believing that some past event can be used to predict future outcomes. People who exhibit the hot hand fallacy expect an increasing trend to continue in the near future. This bias is observable when people mostly buy from funds who were successful in the past because they are convinced it would continue to be successful in the latter periods (Sirri & Tufano, 1998). On the other hand, people who exhibit the gambler's fallacy expect a current trend to 'break and reverse' in the future. This bias is observable when people buy losing stocks, which are stocks who have recently declined in prices, because they expect a reversal of their losses later. However, the efficient market hypothesis (Fama, 1991) showed that trends in the prices of stocks are unreliable to consumers because the stock prices instantaneously adjusts to new information by taking into account the discounted future value of that information. Rufino (2013) demonstrated that the Philippine stock market exhibits this market efficiency and thus shows a random walk phenomenon. Therefore, it is unreliable to use past information to predict future prices.

Groups tend to decide more optimally than individuals in both strategic and nonstrategic situations (Feri et al., 2010; Linder & Morgan, 2005). Despite this optimality, both groups and individuals alike still fall prey to behavioral biases. Investors who are exposed to these fallacies end up holding less diversified portfolios, which can affect their expected returns and exposures to risk. However, Stöckl et al. (2015) pointed out that groups are actually less prone to the hot hand fallacy and gambler's fallacy as compared to individual investors. Thus, we are looking at a possible 'mitigation effect' on the overall proneness to behavioral biases by making decisions as a group. With this, we can say that there is a so-called wisdom in groups. Yet, the desired group size in achieving this mitigation effect has not yet been established.

To address these problems, we use a two-dimensional analysis on investment behavior for our objectives on 1) the investment strategy and 2) the behavioral biases that affect these decisions. Generally, we say that individuals and groups make investment decisions differently. However, since the scope of the differences is broad, we narrowed it down to two research questions that are aligned with the objectives. The first research question focuses on the investment strategy; we ask ourselves "Do individual and group investors make investment decisions differently? How reliant are they on experts and who chooses the riskier option more?" The second research question focuses on behavioral biases; we ask "Can we mitigate the overall proneness to the hot hand and gambler's fallacies by making decisions as a group? If so, how large does a group need to be in order to feel this mitigating effect?"

In the perspectives of the investors, this study will be able to aid them in critical investment decisions. As they become more aware of their proneness to these behavioral biases when investing their money, individual investors can look for ways to reduce their exposure to the fallacies by joining investment groups that satisfy the minimum group size in order to experience the mitigation effect. Making investment decisions as a group can be thought of as "investment clubs" where non-professional investors combine their investable wealth through a partnership or a limited liability company, make investment decisions together, and split any earnings among themselves. Investment clubs differ from mutual funds since the latter is funded by shareholders who waive their rights to manage the portfolios to a professional fund manager. In the perspective of the academe, this research shall fill the gaps in the current literature by establishing the minimum group size in order to reduce the overall proneness to the behavioral biases.

This study will make use of a coin-toss investment simulation model where participants aim to correctly predict which side of the coin will appear. Their investment decisions will determine their investment strategy and overall proneness to the hot hand fallacy and gambler's fallacy. The experiment proper was conducted in De La Salle University- Manila during the third term of the Academic Year 2017-2018. Market investors will be represented by bonafide undergraduate students ages 18 to 23 years old. The simulation will be divided into three treatments (INDIV, GROUP2, GROUP3) with a total of 180 participants for which we have gathered 3,600 decisions for analysis. In this study, only pairs and trios were tested for group classifications. Therefore, we can only provide evidence for the minimum group size for the mitigation effect to be felt instead of the boundary condition that provides the optimal group size with the greatest mitigation effect.

Currently, there have been numerous studies that revolve around the hot hand and gambler's fallacies. Fischbein (1975) showed evidence of the hot hand fallacy when successive outcomes of heads would lead individuals to believe that the probability of another head appearing to increase. On the other hand, Tversky and Kahneman (1971) showed that when three successive heads appear, people would infer that the next outcome is a tail as a manifestation of the gambler's fallacy. Further studies by Stöckl et al. (2015) showed that individuals and group alike were both prone to the hot hand and gambler's fallacies. However, O'Leary (1993) was able to discover that groups are less prone to biases that occur naturally in individuals. Therefore, the overall proneness of groups to the hot hand and gambler's fallacies were significantly less than individuals. However, the desired group size to feel the mitigation of the proneness to fallacies has not yet been thoroughly studied.

Since some investment decisions can be done in a group environment, Simmel (1950) and Weick (1969) noted the importance in determining the optimal group size that can mitigate the existence of these fallacies, especially in smaller groups. When it comes to the crucial transitions in group sizes, which consists one to two, two to three, three to four, four to seven, and seven to nine members, trios are one of the most crucial group sizes. In organization theory, Weick (1969) refers to groups of three as the basic unit of analysis since it is the smallest possible group size that allows two group members to be allies against one. Groups of three allow for cooperation, control and competition.

For larger group sizes, another group of three can be formed as subgroups. Since groups of three are still considered when it comes to forming larger groups, this group size will be the focus of the paper. O'Leary (2011) performed an experiment comparing groups of three and individuals and discovered that groups mitigate the effect of these biases but do not completely remove its impact, as the performance of groups was not perfect.

The wisdom of crowds is an event where a group of people make better decisions compared to individuals who are experts in a particular subject matter (Surowiecki, 2005). Crowds tend to be wiser than individuals but too many members in one group can lead to herd behavior which causes poor decision-making. Furthermore, Goldstein, Mcafee, and Suri (2016) applied the wisdom of crowds in a experimental setting where a smaller group of 30 people were grouped together from a crowd composed of 100 members and compared the groups' decisions against an expert in Fantasy soccer. They discovered that the smarter and smaller crowd beat the wisdom of the larger crowd.

To fulfill our first objective, we try to identify the different investment strategies by using a probit model on the three treatments. The probit model was chosen because the dependent variable is either 1 or 0. The regression results would allow the researchers to assess the likelihood of choosing between the RISKown, RISKexpert, and RISKfree options.

Specifically, the probit regression coefficients would determine the relationship of the dependent and independent variables. Additionally, we use the marginal effects of the coefficients to determine the respective likelihood of choosing each of the three options mentioned earlier. For additional robustness checks on the results provided by the individual probit model, we also run a multinomial probit model to take into account the fact that subjects can only choose one out of three options.

On the other hand, to fulfill the second objective regarding behavioral biases, we used a tobit regression model since some of the dependent variables are unobservable. This means that if the Classical Linear Regression Model (CLRM) is to be used, it would drop the unobservable dependent variables, therefore making the results unreliable by omission. Instead, we use the tobit model in order to include and still consider these unobservable dependent variables. Furthermore, since the experiment was designed in such a way that the subjects do not have any choice but to invest, it is considered to be a single hurdle process instead of a double hurdle process which makes the tobit model more superior to both the craggit and heckit models.

Theoretical models such as the rational choice theory shows that individuals are rational in making their decisions by knowing which option they prefer and going for the option that would give them the highest utility. On the other hand, the prospect theory accounts for decision-making behavior under risk and uncertainty. As individuals' have successful experiences, they tend to prefer risk-free assets over risky assets. Consequently, when these individuals are experiencing losses, they tend to invest more on risky assets than risk-free ones (Kahneman and Tversky, 1979). However, exposures to the hot hand fallacy and gambler's fallacy lead to suboptimal outcomes and decisions.

Aside from confirming past literatures which stated that groups are less prone to behavioral biases as compared to individuals, the research gap this study aims to bridge is the establishment of the minimum group size in mitigating the proneness to the hot hand fallacy and gambler's fallacy. We also analyzed the differences in decision-making between individuals and groups of varying numbers particularly in pairs and trios as well as exploring any gender effects. For the first research question on investment strategy, we test the hypothesis that groups rely less on experts and choose riskier options as compared to individuals. For the second research question on behavioral fallacies, we test the hypothesis that groups are less prone to the hot hand and gambler's fallacies, and there exist a minimum group size for the mitigation effect.

2. Experimental Design

Sample. We gathered 180 undergraduate students (92 males and 88 females) through an online signup link shared on social media. Participants aged 18-23 are all from De La Salle University-Manila. The participants are divided into three treatments: individuals (INDIV), groups of two (GROUP2), and groups of three (GROUP3).

Task. Throughout the experiment, the subjects had to decide on their own whether they wanted to choose heads or tails by themselves (RISKown), delegate the decision to an expert (RISKexpert), or choose the risk-free alternative (RISKfree) for 40 periods.

Design. The experiment was conducted in a computer laboratory setting using Google Forms and Google Sheets as the online softwares. At the start of the experiment, each participant was given an initial endowment of 500 Taler, which is the experimental currency. The RISKfree alternative guaranteed the participants 5 Taler regardless of the outcome of the coin toss. The risky alternative is replicated using a coin toss where heads and tails have equal probabilities. When choosing the risky alternative, the participants need to select one side of the coin, and the goal is to correctly predict the outcome of the coin toss to receive a positive payoff. The participants have two options when deciding for the risky alternative. First is by choosing RISKown where they have to guess on their own whether the outcome of the coin toss is heads or tails, or second, by choosing RISKexpert where they delegate the decision to one of the five "experts", who will then randomly select heads or tails for the participants. However, participants are not informed that the experts are mere randomizers. We use the same coin toss realizations, which were drawn randomly in advance, for every session to ensure comparability across all observations.

For every correct RISKown decision, 100 Taler will be added while 50 Taler will be deducted if otherwise. For RISKfree, 5 Taler will be immediately added to their accounts. For RISKexpert, a management fee of 5 Taler will be deducted to the payoff regardless of the outcome of the coin toss. 95 Taler will be added if the expert made the correct decision while 55 will be deducted if otherwise.

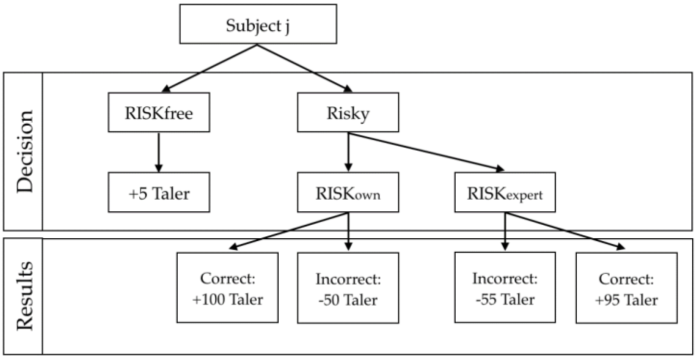


Figure 1. Flowchart of the entire experimental design

Before the experiment, the subjects were informed that three screens will be used throughout the 40 periods. The first screen is the decision screen where the participants' decisions will be inputted for each and every decision made. The second screen is the results screen where the summary of all their decisions, their successes and failures in predicting the outcomes, the respective payoffs in every period, and the running balance of their accounts are shown. The third screen is the trading screen which contains information on the current period, the results of the coin toss for every period including all past periods, and the successes and failures of the experts in predicting the results of the immediate previous four periods. The track record of the experts is updated per period and can be accessed at any time during the experiment.

For the INDIV treatment, participants decided on their own were only needed to make one decision per period. They were not allowed to seek help from other individual participants. For the group treatments, all members of the group had to collectively decide on which decision to make every period. Communication was allowed but only within the group members. Each member had separate accounts for the group's total balance which meant that the earnings were not divided equally and that their total balances were paid to each member at the end. They also had separate screen which means that each group member still had to enter his decisions in his decision screen. Groups only have a chance of getting a positive payoff if all members choose the same decision. In the event that they do not come up with the same decision in their first try in any period, they are given a second chance to decide and come up with a unanimous decision. A penalty of 50 Taler will be deducted from their individual accounts in the event that their decisions are still inconsistent. A total of 3,600 observations (1,200 per treatment) were gathered at the end of the experiment.

For the hypotheses in both research questions, we consider the INDIV treatment as the control group. All characteristics of INDIV are present in the GROUP2 and GROUP3 treatments including the assumption of common knowledge in investing and optimization. However, the variable characteristics of communication and joint decisions are only present in the GROUP treatments.

Payouts. At the end of the experiment, the total Taler balances were converted to Philippine Peso using an exchange rate [1 PhP=15 Taler(for INDIV and GROUP2) and 20 Taler(for GROUP3)]. Cash was immediately paid after the experiment. In addition, a participation fee of PhP 20 as a compensation for their time was also added.

3. Results

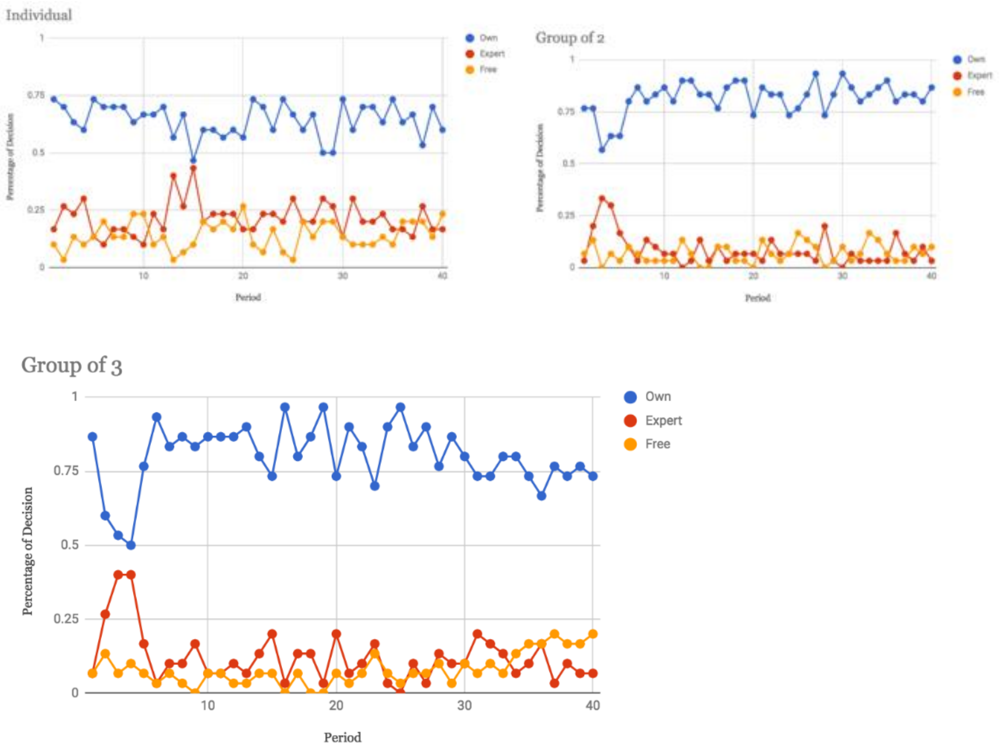


Figure 2. Proportion of total decisions per period allocated to RISKown, RISKexpert, and RISKfree options.

Did groups rely more on experts than individuals? Did they choose riskier options as compared to the risk-free options?

The proportion of total decisions allocated to RISKown, RISKexpert, and RISKfree options with respect to each of the treatments (INDIV, GROUP2, GROUP3) is shown in figure 2. INDIV participants have a higher reliance on the experts and choose the RISKfree options more rather than making the decisions by themselves as compared to both GROUP2 and GROUP3 participants. We found from running a Mann Whitney U-test that there are significant differences between the samples. For the decisions made by the participants themselves (RISKown), there is clearly gender and group effects that differentiates the treatments. For decisions delegated to experts (RISKexpert), *FF*, *MMM*, *FMM*, and *FFF* are derived from the same population. For decisions involving no risk (RISKfree), only *FMM* and *FFF* come from the same population.

Table 1: Probit Regression Model 1

	RISKown				RISKexpert				RISKfree			
Variable	Probit Coefficient		Marginal Effect		Probit Coefficient		Marginal Effect		Probit Coefficient		Marginal Effect	
Constant	0.2469	***	0.7517	***	-0.4266	***	0.1393	***	-1.3624	***	0.1055	***
Group of 2	0.3489	***	0.1057	***	-0.7073	***	-0.1479	***	0.2009	*	0.0356	*
Group of 3	0.9863	***	0.2989	***	-1.1689	***	-0.2445	***	-0.3337	**	-0.0592	**
Period	0.0023		0.0007		-0.0089	***	-0.0019	***	0.0066	***	0.0012	***
F	0.1373	*	0.0416	*	-0.3727	***	-0.0780	***	0.2657	***	0.0471	***

FF	0.4120	***	0.1249	***	-0.1062		-0.0222		-0.5642	***	-0.1001	***
FM	0.3284	***	0.0995	***	-0.1124		-0.0235		-0.3551	***	-0.0630	***
FFM	-0.7161	***	-0.2170	***	0.9738	***	0.2037	***	0.1105		0.0196	
FMM	-0.4634	***	-0.1404	***	0.5135	***	0.1074	***	0.2877	**	0.0510	**
FFF	-0.4381	***	-0.1328	***	0.6852	***	0.1433	***	0.0007		0.0001	
*** - Significant at 99% confidence level												
** - Significant at 95% confidence level												
* - Significant at 90% confidence level												

We then run a probit regression model (Model 1) to predict the probability that treatments would select either RISKown, RISKexpert, or RISKfree. From Model 1, we find evidence that GROUP2 and GROUP3 are .1057 and .2989 more likely to make the decisions on their own (RISKown) than INDIV. Female individuals are .0416 more likely to take the risk on their own than their male counterparts. Female pairs are also more likely to choose RISKown than a male dominated pair or a mixed pair. An all-male trio is most likely to make the decision on their own than the *FFF*, *FFM*, and *FMM* counterparts. Only *Period* was the insignificant variable for the RISKown. For delegating the decisions to experts, we find that participants under the GROUP2 and GROUP3 treatments are less likely to delegate the decision by .1479 and .2444 respectively. Females are also .0779 less likely to delegate the decision than males. Female-male pairs and all-male trios are the least likely to delegate the decision to an expert. For choosing the RISKfree option, GROUP2 are .0356 more likely to chose the risk-free option while GROUP3 are less likely to choose this. Female individuals also prefer the risk-free options than male individuals. On the other hand, females in pairs are least likely to choose risk-free than male-male or female-male participants. All male trios are the least likely to choose the risk-free option.

Table 2: Probit Regression Model 2												
	RISKown				RISKexpert				RISKfree			
Variable	Probit Coefficient		Marginal Effect		Probit Coefficient		Marginal Effect		Probit Coefficient		Marginal Effect	
Constant	0.3456	***	0.7517	***	-0.5971	***	0.1394	***	-1.3170	***	0.1056	***
Group of 2	0.0749		0.0227		-0.3959	**	-0.0826	**	0.3114	*	0.0551	*
Group of 3	0.9401	***	0.2843	***	-0.9092	***	-0.1897	***	-0.6400	***	-0.1132	***
Period	-0.0024		-0.0007		-0.0006		-0.0001		0.0045		0.0008	
F	0.1367	*	0.0413	*	-0.3695	***	-0.0771	***	0.2649	***	0.0469	***
FF	0.4164	***	0.1259	***	-0.1118		-0.0233		-0.5611	***	-0.0992	***
FM	0.3301	***	0.0994	***	-0.1175		-0.0245		-0.3499	***	-0.0619	***
FFM	-0.7156	***	-0.2164	***	0.9883	***	0.2062	***	0.1138		0.0201	
FMM	-0.4641	***	-0.1404	***	0.5217	***	0.1089	***	0.3038	**	0.0537	**
FFF	-0.4401	***	-0.1331	***	0.6917	***	0.1444	***	0.0118		0.0021	
Group2Period	0.0136	***	0.0036	***	-0.0158	***	-0.0033	***	-0.0055		-0.0010	
Group3Period	0.0022		0.0006		-0.0135	**	-0.0028	**	0.0133	**	0.0024	**
*** - Significant at 99% confidence level												

** - Significant at 95% confidence level
* - Significant at 90% confidence level

We then run another probit regression model (Model 2) for all three decisions with the addition of interacting variables *Group2Period* and *Group3Period* to account for the learning differences between the respective treatments. For the RISKown, the *Group2Period* indicates that the interacting variable increases the probability of choosing the RISKown by .0036. Likewise, both the interacting variables lead to a significant decrease in the probability of delegating the decision to an expert (RISKexpert). However, the *Group3Period* interacting variable leads to a .0023 increase in the risk-free option as well.

Table 3: Multinomial Probit Regression Model 1												
	Risk Own				Risk expert			Risk free				
Variable	Probit Coefficient		Marginal Effect		Probit Coefficient	Marginal Effect		Probit Coefficient		Marginal Effect		
Constant	0.4676	***	0.7540	***	Base Outcome	0.1401	***	-0.9817	***	0.1059	***	
Group of 2	0.9290	***	0.1142	***	Base Outcome	-0.1475	***	0.9569	***	0.0332	*	
Group of 3	1.7068	***	0.2985	***	Base Outcome	-0.2437	***	0.8977	***	-0.0548	**	
Period	0.0105	***	0.0007		Base Outcome	-0.0019	***	0.0169	***	0.0012	***	
F	0.4508	***	0.0354		Base Outcome	-0.0777	***	0.6585	***	0.0423	***	
FF	0.2927	*	0.1199	***	Base Outcome	-0.0206		-0.5133	**	-0.0993	***	
FM	0.2631		0.0875	***	Base Outcome	-0.0248		-0.2641		-0.0627	***	
FFM	-1.3684	***	-0.2142	***	Base Outcome	0.2031	***	-0.9639	***	0.0111		
FMM	-0.7643	***	-0.1474	***	Base Outcome	0.1049	***	-0.2610		0.0425		
FFF	-0.9400	***	-0.1347	***	Base Outcome	0.1433	***	-0.7825	**	-0.0086		
*** - Significant at 99% confidence level												
** - Significant at 95% confidence level												
* - Significant at 90% confidence level												

Table 4: Multinomial Probit Regression Model 2											
	Risk own				Risk expert			Risk free			
Variable	Probit Coefficient		Marginal Effect		Probit Coefficient	Marginal Effect		Probit Coefficient		Marginal Effect	
Constant	0.7001	***	0.7540	***	Base Outcome	0.1401	***	-0.7429	***	0.1059	***
Group of 2	0.4549	**	0.0298		Base Outcome	-0.0800	**	0.7219	***	0.0502	*
Group of 3	1.3993	***	0.2927	***	Base Outcome	-0.1843	***	0.2629		-0.1084	***
Period	-0.0009		-0.0008		Base Outcome	-0.0001		0.0054		0.0008	
F	0.4459	***	0.0351		Base Outcome	-0.0766	***	0.6504	***	0.0416	***
FF	0.3014	*	0.1205	***	Base Outcome	-0.0220		-0.5028	**	-0.0985	***
FM	0.2691		0.0874	***	Base Outcome	-0.0259		-0.2529		-0.0616	***

FFM	-1.3892	***	-0.2153	***	Base Outcome	0.2061	***	-0.9938	***	0.0092	
FMM	-0.7726	***	-0.1497	***	Base Outcome	0.1054	***	-0.2612		0.0443	**
FFF	-0.9477	***	-0.1366	***	Base Outcome	0.1438	***	-0.7787	**	-0.0073	
Group2Period	0.0238	***	0.0042	***	Base Outcome	-0.0034	***	0.0119		-0.0008	
Group3Period	0.0161	**	0.0006		Base Outcome	-0.0030	**	0.0302	***	0.0024	**
*** - Significant at 99% confidence level											
** - Significant at 95% confidence level											
* - Significant at 90% confidence level											

To confirm the results of the individual probit models, we ran a multinomial probit regression model for both the Model 1 and the Model 2. The multinomial probit models take into account that the subjects can only choose one out of the three alternatives. Therefore, the three alternatives are mutually exclusive. The results produced by the different probit models yield the same interpretation. Groups are still less risk averse than individuals with GROUP3 being the least risk averse. They are the most inline with the profit maximizing strategy. Similar with the individual probit regression models, we could see that both GROUP2 and GROUP3 rely less on experts overtime. Lastly, for additional robustness checks, we also ran several logit regression models (see Appendix D). We can see that both the probit and the logit regression models yield the same result and interpretations, hence we can conclude that the results produced by the two models are robust.

Generally, we see that there are differences among the investment decisions between individuals and groups. INDIV has a greater inclination of relying on experts and choosing the risk-free option than groups. Particularly, male individuals delegate the decisions more than their female counterparts while female individuals have a greater preference for the risk-free option. On the other hand, GROUP2 and GROUP3 are more geared towards making the decisions by themselves. Female pairs and male trios make the decision on their own more frequently than their pair and trio counterparts. Female pairs and all male trios are least likely to select the risk-free option. In terms of the existence of a learning curve, only GROUP2 had a significant learning curve.

Table 5. Tobit Regression Models												
	Hot hand Fallacy						Gambler's Fallacy					
Variables	INDIV		GROUP2		GROUP3		INDIV		GROUP2		GROUP3	
Constant	-0.0108		-0.0188		-0.0133		0.0235	***	0.0210	***	0.0217	***
Streak 0	0.0754	***	0.0723	***	0.0674	***	-		-		-	
Streak 1	0.0613	***	0.0604	***	0.0605	***	-		-		-	
Streak 2	0.0378	**	0.0412	**	0.0549	***	-0.0179	***	-0.0121	***	-0.0138	***
Streak 3	-		-		-		-0.0275	***	-0.0203	***	-0.0196	***
Streak 4	0.0497	**	0.0782	***	0.0306	*	-0.0374	***	-0.0273	***	-0.0257	***
Streak 5	0.0613	**	0.1079	***	0.1035	***	-0.0345	***	-0.0259	***	-0.0262	***
Streak 6	0.0380		0.0980	***	0.0897	***	-		-		-	
*** - Significant at 99% confidence level												
** - Significant at 95% confidence level												
* - Significant at 90% confidence level												

Does the Hot Hand Fallacy exist in both individuals and groups?

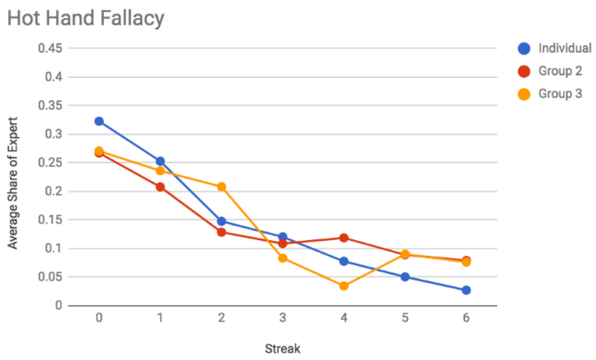
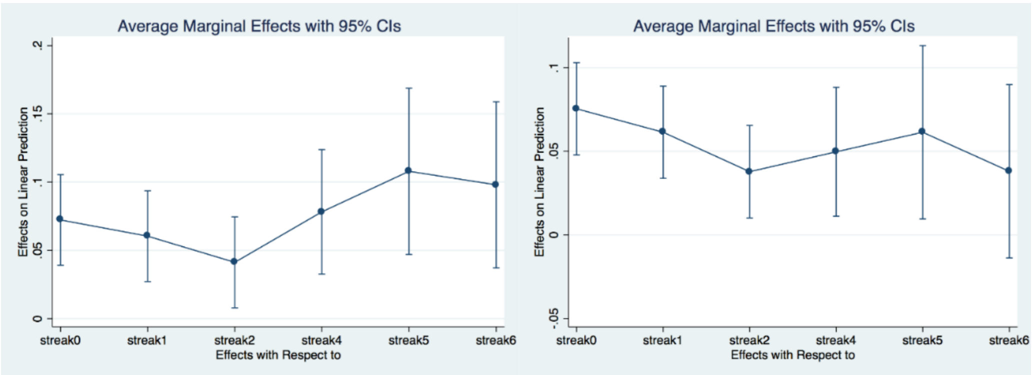
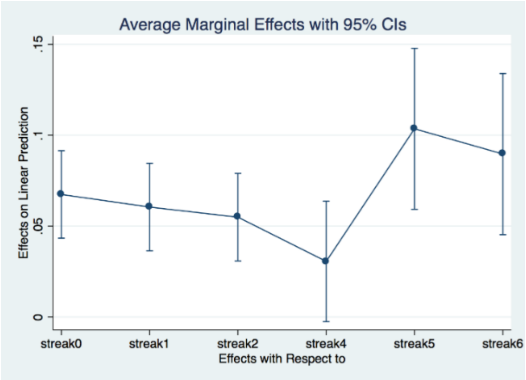


Figure 3. Average share of Expert against the streak of correct predictions

Figure 3 provides a graph that shows the percentage of decisions that are delegated to the experts conditional on the streaks of successful expert decisions. We plot the average share of decisions that expert *i* gains from all decisions under RISKexpert against his recent streaks of successful predictions. An unbiased decision by the market participants would lead to each expert gaining ⅕ of all decisions since we have a total of five experts in the industry. When a decision is unbiased, the past performances by the experts should not matter. We have observed and discovered a distinctly different pattern for the hot hand fallacy as compared to previous literatures which had a steady increasing trend on the percentage decisions delegated to the experts as the streaks increased (Stöckl et al., 2014). In all treatments, one can easily see the percentage share of decisions that expert *i* loses despite his increasing correct predictions during lower streaks (below streaks of 3) while having an increasing percentage share for higher streaks (above streaks of 3). Since the participants are also given the experts' success rates from negative four periods back before the game actually started, it can be observed that the reliance on experts is quite high when there has been no streaks yet. The declining trend can be attributed to the participant's analysis and updating of beliefs regarding the experts ability since they are not explicitly aware of how experts make their decisions. The hot hand fallacy is observed when the treatments relied on the experts once again from the moment they have reached a certain number of correct predictions.



a) Individual b) Group2



c) Group3

Figure 4. Average Marginal Effects for the Hot Hand Fallacy Tobit Regression

These were verified by running a tobit regression model where a streak of three correct predictions is the base dummy variable. For INDIV, there is a significant declining trend in the percentage of decisions delegated to expert *i* during periods of streaks of 0, streaks of 1, to streaks of 2 correct predictions. An expert *i* has a higher percentage share by 0.0754 when there is no streak while there is a lower percentage share for streaks of 1 and 2. However when expert *i* experiences a streak of 4 and streak of 5 only, his percentage share of decisions increased. For GROUP2, the same declining trend is significant except only until streaks of 2. An expert *i* has a higher percentage share by 0.0722 when there is no streak yet while his percentage share decreases for streak of 1 and streak of 2. When expert *i* experiences a streak of 4, streak of 5, and streak of 6, there is a higher percentage share increase by .0782, .1079, and .0980 respectively. For GROUP3, the declining trend is until streaks of 4. The percentage share increases by .1035 for streak of 5 although it declines to .0897 for streak of 6. We have observed that GROUP2 is the most prone to the hot hand fallacy.

Does the Gambler's Fallacy exist in both individuals and groups?

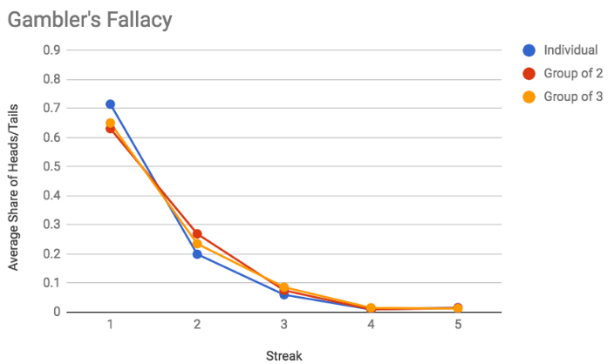
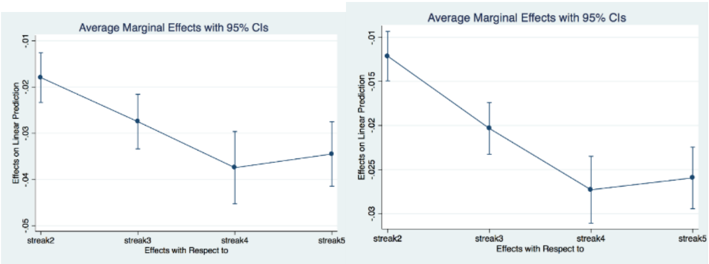
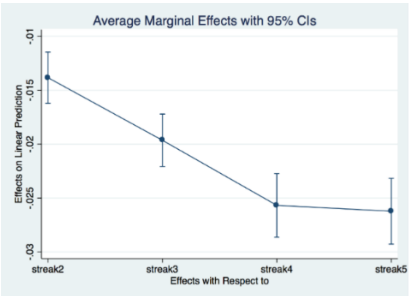


Figure 5. Average share of Heads/Tails against the frequency of appearance on past realizations/Streak.

Figure 5 shows the graph for the average frequency of choosing heads or tails depending on the streaks of head/tail realizations drawn immediately before from all decisions under RISKown. An unbiased behavior by the market participants should yield half of all RISKown decisions per each side of the coin. When a decision is unbiased, the past realizations of the coin should not influence the decision of the participants. In all treatments, we can see a declining trend in the average share of the side of the coin relative to the frequency of its appearance from previous realizations. During the first appearance of the coin side from previous realizations, if the participant selects the same side it is considered as a streak of 1. If the same side appears again and the participant selects the same side, it is considered as a streak of 2 and so on. During a streak of 1, all treatments would still opt to choose the same side as evidenced by the high rate of average share of the head/tail side. However, after a streak of two consecutive appearance, the average share of the coin side declines for all treatments with INDIV having the most decline. We can observe the gambler's fallacy when a side of a coin is chosen less frequently after the same side appeared in a streak of several identical realizations.



a) Individual b) Group2



c) Group3

Figure 6. Average Marginal Effects for the Gambler's Fallacy Tobit Regression

We run a tobit regression to verify the results with the dependent variable as the percentage share of the each side relative to the streaks of its previous appearances with the streak of 1 as the base dummy variable. For all treatments, the coefficients are getting more negative as the streaks increases. This is an indication of the gambler's fallacy as the percentage share of the coin side decreases as it appears more frequently in the previous realizations. For INDIV, the participants would have an average percentage share of a particular coin side of .0235 for a streak of 1. When the streak increases to 2, the percentage share of that coin side decreases by -.0179. When the streak increases to 3, the percentage share decreases further by -.0275. The greatest percentage share decrease happens in a streak of 4 with a decrease of -.0374. For GROUP2, the participants would have an average percentage share of .0210 for a streak of 1. The percentage share decreases by -.0121 when the streak increases to 2. The percentage share further declines by -.0203. The greatest percentage share decrease still happens in a streak of 4 by -.0275. For GROUP3, the participants would have an average percentage share of a particular coin side of .0217 for a streak of 1. When the streak increases to 2, the percentage share of that coin side decreases by -.0138. When the streak increases to 3, the percentage share decreases further by -.0196. Unlike the other treatments, GROUP3 has the highest percentage decrease during a streak of 5.

Table 6. Treatment Average Earnings (in Taler)

Treatment	Average Taler Earnings
INDIV	1,258.83
GROUP2	1,323.83
GROUP3	1,507.00

GROUP3 has the highest average earnings while INDIV has the lowest. If an individual decides to become part of a pair or a trio, he would receive a marginal earning of 65 Taler by being in pairs or 248.17 Taler by being in trios on average. If pairs were to include one more member in the group to become a trio, each member of the group could have earned an additional 183.17 Taler on average.

4. Discussion

In terms of investment strategy, we find that groups make investment decisions differently from individuals. Particularly, they prefer to make the investment decisions on their own, not rely on experts, and choose the risk free option less. We also looked into some gender differences. In terms of the behavioral fallacies, we find that both groups and individuals alike are still prone to the hot hand fallacy and gambler's fallacy, but groups are marginally less prone indicating that there is a mitigation effect by being in a group. We find that trios are the least prone to the both biases, but does not completely eliminate it.

Our study presents new findings in both the hot hand fallacy and gambler's fallacy. We discovered pairs are the most prone to the hot hand fallacy while individuals are the most prone to the gambler's fallacy unlike in Stockl et. al (2015) where pairs are less prone to both fallacies as compared to individuals. We also discovered a different trend for the hot hand fallacy where participants actually chose the experts less during lower streaks of correct predictions (below streaks of 3) and relied on them during periods of higher streaks only (above streaks of 3) unlike in previous literatures which observed an upward trend in the share among all experts when their streaks increased even at low levels (Stockl et al., 2015)

A key assumption of the study is that participants possess common knowledge in maximizing returns. Therefore, using students as participants for the study is deemed to be sufficient as using actual investors. We give particular importance on the design of the experts in the experiment. It may be criticized that their expertise is unconvincing since the study only has two outcomes, thus the probability for correctly predicting the coin toss by oneself is relatively high. This indicates that convincing the participants that the experts possess expertise in an environment where one outcome has a 50% probability of happening may have posed a challenge. Additionally, the researchers recognize that the pricing of the services offered by an expert could be further improved by dividing the cost into fixed and variable fees, especially since this represents a more accurate picture of how experts are in reality. Lastly, the researchers recommend that further studies expand the group size in order to discover the boundary condition or the optimal group size with the greatest mitigation effect to the hot hand and gambler's fallacies. This study only focused on discovering the minimum group size where the mitigation effect begins.

To address our first research question on investment strategy, we tested the hypothesis that groups rely less on experts and choose riskier options as compared to individuals. Our main finding is that groups are far better optimizers of their investment returns than individuals since the latter had chosen more RISKexpert and RISKfree options, which provided a lesser payoff as compared to choosing the RISKown option.

In terms of a profit maximizing strategy in decision making, we look at the RISKown option. We discovered that groups make investment decisions closer to the expected return maximization than individuals since RISKown decisions yield higher returns than RISKexpert and RISKfree. Both GROUP treatments chose the RISKown option more and the RISKexpert and RISKfree options less. However, GROUP3 has the best profit maximizing strategy among all treatment from the Probit Regression Model 1. Conversely, this signifies that INDIV strays from the payoff maximizing strategy because they relied more on the experts and selected the guaranteed return more often.

In terms of expert reliance, we look at the RISKexpert option. It is important to remember that the so-called "experts" are mere randomizers, but the participants are not made aware of this. By virtue of their titles as experts, participants had relied on their predictions and paid a certain fee. We found that both GROUP treatments had relied less on the experts as evidenced by their negative significant coefficients under the regression model with GROUP3 relying on the experts the least. On the other hand, we find that INDIV had the highest reliance on experts. This implies that INDIV prefers to be less in control with the outcome of their decisions because they have higher proportions of decisions delegated to RISKexpert and more RISKfree options. Additionally, we find that INDIV had steadily relied on the experts throughout the duration of the experiment. There is no period where INDIV participants did not invest in an expert.

In terms of risk aversion, we look at the RISKfree option. The higher frequency of RISKfree decisions indicates a higher level of risk aversion because the RISKfree option guarantees a small return with no possibility of a loss. INDIV selected the RISKfree option the most and is thus the most risk averse. We find that both GROUP treatments have a lower risk aversion because they chose the RISKfree option less. However, GROUP3 has the lowest risk aversion in comparison to GROUP2. During the last five periods of the experiment, it is interesting to note that all treatments selected the RISKfree option more because they were trying to shield their earnings from any possible losses.

Since the experts in the experiments were mere randomizers, we explored the possibility whether the treatments would realize that there is in fact no expert opinion. This is made possible by adding an interacting variable *Period* to the treatment variables. We find that *Group2Period* and *Group3Period* are actually significant with negative marginal effects indicating that the likelihood of delegating their decision to an expert actually decreases for every period. As time progresses, GROUP2 and GROUP3 treatments relied less on the experts than their INDIV counterparts. This suggests that both GROUP treatment exhibits a learning curve while INDIV does not. By learning curve, we imply that they are slowly beginning to realize that delegating the decision to an expert will not give them the maximum results and that the experts are merely randomizing their predictions.

When we decompose the investment strategy by gender differences, we show significant results for RISKown, RISKexpert, and RISKfree. In terms of RISKown, female INDIV, female GROUP2, and all-male GROUP3 are the most likely to make the decisions on their own respectively. They prefer being in control of the outcome of their decisions and employ the most profit maximizing strategy. In terms of RISKexpert, female INDIV, female-male GROUP2, and all-male GROUP3 are the least likely to delegate the decisions to the experts respectively. They show greater skepticism in the ability of the experts. In terms of RISKfree, female INDIV, male GROUP2, and female-male-male GROUP3 prefer the RISKfree option the most.

To address our second research question on behavioral fallacies, we tested the hypothesis that being in groups can lessen the overall proneness to the hot hand and gambler's fallacies, and there exist a minimum group size for the mitigation effect. Our main finding is that while both individuals and groups are prone to the hot hand fallacy and gambler's fallacy, groups are less influenced as compared to individuals. Specifically, GROUP3 is the least prone to both fallacies.

In analyzing the existence of the hot hand fallacy, we noticed a difference from previous literatures that have observed an increasing trend as the streaks increased (Stöckl et al., 2015). In our results, we noticed a different general pattern of behavior. Initially, experts are chosen more frequently when they correctly predict the most recent previous period as compared to when they correctly predict the outcome of two and three periods back. Additionally, they are again selected more frequently after they have reached a certain number of correct streaks. Therefore, we observed a declining trend in choosing an expert during lower streaks (streaks below 3), but an increasing trend during higher streaks (streaks above 3). The hot hand fallacy is only observed for all treatments when they select an expert after that expert has reached a certain number of correct predictions.

In all treatments, the high frequency of choosing an expert when they predicted the most recent period correctly can be attributed to what we refer to as the "recency effect." Subjects are more influenced by the latest information presented to them when faced with a list of information for immediate free recall (Baddeley and Hitch, 1993). The decline in the frequency during correct streaks of 2 and 3 can be attributed to the wariness of the subjects regarding the expert's ability. Since they were not explicitly informed about how experts come up with their predictions, the subjects are observing the experts and updating their previous beliefs regarding the expert's abilities. It is only when the experts reach a certain number of correct predictions that the subjects begin to select the experts once again. This is an indication of the hot hand fallacy similarly observed by Stock et. al (2015). We find that INDIV and GROUP2 are more prone to the hot hand fallacy because they began to select the experts more frequently once again for streaks of 4 and 5. We consider GROUP3 as the least prone to the hot hand fallacy because they only selected the experts again after a streak of 5.

In analyzing the gambler's fallacy, we observed that the frequency of a particular coin side being selected decreases as the streaks of the opposite side increased. This is evidenced by the negative coefficients in the regression results. As the streak used in the regression increases, all coefficients become more negative, which is a clear indication of the proneness to the gambler's fallacy. This is in line with the findings by Stöckl et al. (2015) where the frequency of selecting the opposite outcome increases as streaks of the same side realized increased.

INDIV is the most prone to the gambler's fallacy among all treatments since it has the most negative coefficients across all streaks and the greatest incremental change between streaks. Consequently, GROUP3 is the least prone to the gambler's fallacy as it has the lowest incremental change in coefficients between streaks.

After analyzing the treatments, we find evidence that groups are superior to individuals in terms of their investment strategy. Specifically, we find that GROUP2 shows significant learning curves on both RISKown and RISKexpert decisions while GROUP3 only has a significant learning curve on RISKexpert decisions. INDIV has no significant learning curves on both RISKown and RISKexpert. Comparing the magnitude of the learnings curves for RISKexpert, GROUP3 has a faster learning rate since it has a greater negative coefficient indicating it began to rely less on the experts earlier than GROUP2. Examining both the hot hand and gambler's fallacies, we find GROUP3 is the least influenced. Contrary to previous literature, we have found evidence that deciding individually can mitigate the effect of the hot hand fallacy rather than working in pairs since results suggest that pairs are actually the most prone to the hot hand fallacy. On the other hand, our findings on the gambler's fallacy support existing literature that INDIV is the most prone to the gambler's fallacy.

Some investment decisions made by individuals can also be made by groups. From the findings of our study, we recommend group investment decisions because of their more optimal investment strategy and less proneness to the hot hand and gambler's fallacies. After collectively taking into consideration all criteria, we find that GROUP3 is the most efficient treatment. It is more superior to GROUP2 because it is the least prone to the fallacies and have better investment strategies whereas GROUP2 is actually the most prone to the hot hand fallacy. Comparing INDIV and GROUP3, where INDIV represents an individual expert and GROUP3 representing an investment trio, we find evidence that investment trios make better investment strategies than experts as seen by our results. The results are in agreement with O'Leary (2011) and the Organization Theory where group performances, notably triads, mitigate the effects of biases, but the impact is not completely removed. For this study, we identify GROUP3 as the minimum group size where the mitigation impact is first felt.

As a policy implication, we recognize the benefits of forming investment clubs. Investment clubs are formed by non-professional investors who pool their money into a common fund and make investment decisions together. After establishing the minimum group size that mitigates the impact of the fallacies, we believe that the ideal investment club size is a triad. After factoring in the risk appetites of the investors, the triad can be further decomposed into specific genders. An all-male triad employs a more profit-maximizing strategy and higher risk appetite. However, for those who prefer a more risk-averse approach despite being less optimal, the composition of the triad should be female-male-male.

Investment clubs are applicable in settings where people would want to earn money through speculative markets such as the stock and bond markets when there are new information and opportunities to invest. We believe that the goals of the investment clubs are in line with the profit maximizing strategy supported by the rational choice theory since gathering a crowd to decide on investment strategies are found to be the smarter decision (Goldstein et al., 2014). This further heightens the implications of the Wisdom of Crowds where pooling non-professional investors together create superior decisions versus relying on the wisdom of the experts on investing.

Informal theory on financial economics provides that one of the investment strategies in maximizing returns is to buy low and sell high to reap capital gains (Raymundo, 2018). However, individual investors behave oppositely as they prefer to buy winning stocks and sell losing stocks (Johnson, Tellis, & Macinnis, 2005) since they are more prone to the hot hand fallacy, believing that the respective increase and decrease in prices will continue in the future periods. Since our results suggest that triads are the least affected by these behavioral biases, therefore they are more inclined in buying at low prices and selling at high prices, thus subsequently aligning their decisions with the objective of reaping the maximum capital gains available in accordance with their reaction on all available information on future prices.

References

[1] Baddeley, A., & Hitch, G. (1993). The recency effect: Implicit learning with explicit retrieval?. *Memory and Cognition.* 21(2), 146-155
[2] Bornstein, G., Kugler, T., & Ziegelmeyer, A. (2004). Individual and group decisions in the centipede game: Are groups more "rational" players?. *Journal of Experimental Social Psychology*, 40(5), 599-605. doi:http://dx.doi.org/10.1016/j.jesp.2003.11.003
[3] Fama, Eugene (1991), "Efficient Capital Markets II," *Journal of Finance*, 46 (December), 1575–1617.

[4] Feri, F., Irlenbusch, B., Sutter, M., (2010). Efficiency gains from team-based coordination – large-scale experimental evidence. *American Economics. Review*. 100 (4),1892–1912.
[5] Fischbein, H. (1975). The intuitive sources of probabilistic thinking in children (Vol. 85). *Springer Science & Business Media*.
[6] Goldstein, D. G., McAfee, R. P., & Suri, S. (2014, June). The wisdom of smaller, smarter crowds. In *Proceedings of the fifteenth ACM conference on Economics and computation* (pp. 471-488). ACM.
[7] Johnson, J., Tellis, G. J., & Macinnis, D. J. (2005). Losers, Winners, and Biased Trades. *Journal of Consumer Research*, *32*(2), 324-329. doi:10.1086/432241.
[8] Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the econometric society*, 263-291.
[9] Odean, T. (1998) .Are investors reluctant to realize their losses. *J. Finance*, 53 (3)), pp. 1175–1789
[10] O'Leary, D. E. (2011). The emergence of individual knowledge in a group setting: mitigating cognitive fallacies. *Group Decision and Negotiation*, *20*(1), 3-18.
[11] Rabin, M. (2002). Inference by Believers in the Law of Small Numbers. T*he Quarterly Journal of Economics*, 117(3), 775-816. doi:10.1162/003355302760193896
[12] Rabin, M., & Vayanos, D. (2009). The Gambler's and Hot Hand Fallacies: Theory and Applications. *V Review of Economic Studies, forthcoming*.
[13] Raymundo, R.B. (2018). *Lecture on Money and Banking.* De La Salle University, Manila City.
[14] Roney, C. J., & Trick, L. M. (2009). Sympathetic magic and perceptions of randomness: The hot hand versus the gambler's fallacy. *Thinking & Reasoning*, *15*(2), 197-210.
[15] Rufino, C. C. (2017). *Lecture on Advanced Financial Econometrics*. De La Salle University, Manila City.
[16] Rufino, C., (2013), Random walks in the different sub-markets of the Philippine Stock Exchange amid modernization. *Philippine Review of Economics* **XLX (1)**, pp. 57-82.
[17] Simmel, G. (1950). *The sociology of georg simmel* (Vol. 92892). Simon and Schuster.
[18] Sirri, E.R., Tufano, P., 1998. Costly search and mutual fund flows. *J. Finance.* 53 (5), 1589–1622.
[19] Stöckl, T., Huber, J., Kirchler, M., & Lindner, F. (2015). Hot hand and gambler's fallacy in teams: Evidence from investment experiments. *Journal of Economic Behavior & Organization*, 117, 327-339. doi:http://dx.doi.org/10.1016/j.jebo.2015.07.004
[20] Surowiecki, J. (2004). The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business. *Economies, Societies and Nations*, *296*.
[21] Weick, K. E. (1969). *The social psychology of organizing* (No. 04; HM131, W3.).

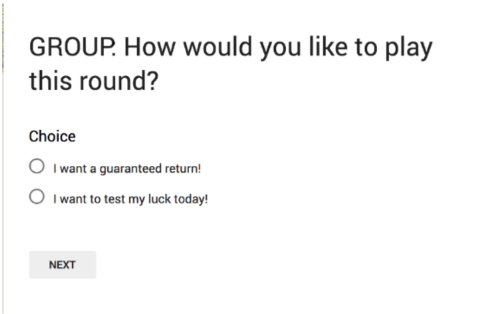
APPENDIX

Appendix A: Coin Realizations and Expert Performance

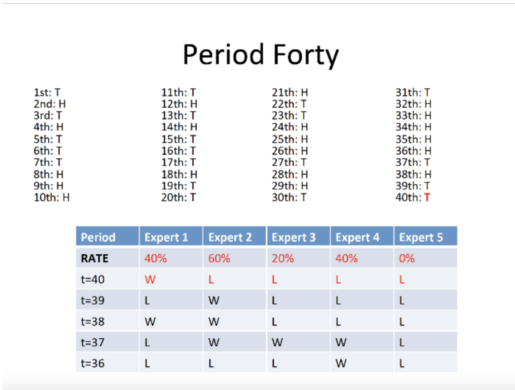
Period	Coin	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5
1	Tail	L	W	W	L	L
2	Heady	W	W	W	L	W
3	Tail	W	L	W	L	L
4	Head	W	L	L	L	W
5	Tail	L	L	L	L	L
6	Tail	L	W	L	W	W
7	Tail	W	W	W	W	L
8	Head	W	L	L	L	L
9	Head	W	L	L	L	L
10	Head	L	W	W	W	L
11	Tail	L	L	W	L	L
12	Head	L	W	W	W	W
13	Tail	L	L	W	L	L
14	Head	W	L	W	L	L
15	Tail	W	W	L	L	L
16	Tail	W	W	W	L	W
17	Tail	L	L	W	L	L
18	Head	L	L	W	W	L
19	Tail	W	L	L	W	W
20	Tail	L	W	L	W	W
21	Head	L	W	W	L	W
22	Tail	W	L	W	L	L
23	Tail	W	W	L	L	L
24	Head	W	L	L	W	L
25	Head	L	L	W	W	W
26	Head	L	W	L	L	W
27	Tail	L	W	L	W	W
28	Head	W	W	W	L	L
29	Head	L	L	L	L	L
30	Tail	W	W	W	W	L
31	Tail	L	L	W	W	W
32	Head	W	L	L	L	L
33	Head	L	W	W	W	W
34	Head	L	W	L	W	W
35	Head	L	W	W	L	W
36	Head	L	L	L	W	L
37	Tail	L	W	W	W	L
38	Head	W	W	L	L	L
39	Tail	L	W	L	L	L
40	Tail	W	L	L	L	L

Source: Stockl et al. (2015)

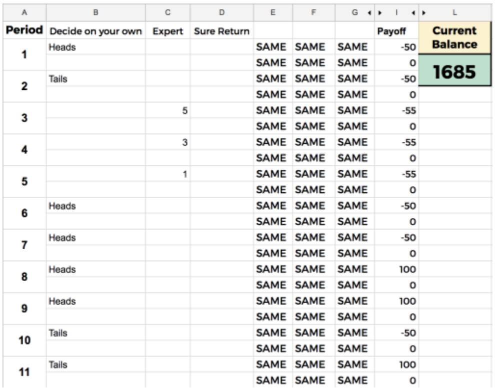
Appendix B: Screens used during the experiment



Decision Screen



Results Screen



Trading Screen

Appendix C: Talers earned

Tag Number	Taler Earned	Tag Number	Taler Earned	Tag Number	Taler Earned
I-000	1,570.00	G-000.a	545.00	K-000.a	-
I-001	1,350.00	G-001.a	1,685.00	K-001.a	1,350.00
I-002	550.00	G-002.a	750.00	K-002.a	2,290.00
I-003	1,640.00	G-003.a	1,015.00	K-003.a	2,090.00
I-004	1,020.00	G-004.a	1,800.00	K-004.a	1,550.00
I-005	2,170.00	G-005.a	1,490.00	K-005.a	1,350.00
I-006	930.00	G-006.a	555.00	K-006.a	925.00
I-007	895.00	G-007.a	1,500.00	K-007.a	2,440.00
I-008	1,500.00	G-008.a	900.00	K-008.a	1,000.00

I-009	665.00	G-009.a	1,150.00	K-009.a	1,700.00
I-010	760.00	G-010.a	1,405.00	K-010.a	1,485.00
I-011	1,345.00	G-011.a	2,195.00	K-011.a	1,855.00
I-012	995.00	G-012.a	1,085.00	K-012.a	1,465.00
I-013	1,995.00	G-0.13.a	1,870.00	K-013.a	715.00
I-014	1,115.00	G-0.14.a	1,660.00	K-014.a	2,250.00
I-015	1,760.00	G-015.a	1,090.00	K-015.a	1,500.00
I-016	1,200.00	G-016.a	1,950.00	K-016.a	800.00
I-017	1,160.00	G-017.a	2,095.00	K-017.a	1,730.00
I-018	1,420.00	G-018.a	1,800.00	K-018.a	2,005.00
I-019	1,550.00	G-019.a	2,085.00	K-019.a	1,745.00
I-020	1,350.00	G-020.a	645.00	K-020.a	1,645.00
I-021	520.00	G-021.a	-	K-021.a	375.00
I-022	585.00	G-022.a	1,035.00	K-022.a	1,330.00
I-023	1,315.00	G-023.a	1,160.00	K-023.a	1,205.00
I-024	1,135.00	G-024.a	1,450.00	K-024.a	1,445.00
I-025	1,520.00	G-025.a	280.00	K-025.a	2,005.00
I-026	-	G-026.a	1,375.00	K-026.a	1,455.00
I-027	1,150.00	G-027.a	325.00	K-027.a	1,640.00
I-028	850.00	G-028.a	950.00	K-028.a	1,280.00
I-029	2,095.00	G-029.a	1,920.00	K-029.a	1,350.00
I-030	1,655.00	G-030.a	1,950.00	K-030.a	1,235.00

Appendix D: Logit Regression Results

Logit Regression Model 1												
	RISKown				RISKexpert				RISKfree			
Variable	Logit Coefficient		Marginal Effect		Logit Coefficient		Marginal Effect		Logit Coefficient		Marginal Effect	
Constant	0.3986	***	0.7517	***	-0.6860	***	0.1393	***	-2.3409	***	0.1056	***
Group of 2	0.5733	***	0.1022	***	-1.2720	***	-0.1455	***	0.3692	*	0.0342	*
Group of 3	1.7248	***	0.3075	***	-2.2164	***	-0.2535	***	-0.6565	**	-0.0608	**
Period	0.0036		0.0006		-0.0150	***	-0.0017	***	0.0122	**	0.0011	**
F	0.2227	*	0.0397	*	-0.6441	***	-0.0737	***	0.4901	***	0.0454	***
FF	0.7219	***	0.1287	***	-0.2001		-0.0229		-1.1068	***	-0.1024	***
FM	0.5712	***	0.1018	***	-0.2145		-0.0245		-0.6706	***	-0.0621	***
FFM	-1.2813	***	-0.2284	***	1.8813	***	0.2151	***	0.5675		0.0210	
FMM	-0.8532	***	-0.1521	***	1.0511	***	0.1202	***	0.2269	**	0.0525	**
FFF	-0.8111	***	-0.1446	***	1.3845	***	0.1583	***	-0.0126		-0.0012	
*** - Significant at 99% confidence level												
** - Significant at 95% confidence level												
* - Significant at 90% confidence level												

Logit Regression Model 2												
	RISKown				RISKexpert				RISKfree			
Variable	Logit Coefficient		Marginal Effect		Logit Coefficient		Marginal Effect		Logit Coefficient		Marginal Effect	
Constant	0.5556	***	0.7517	***	-0.9738	***	0.1394	***	-2.2527	***	0.1056	***
Group of 2	0.0969		0.0172		-0.6505	**	-0.0743	**	0.5894	*	0.0551	*
Group of 3	1.6545	***	0.2943	***	-1.7473	***	-0.1995	***	-1.3307	***	-0.1132	***
Period	-0.0040		-0.0007		-0.0007		-0.0001		0.0081		0.0008	
F	0.2227	*	0.0396	*	-0.6409	***	-0.0732	***	0.4895	***	0.0469	***
FF	0.7276	***	0.1294	***	-0.2020		-0.0231		-1.1050	***	-0.0992	***
FM	0.5759	***	0.1025	***	-0.2165		-0.0247		-0.6693	***	-0.0619	***
FFM	-1.2809	***	-0.2279	***	1.8904	***	-0.2159	***	0.2292		0.0201	
FMM	-0.8530	***	-0.1518	***	1.0545	***	0.1204	***	0.5742	**	0.0537	**
FFF	-0.8109	***	-0.1443	***	1.3899	***	0.1587	***	-0.0127		0.0021	
Group2Period	0.0238	***	0.0042	***	-0.0330	***	-0.0038	***	-0.0106		-0.0010	
Group3Period	0.0034		0.0006		-0.0244	**	-0.0028	**	0.0292	**	0.0024	**
*** - Significant at 99% confidence level												
** - Significant at 95% confidence level												
* - Significant at 90% confidence level												

Multinomial Logit Regression Model 1												
	Risk Own				Risk expert				Risk free			
Variable	Logit Coefficient		Marginal Effect		Logit Coefficient		Marginal Effect		Logit Coefficient		Marginal Effect	
Constant	0.5421	***	0.7540	***	Base Outcome		0.1402	***	-1.3863	***	0.1059	***
Group of 2	1.2449	***	0.1129	***	Base Outcome		-0.1451	***	1.3855	***	0.0322	*
Group of 3	2.3233	***	0.3065	***	Base Outcome		-0.2517	***	1.3470	***	0.0548	**
Period	0.0136	***	0.0006		Base Outcome		-0.0017	***	0.0234	***	0.0011	**
F	0.5934	***	0.0323		Base Outcome		-0.0734	***	0.9379	***	0.0411	**
FF	0.3237		0.1206	***	Base Outcome		-0.0196		-0.8184	**	-0.1010	***
FM	0.3137		0.0864	**	Base Outcome		-0.0250		-0.3993		-0.0613	***

FFM	-1.9147	***	-0.2244	***	Base Outcome	0.2131	***	-1.4751	***	0.0113	
FMM	-1.1060	***	-0.1600	***	Base Outcome	0.1170	***	-0.4588		0.0430	
FFF	-1.3910	***	-0.1463	***	Base Outcome	0.1581	***	-1.2869	**	-0.0118	
*** - Significant at 99% confidence level											
** - Significant at 95% confidence level											
* - Significant at 90% confidence level											

Multinomial Logit Regression Model 2											
	Risk own				Risk expert			Risk free			
Variable	Logit Coefficient		Marginal Effect		Logit Coefficient	Marginal Effect		Logit Coefficient		Marginal Effect	
Constant	0.8359	***	0.7540	***	Base Outcome	0.1402	***	-1.0552	***	0.1059	***
Group of 2	0.5675	**	0.0225		Base Outcome	-0.0718	**	1.0052	***	0.0493	*
Group of 3	1.8792	***	0.3089	***	Base Outcome	-0.1910	***	0.2939		-0.1179	***
Period	-0.0011		-0.0008		Base Outcome	0.0000		0.0072		0.0008	
F	0.5897	***	0.0321		Base Outcome	-0.0729	***	0.9318	***	0.0408	**
FF	0.3288		0.1208	***	Base Outcome	-0.0201		-0.8128	**	-0.1007	***
FM	0.3172		0.0865	**	Base Outcome	-0.0254		-0.3952		-0.0611	***
FFM	-1.9204	***	-0.2236	***	Base Outcome	0.2136	***	-1.4921	***	0.0101	
FMM	-1.1053	***	-0.1597	***	Base Outcome	0.1167	***	-0.4573		0.0430	
FFF	-1.3951	***	-0.1459	***	Base Outcome	0.1584	***	-1.2989	**	-0.0126	
Group2Period	0.0357	***	0.0047	***	Base Outcome	-0.0039	***	0.0213		-0.0008	
Group3Period	0.0232	**	0.0003		Base Outcome	-0.0031	**	0.0493	***	0.0028	**
*** - Significant at 99% confidence level											
** - Significant at 95% confidence level											
* - Significant at 90% confidence level											

```
. ranksum  own, by( MM)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

          MM |      obs    rank sum    expected
-------------+---------------------------------
           0 |     8599    37904848    38695500
           1 |      400     2590652     1800000
-------------+---------------------------------
    combined |     8999    40495500    40495500

unadjusted variance    2.580e+09
adjustment for ties   -9.523e+08
                      ----------
adjusted variance      1.627e+09

Ho: own(MM==0) = own(MM==1)
             z = -19.599
    Prob > |z| =   0.0000
```

```
. ranksum  own, by( M)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

           M |      obs    rank sum    expected
-------------+---------------------------------
           0 |     8439    37180853    37975500
           1 |      560   3314647.5     2520000
-------------+---------------------------------
    combined |     8999    40495500    40495500

unadjusted variance    3.544e+09
adjustment for ties   -1.308e+09
                      ----------
adjusted variance      2.236e+09

Ho: own(M==0) = own(M==1)
             z = -16.805
    Prob > |z| =   0.0000
```

```
. ranksum  own, by( F)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

            F |      obs    rank sum    expected
 -------------+---------------------------------
            0 |     8359    36560134    37615500
            1 |      640   3935366.5     2880000
 -------------+---------------------------------
     combined |     8999    40495500    40495500

unadjusted variance   4.012e+09
adjustment for ties  -1.481e+09
                     ----------
adjusted variance     2.531e+09

Ho: own(F==0) = own(F==1)
             z = -20.977
    Prob > |z| =   0.0000
```

```
. ranksum  own, by( FM)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

           FM |      obs    rank sum    expected
 -------------+---------------------------------
            0 |     8679    38287094    39055500
            1 |      320   2208406.5     1440000
 -------------+---------------------------------
     combined |     8999    40495500    40495500

unadjusted variance   2.083e+09
adjustment for ties  -7.689e+08
                     ----------
adjusted variance     1.314e+09

Ho: own(FM==0) = own(FM==1)
             z = -21.198
    Prob > |z| =   0.0000
```

Appendix E: Mann Whitney U-test for RISKOwn

```
. ranksum  own, by(FF)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

           FF |      obs    rank sum    expected
 -------------+---------------------------------
            0 |     8519    37140145    38335500
            1 |      480     3355355     2160000
 -------------+---------------------------------
     combined |     8999    40495500    40495500

unadjusted variance   3.067e+09
adjustment for ties  -1.132e+09
                     ----------
adjusted variance     1.935e+09

Ho: own(FF==0) = own(FF==1)
             z = -27.176
    Prob > |z| =   0.0000
```

```
. ranksum  own, by(FFF)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

          FFF |      obs    rank sum    expected
 -------------+---------------------------------
            0 |     8799    39146180    39595500
            1 |      200     1349320      900000
 -------------+---------------------------------
     combined |     8999    40495500    40495500

unadjusted variance   1.320e+09
adjustment for ties  -4.872e+08
                     ----------
adjusted variance     8.326e+08

Ho: own(FFF==0) = own(FFF==1)
             z = -15.572
    Prob > |z| =   0.0000
```

```
. ranksum  own, by(FFM)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

          FFM |      obs    rank sum    expected
 -------------+---------------------------------
            0 |     8719    38714440    39235500
            1 |      280     1781060     1260000
 -------------+---------------------------------
     combined |     8999    40495500    40495500

unadjusted variance   1.831e+09
adjustment for ties  -6.759e+08
                     ----------
adjusted variance     1.155e+09

Ho: own(FFM==0) = own(FFM==1)
             z = -15.332
    Prob > |z| =   0.0000
```

```
. ranksum  own, by(FMM)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

          FMM |      obs    rank sum    expected
 -------------+---------------------------------
            0 |     8559    37540495    38515500
            1 |      440   2955005.5     1980000
 -------------+---------------------------------
     combined |     8999    40495500    40495500

unadjusted variance   2.824e+09
adjustment for ties  -1.043e+09
                     ----------
adjusted variance     1.782e+09

Ho: own(FMM==0) = own(FMM==1)
             z = -23.098
    Prob > |z| =   0.0000
```

```
. ranksum  own, by( M)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

            M |      obs    rank sum    expected
 -------------+---------------------------------
            0 |     8439    37180853    37975500
            1 |      560   3314647.5     2520000
 -------------+---------------------------------
     combined |     8999    40495500    40495500

unadjusted variance   3.544e+09
adjustment for ties  -1.308e+09
                     ----------
adjusted variance     2.236e+09

Ho: own(M==0) = own(M==1)
             z = -16.805
    Prob > |z| =   0.0000
```

```
. ranksum  own, by(MMM)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

          MMM |      obs    rank sum    expected
 -------------+---------------------------------
            0 |     8719    38480466    39235500
            1 |      280     2015034     1260000
 -------------+---------------------------------
     combined |     8999    40495500    40495500

unadjusted variance   1.831e+09
adjustment for ties  -6.759e+08
                     ----------
adjusted variance     1.155e+09

Ho: own(MMM==0) = own(MMM==1)
             z = -22.216
    Prob > |z| =   0.0000
```

```
. ranksum  expert, by( M)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

           M |      obs    rank sum    expected
-------------+---------------------------------
           0 |     8439    37432136    37975500
           1 |      560     3063364     2520000
-------------+---------------------------------
    combined |     8999    40495500    40495500

unadjusted variance   3.544e+09
adjustment for ties  -2.984e+09
                     ----------
adjusted variance     5.601e+08

Ho: expert(M==0) = expert(M==1)
             z = -22.960
    Prob > |z| =   0.0000
```

```
. ranksum  expert, by( MM)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

          MM |      obs    rank sum    expected
-------------+---------------------------------
           0 |     8599    38624919    38695500
           1 |      400     1870581     1800000
-------------+---------------------------------
    combined |     8999    40495500    40495500

unadjusted variance   2.580e+09
adjustment for ties  -2.172e+09
                     ----------
adjusted variance     4.076e+08

Ho: expert(MM==0) = expert(MM==1)
             z =  -3.496
    Prob > |z| =   0.0005
```

```
. ranksum  expert, by( FF)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

          FF |      obs    rank sum    expected
-------------+---------------------------------
           0 |     8519    38284999    38335500
           1 |      480     2210501     2160000
-------------+---------------------------------
    combined |     8999    40495500    40495500

unadjusted variance   3.067e+09
adjustment for ties  -2.582e+09
                     ----------
adjusted variance     4.846e+08

Ho: expert(FF==0) = expert(FF==1)
             z =  -2.294
    Prob > |z| =   0.0218
```

```
. ranksum  expert, by( MMM)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

         MMM |      obs    rank sum    expected
-------------+---------------------------------
           0 |     8719    39256286    39235500
           1 |      280   1239214.5     1260000
-------------+---------------------------------
    combined |     8999    40495500    40495500

unadjusted variance   1.831e+09
adjustment for ties  -1.542e+09
                     ----------
adjusted variance     2.893e+08

Ho: expert(MMM==0) = expert(MMM==1)
             z =   1.222
    Prob > |z| =   0.2217
```

```
. ranksum  own, by( F)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

           F |      obs    rank sum    expected
-------------+---------------------------------
           0 |     8359    36560134    37615500
           1 |      640   3935366.5     2880000
-------------+---------------------------------
    combined |     8999    40495500    40495500

unadjusted variance   4.012e+09
adjustment for ties  -1.481e+09
                     ----------
adjusted variance     2.531e+09

Ho: own(F==0) = own(F==1)
             z = -20.977
    Prob > |z| =   0.0000
```

```
. ranksum  expert, by( FM)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

          FM |      obs    rank sum    expected
-------------+---------------------------------
           0 |     8679    39023333    39055500
           1 |      320   1472167.5     1440000
-------------+---------------------------------
    combined |     8999    40495500    40495500

unadjusted variance   2.083e+09
adjustment for ties  -1.754e+09
                     ----------
adjusted variance     3.291e+08

Ho: expert(FM==0) = expert(FM==1)
             z =  -1.773
    Prob > |z| =   0.0762
```

```
. ranksum  expert, by( F)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

           F |      obs    rank sum    expected
-------------+---------------------------------
           0 |     8359    37303693    37615500
           1 |      640   3191807.5     2880000
-------------+---------------------------------
    combined |     8999    40495500    40495500

unadjusted variance   4.012e+09
adjustment for ties  -3.378e+09
                     ----------
adjusted variance     6.340e+08

Ho: expert(F==0) = expert(F==1)
             z = -12.383
    Prob > |z| =   0.0000
```

```
. ranksum  expert, by( FFM)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

         FFM |      obs    rank sum    expected
-------------+---------------------------------
           0 |     8719    39040310    39235500
           1 |      280   1455190.5     1260000
-------------+---------------------------------
    combined |     8999    40495500    40495500

unadjusted variance   1.831e+09
adjustment for ties  -1.542e+09
                     ----------
adjusted variance     2.893e+08

Ho: expert(FFM==0) = expert(FFM==1)
             z = -11.475
    Prob > |z| =   0.0000
```

Appendix F: Mann Whitney U-test for RISKExpert

```
. ranksum  expert, by( FMM)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

         FMM |      obs    rank sum    expected
-------------+---------------------------------
           0 |     8559    38418963    38515500
           1 |      440     2076537     1980000
-------------+---------------------------------
    combined |     8999    40495500    40495500

unadjusted variance    2.824e+09
adjustment for ties   -2.378e+09
                      ----------
adjusted variance      4.463e+08

Ho: expert(FMM==0) = expert(FMM==1)
             z =  -4.570
    Prob > |z| =   0.0000
```

```
. ranksum  free, by( M)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

           M |      obs    rank sum    expected
-------------+---------------------------------
           0 |     8439    37802931    37975500
           1 |      560     2692569     2520000
-------------+---------------------------------
    combined |     8999    40495500    40495500

unadjusted variance    3.544e+09
adjustment for ties   -3.114e+09
                      ----------
adjusted variance      4.300e+08

Ho: free(M==0) = free(M==1)
             z =  -8.322
    Prob > |z| =   0.0000
```

```
. ranksum  expert, by(FFF)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

         FFF |      obs    rank sum    expected
-------------+---------------------------------
           0 |     8799    39519714    39595500
           1 |      200      975786      900000
-------------+---------------------------------
    combined |     8999    40495500    40495500

unadjusted variance    1.320e+09
adjustment for ties   -1.111e+09
                      ----------
adjusted variance      2.086e+08

Ho: expert(FFF==0) = expert(FFF==1)
             z =  -5.248
    Prob > |z| =   0.0000
```

```
. ranksum  free, by(FM)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

          FM |      obs    rank sum    expected
-------------+---------------------------------
           0 |     8679    38994814    39055500
           1 |      320   1500686.5     1440000
-------------+---------------------------------
    combined |     8999    40495500    40495500

unadjusted variance    2.083e+09
adjustment for ties   -1.830e+09
                      ----------
adjusted variance      2.527e+08

Ho: free(FM==0) = free(FM==1)
             z =  -3.817
    Prob > |z| =   0.0001
```

```
. ranksum  free, by(F)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

           F |      obs    rank sum    expected
-------------+---------------------------------
           0 |     8359    37251154    37615500
           1 |      640     3244346     2880000
-------------+---------------------------------
    combined |     8999    40495500    40495500

unadjusted variance    4.012e+09
adjustment for ties   -3.525e+09
                      ----------
adjusted variance      4.868e+08

Ho: free(F==0) = free(F==1)
             z = -16.513
    Prob > |z| =   0.0000
```

```
. ranksum  free, by(FF)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

          FF |      obs    rank sum    expected
-------------+---------------------------------
           0 |     8519    38305214    38335500
           1 |      480   2190286.5     2160000
-------------+---------------------------------
    combined |     8999    40495500    40495500

unadjusted variance    3.067e+09
adjustment for ties   -2.695e+09
                      ----------
adjusted variance      3.721e+08

Ho: free(FF==0) = free(FF==1)
             z =  -1.570
    Prob > |z| =   0.1164
```

Appendix G: Mann Whitney U-test for RISKfree

```
. ranksum  free, by( MM)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

           MM |      obs    rank sum    expected
-------------+---------------------------------
            0 |     8599    38497031    38695500
            1 |      400   1998469.5     1800000
-------------+---------------------------------
     combined |     8999    40495500    40495500

unadjusted variance   2.580e+09
adjustment for ties  -2.267e+09
                     ----------
adjusted variance     3.130e+08

Ho: free(MM==0) = free(MM==1)
             z = -11.218
    Prob > |z| =   0.0000
```

```
. ranksum  free, by(FFF)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

          FFF |      obs    rank sum    expected
-------------+---------------------------------
            0 |     8799    39579506    39595500
            1 |      200      915994      900000
-------------+---------------------------------
     combined |     8999    40495500    40495500

unadjusted variance   1.320e+09
adjustment for ties  -1.160e+09
                     ----------
adjusted variance     1.601e+08

Ho: free(FFF==0) = free(FFF==1)
             z =  -1.264
    Prob > |z| =   0.2063
```

```
. ranksum  free, by( FFM)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

          FFM |      obs    rank sum    expected
-------------+---------------------------------
            0 |     8719    39194211    39235500
            1 |      280   1301289.5     1260000
-------------+---------------------------------
     combined |     8999    40495500    40495500

unadjusted variance   1.831e+09
adjustment for ties  -1.609e+09
                     ----------
adjusted variance     2.222e+08

Ho: free(FFM==0) = free(FFM==1)
             z =  -2.770
    Prob > |z| =   0.0056
```

```
. ranksum  free, by( FMM)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

          FMM |      obs    rank sum    expected
-------------+---------------------------------
            0 |     8559    38396623    38515500
            1 |      440   2098877.5     1980000
-------------+---------------------------------
     combined |     8999    40495500    40495500

unadjusted variance   2.824e+09
adjustment for ties  -2.482e+09
                     ----------
adjusted variance     3.427e+08

Ho: free(FMM==0) = free(FMM==1)
             z =  -6.422
    Prob > |z| =   0.0000
```

```
. ranksum  free, by( MMM)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

          MMM |      obs    rank sum    expected
-------------+---------------------------------
            0 |     8719    39212209    39235500
            1 |      280   1283291.5     1260000
-------------+---------------------------------
     combined |     8999    40495500    40495500

unadjusted variance   1.831e+09
adjustment for ties  -1.609e+09
                     ----------
adjusted variance     2.222e+08

Ho: free(MMM==0) = free(MMM==1)
             z =  -1.563
    Prob > |z| =   0.1181
```