Belief Formation Under Signal Correlation*

Tanjim Hossain
University of Toronto

Ryo Okui
New York University - Shanghai
University of Gothenburg

July 4, 2018

Abstract

Using a set of incentivized laboratory experiments, we characterize how people form beliefs about a random variable based on independent and correlated signals. Subjects form their beliefs in an unbiased, yet suboptimal, way. They do not completely neglect precision level or correlation structure of signals. While subjects do *overvalue* moderately or strongly correlated signals as previously documented in the literature, we find that they actually *undervalue* weakly correlated signals. Additionally, they do not fully benefit from wisdom of the crowd — they undervalue aggregated information about others' actions in favor of their private information.

Keywords: Correlated and independent signals, information processing, bounded rationality, correlation neglect and overadjustment, belief elicitation, wisdom of the crowd.

JEL classification: C91, D81, D83, D84.

*We thank Andro Rilović, Clarice Zhao, Jiaying (Joey) Yang, and Yue Yu for excellent research assistance. Masaki Aoyagi, Guodong Chen, Zhenxing Huang, Gilat Levy, Takeshi Murooka, Imran Rasul, and Eric Set provided us with useful comments. We thank participants at a number of seminars and conference presentations for helpful comments and suggestions. We are grateful to Sudipto Dasgupta who initially got us interested in the problem. A part of this research was conducted while Okui was at Vrije Universiteit Amsterdam and Kyoto University. We appreciate the Center for Research in Experimental Economics and Political Decision Making (CREED) at the University of Amsterdam for letting us to use their laboratory. All remaining errors are of course ours. Okui acknowledges financial support from the Japan Society for the Promotion of Science (JSPS) under KAKENHI Grant Nos. 25780151, 25285067, 15H03329 and 16K03598. Please direct all correspondence to Tanjim Hossain at tanjim.hossain@utoronto.ca.

1 Introduction

"In retrospect, a key mistake in the forecast updating that Kremp and I did, was that we ignored the correlation in the partial information from early-voting tallies," admitted statistician Andrew Gelman on his blog regarding his 2016 US presidential election forecasts.[1] Even an academic statistician like Gelman may not take correlation among information sources perfectly into account when aggregating them. Processing information about the realization of a random variable from correlated and independent sources is a common, yet complex, problem. Examples of correlated signals are abound—A set of stock analysts may drift towards commonly known prevailing consensus and neglect their own idiosyncratic information in generating their forecasts, making the forecasts highly correlated (Trueman, 1994, Huang et al., 2017). Multiple newspapers relying on the same original news source such as The Associated Press for writing their own articles would publish news stories from a similar angle.[2] Social networks may give rise to signal correlation (Eyster and Rabin, 2004). Information about experience goods received from a set of friends will be correlated where the degree of correlation may depend on the connectivity among friends (Lee and Iyengar, 2016). Business and leisure travellers typically use different sets of information sources (Chen, 2000), which may lead to correlation among reviews by business travellers and among reviews by leisure travellers. To convince a jury, attorneys present multiple witnesses who provide similar information (Harkins and Petty, 1987).

Using a comprehensive set of laboratory experiments, we investigate how people value signals of varying levels of precision and (positive) correlation and process the information she receives to form her beliefs. In theory, if the signal generation process is known, a homo economicus will just calculate the posterior distribution of the random variable based on her signals in a Bayesian manner. In practice, however, how people perform such a technically challenging task remains an empirical question. We find that subjects do not completely neglect signal correlation, but adjust their beliefs for correlation suboptimally. Specifically, they *overvalue* moderately or strongly correlated signals, consistent with the findings from the received literature.[3] Our new finding is that subjects *undervalue* signals that are weakly correlated.

Overview of the Experimental Design and Findings: In our experiments, subjects receive forecasts by 10 different analysts regarding the *earning per share* (EPS) for a stock and are asked to estimate the true value of the EPS. Each experimental session consists of 60 periods where the stock in each period is different and independent of each other. For a given stock, all the forecasts are unbiased—they equal the sum of the true value and two zero-mean normally-distributed error terms. Using a novel information generation

[1]November 8, 2016, http://andrewgelman.com/2016/11/08/election-forecasting-updating-error-ignored-correlations-data-thus-producing-illusory-precision-inferences/

[2]For example, to report about the US military operation WebOps, Daily Mail and US News & World Report directly published the original AP article: https://apnews.com/b3fd7213bb0e41b3b02eb15265e9d292/us-military-botches-online-fight-against-islamic-state. On the other hand, Fox News and the Independent wrote their own stories based on the AP article. In Japan, media frequently get information from the government through the "kisha" club and then they write up their own articles as reported in Au and Kawai (2012).

[3]See, for example, Enke and Zimmermann (2017).

process, we create two types of analysts—independent and correlated. For independent analysts, all error terms are independently drawn, which makes their forecasts, conditional on the EPS, independent of each other. For a set of correlated analysts, on the other hand, one error term is common for all these analysts while the other is independent across the analysts. Thus, signals from this set are correlated with each other, even conditional on the true value of the EPS.

We vary the variances of the distributions of the error terms across sessions—we use five different sets of parameters for the error term distributions in our experiments. By doing so, we can vary the precision level of a forecast and also the correlation level of a set of correlated forecasts. In each case, analysts are from two different groups where the signal generation process for all analysts from a particular group is the same. Under three sets of parameters, one group provide independent forecasts while the other group provide correlated forecasts. Under the fourth set of parameters, both groups provide correlated forecasts and under the last, both groups provide independent forecasts. Within a session, we use the same distributions of the error terms in every period. However, the numbers of analysts from each group vary across periods.

In any given period, after providing 10 forecasts about the EPS of that period's stock, we ask subjects to predict the true value of the EPS. The subjects are fully informed of the data generating process of signals including variance and correlation structure. Let us refer to the reported belief as a subject's *initial prediction*. We incentivize truthful reporting of subjects' beliefs by using the binarized scoring rule by Hossain and Okui (2013). The optimal prediction for the true EPS based on the 10 forecasts is a weighted average of the average of the forecasts from one group and that from the other group. The optimal weights depend on the numbers of the two types of forecasts and the variances of and correlations across the error terms. Moreover, any weighted average (even with suboptimal weights) of the two averages is an unbiased prediction. From the initial predictions that the subjects report, we estimate the weights they put on the averages from two groups in calculating their prediction.

While subjects in our experiments choose unbiased predictions, they typically choose weights sub-optimally. The direction of suboptimality, however, varies based on the nature of the correlation and the relative precision of the independent and correlated signals. In a new finding to the literature, the subjects put sub-optimally low weight on the correlated forecasts when the correlation is weak. We refer to this as *correlation overadjustment*.[4] We also find that they under-adjust strongly or moderately correlated signals. Such *correlation neglect* under strong correlation is consistent with the findings in the received literature. On the other hand, subjects do not neglect correlation completely in any of the settings. When both groups of analysts provide independent forecasts with different levels of precision, subjects assign higher than optimal weights to the less precise forecasts; but, they do not completely ignore variance. Comparing this setting to those where one group of forecasts are correlated, we see that the chosen weights are directionally consistent

[4]This is different from overadjustment defined in epidemiology or public health literature where it represents using unnecessary variables in a regression model as defined by Breslow (1982).

to the predictions from Bayesian updating. These results suggest that imperfect adjustment for correlation (sometimes under-adjusting and sometimes over-adjusting) and imperfect calculation of optimal weights explain our results.

We investigate how subjects utilize aggregated information from other subjects using a second exercise. In each period, after subjects reported their initial prediction based on their private information as reported above, we report the average of initial predictions of all subjects in that session for that stock. We refer to this average as the *session average* for that period. Then the subjects are asked to submit a *revised prediction* based on the average of initial predictions as well as the 10 forecasts. In a given period, each subject receive a different set of 10 forecasts for the same stock and subjects are aware of it. Thus, the session average provides new and public information to all subjects. If a subject assumes that all other subjects' predictions are unbiased and have variances not higher than her own prediction's variance, it is optimal to ignore her own initial prediction and report the session average (which includes her own initial prediction) as the revised prediction. In other words, it would be optimal for them to follow the "wisdom of the crowd." On the other hand, the session average results from how other subjects incorporate their private information in forming their predictions. As a consequence, the precision of the session average can be ambiguous to subjects. If they are pessimistic or cautious, the session average may not be perceived to be as valuable as theory would predict.

We find that subjects do take their initial prediction into account when calculating their revised prediction.[5] Moreover, the weights they put on their own initial prediction are quite similar across session specifications. Under all specifications, subject put between 19%-26% weight on their initial prediction and around 74%-81% weight on the average initial prediction of all the subjects in the session (which includes their own initial prediction). From the revised predictions, we can estimate that subjects believe that their own prediction have lower variance by a factor of 25 to 50. This provides us with a measurement of how ambiguity aversion can reduce the value of information.

People frequently deal with sophisticated information structures in everyday life. A particularly interesting question is how they process correlated and independent information. We present results from a comprehensive set of experiments that systematically investigate people's belief formation rules under a number of different information structures by directly eliciting belief instead of indirectly inferring it from a specific market situation. While we use a specific setting involving financial assets, the underlying theoretical framework is versatile enough that the experiments provide insights on information processing at a general level that can be useful for many different applications. We show that people are not "hyper-rational," and their "boundedly rational" behavior follows a specific pattern. Moreover, we provide some insights into on

[5]Thus, our finding is consistent with that of Nöth and Weber (2003) and others who find that people underweight public information in favor of their private information. However, our interpretation (see Section 4.2) is related to ambiguity aversion, which is different from their overconfidence interpretation.

how people exactly treat correlation. In particular, overadjustment to weak correlation is a new finding. The lessons from this paper can be quite generally applied and is not only restricted to financial decision-making.

In the following subsection, we discuss how our work relates to the extant literature. Section 2 presents the theoretical framework. Section 3 explains the setting of our experiments. We examine the results of the experiments in Section 4. Section 5 concludes the paper. Proofs of the mathematical propositions are included in the Appendix.

Relation to the Literature: Correlation neglect has gained attention from researchers recently. We argue that correlation being strong is a hidden premise in the literature of correlation neglect. An opposite phenomenon, correlation overadjustment, can also be observed under weak correlation.

Perhaps the most closely related paper is Enke and Zimmerman (2017), who demonstrate correlation neglect when correlation between signals arises due to repetition of some information. They suggest that correlation neglect may be remedied by making subjects notice the presence of correlation. We create a richer experimental setting that allows us to explore the extent of correlation neglect more comprehensively. While we find some degree of correlation neglect under strong correlation, which is consistent with their findings, we also find that subjects undervalued weakly correlated signals (correlation overadjustment). Maines (1996) also documents correlation neglect in an experimental setting where, similar to ours, subjects evaluate analyst forecasts. Luan, Sorkin, and Itzkowitz (2004) study how people process information using a small scale experiment and find that participants overvalued signals with high accuracy and high correlation. In psychology, Soll (1999) explains correlation neglect or even a preference for redundant information using different intuitive theories of information.

In the context of financial markets, Eyster and Weizsäcker (2016) provide subjects with multiple assets where the returns from some of them are independent and the rest are correlated and investigate their portfolio choice. They find evidence that subjects treat correlated variables as uncorrelated and give equal weights to all assets. We provide more generalizable insights regarding information processing by directly analyzing belief formation rules. Kroll, Levy, and Rapoport (1988) and Kallir and Sonsino (2009) also investigate the role of correlation in asset allocation.

There have been several attempts to characterize choices with misperception of correlation in decision theoretic frameworks. Ellis and Piccione (2017) provide an axiomatic framework to represent asset choice under such settings. Levy and Razin (2018) consider situations in which correlation structure is ambiguous and show that correlation neglect is observed when each signal is precise but otherwise more conservative behaviors are induced. Their theory is not directly applicable to our experimental setting and Section 4.1.3 discusses potential relationships between our findings and their theoretical results. In applications of correlation neglect, Levy and Razin (2015) analyze the case when voters ignore correlation in signals about the state of the world. Ortoleva and Snowberg (2015) show that correlation neglect may give rise to overconfidence, which can be connected to ideological extremeness in politics.

Our paper is somewhat related to the "multiple source effect" studied in psychology (Harkins and Petty, 1981a, 1981b, 1987). It is a phenomenon that information presented by multiple sources may have a large impact. Such a setting may be considered as a special case of correlated signals. However, this literature tends to focus on the comparison between multiple sources and a single source (or simple repetition). We present both correlated signals and independent signals. This concept of multiple source effect has been applied to marketing, in particular advertising (see, e.g., Moore, Reardon, and Qualls, 1989, Tang, Newton, and Wang, 2007 and Lim et al, 2015), and is used in evaluating the effectiveness of digital marketing (e.g., Chang and Thorson, 2004, Lee and Nass, 2004, and Voorveld, 2011). Our results would naturally be useful in those applications.

2 Theoretical Framework

This section discusses the underlying theoretical framework for our experiments. First, we present a general framework that can describe all of the treatments used in our experiments. We then discuss the optimal behavior under Bayesian updating. We also consider how subjects incorporate ambiguity about other subjects' actions in updating their beliefs.

Suppose a rational person, who we will refer to as a subject, has an uninformative prior about the distribution of a random variable T. She receives N signals about the realization of T. Of these signals, n_A signals belong to group A and the remaining n_B signals belong to group B (i.e., $n_B = N - n_A$). The information generation process is as follows — signal $j \in \{1, 2, \ldots, n_l\}$ from group $l \in \{A, B\}$ is denoted by X_j^l where $X_j^l = T + \epsilon_{lIj} + \epsilon_{lC}$. Here, ϵ_{lIj} is an independently drawn error term that is different for each signal j within group l and ϵ_{lC} is an error term that is common to all the signals belonging to group l. Note that errors ϵ_{lIj} and ϵ_{lC} are drawn from normal distributions with mean 0 and variances of σ_{lI}^2 and σ_{lC}^2, respectively. Thus, all signals are unbiased. If $\sigma_{lC}^2 > 0$, then all signals belonging to group l, conditional on T, are correlated with each other. On the other hand, if $\sigma_{lC}^2 = 0$, then all signals belonging to group l, conditional on T, are independent of each other.[6] By varying the variances σ_{AI}^2, σ_{AC}^2, σ_{BI}^2, and σ_{BC}^2, we can vary the precision and correlation levels of the two groups of signals. We refer to these variances and the number of signals from the two groups as the *information generation process parameters*—$(\sigma_{AI}^2, \sigma_{AC}^2, \sigma_{BI}^2, \sigma_{BC}^2, n_A, n_B)$.

Setting 1 below summarizes our framework. For all theoretical derivations, we will assume that the signal structure follows this setting.

[6]There is a slight difference between the setting here and the instructions used in the experiments. When a signal consists of only independent components, the instructions state that the signal is still the sum of two error terms (we set the sum of their variances equal to the variance of ϵ_{lIj}) to make the description of signal generation across the two groups of signals comparable. However, there is no difference in the mathematical structure.

Setting 1. *For a given* $(\sigma_{AI}^2, \sigma_{AC}^2, \sigma_{BI}^2, \sigma_{BC}^2, n_A, n_B)$*, signals* X_j^l *for* $j = 1, \ldots, n_l$ *and* $l \in \{A, B\}$ *are generated by* $X_j^l = T + \epsilon_{lIj} + \epsilon_{lC}$*, where* $\epsilon_{lIj} \sim N(0, \sigma_{lI}^2)$ *and* $\epsilon_{lC} \sim N(0, \sigma_{lC}^2)$ *and* ϵ_{lIj} *and* ϵ_{lC} *are mutually independent.*

Based on receiving n_A group A and n_B group B signals about the realization of T, the most efficient unbiased estimator of the realization is a weighted average of $\bar{X}_A$ and $\bar{X}_B$ where $\bar{X}_l$ is the mean of the n_l signals from group $l \in \{A, B\}$. This estimator is presented below.

Proposition 1. *Suppose that the prior of* T *is* $T \sim N(\mu_p, \sigma_p^2)$*. The mean of the limit of the posterior distribution of* T *conditional on* X_j^l *for* $j = 1, \ldots, n_l$ *and* $l \in \{A, B\}$ *under* $\sigma_p^2 \to \infty$ *(i.e., when the prior tends to uninformative one) is*

$$w_A^0 \bar{X}_A + w_B^0 \bar{X}_B,$$

where

$$w_A^0 = \frac{n_A(n_B\sigma_{BC}^2 + \sigma_{BI}^2)}{n_A(n_B\sigma_{BC}^2 + \sigma_{BI}^2) + n_B(n_A\sigma_{AC}^2 + \sigma_{AI}^2)}, \quad w_B^0 = \frac{n_B(n_A\sigma_{AC}^2 + \sigma_{AI}^2)}{n_A(n_B\sigma_{BC}^2 + \sigma_{BI}^2) + n_B(n_A\sigma_{AC}^2 + \sigma_{AI}^2)},$$

$\bar{X}_A = \sum_{j=1}^{n_A} X_j^A / n_A$ *and* $\bar{X}_B = \sum_{j=1}^{n_B} X_j^B / n_B$*.*

The proof is in the Appendix, and it is a standard Bayesian posterior calculation. The optimal point estimate for the realized value of T is based on the averages of signals from each group. All signals from a given group, including those whose realized values happen to be extreme, receive the same weight in the estimate. Moreover, for a given set of parameters of the information generation process, the optimal weights are independent of the values of the observed signals.

Based on the posterior distribution of this estimator, we can also calculate the optimal combination of signals from the two groups when one receives N signals in total.

Proposition 2. *Suppose that the prior of* T *is* $T \sim N(\mu_p, \sigma_p^2)$*. Suppose that the total number of signals a subject receives is fixed at* N*. The number of* A *signals that minimizes the posterior variance under* $\sigma_p^2 \to \infty$ *is either the floor or ceiling of*

$$\frac{\sigma_{AI}\sigma_{BI}^2 + \sigma_{AI}\sigma_{BC}^2 N - \sigma_{BI}\sigma_{AI}^2}{\sigma_{AI}\sigma_{BC}^2 + \sigma_{BI}\sigma_{AC}^2}$$

if the value of the above formula is in the interval $(0, N)$*. It is* $n_A = N$ *if the above formula exceeds* N *and* $n_A = 0$ *if the above formula gives a negative value.*

The proof is in the Appendix, and the proposition is obtained by solving the first order condition for maximizing the inverse of the posterior variance. An implication of this result is that when both groups provide independent signals, i.e., $\sigma_{AC}^2 = \sigma_{BC}^2 = 0$, the subject should choose signals only from the group with lower σ_{lI}^2. When one group of signals are independent and the other correlated, the optimal number of correlated signals when we fix the number of total signals, is independent of N (as long as N is large enough). For example, if $\sigma^2_{AC} > 0$ and $\sigma^2_{BC} = 0$, the optimal number of group A signals is given by the floor or ceiling of $(\sigma_{AI}\sigma_{BI} - \sigma^2_{AI})/\sigma^2_{AC}$.

Finally, we consider a case where subjects receive additional information about T in addition to the N signals and provide a revised prediction. Suppose a subject first reports her estimate of the realized value of T, based on the N private signals she receives. We refer to this estimate as her *initial* prediction. Now suppose $M - 1 \geq 1$ other subjects also receive N signals each about the realized value of T and report their initial predictions. Each of these M subjects observe a completely different set of signals, but their signals are generated using the same information generation process. After all subjects report their initial predictions, suppose all of them are informed of the publicly available session average, i.e. the average of the M subjects' estimates of realized T based on their private signals.

Optimal estimate of the realization of T with this public information depends on a subject's belief about how other subjects generate their initial predictions. Thus, while the session average provides information based on many more signals, the exact values of those signals and how these signals are processed by other subjects are ambiguous. This ambiguity may lead to a subject behaving as if other subjects' initial predictions have higher variance than her own initial prediction. If a subject believes that the variance of all the subjects' estimates are the same, the most efficient estimator is independent of her private signals and would equal just the session average. The optimal strategy is to follow the "wisdom of the crowd." On the other hand, ambiguity about other subjects' information aggregation rules may lead subjects to attribute a much larger variance to the session average. In that case, the optimal estimate will be a weighted average of the subject's own initial estimate and the session average.

The following proposition formalizes the above argument. Note that the effect of ambiguity is summarized by λ that inflates the variance of the average of other subjects' predictions.

Proposition 3. *Suppose that a subject's initial prediction is unbiased, her beliefs follow a normal distribution, and she believes that other subjects' beliefs also follow a normal distribution with unbiased initial predictions and the same variance as the subject's own initial prediction. Then, the posterior mean with the public information equals the session average. However, if a subject's belief about the session average can be summarized by all other initial predictions being unbiased with variance of $\lambda\sigma^2$ where σ^2 is the variance of her own initial prediction, the posterior mean is $(M/(M-1+\lambda))\bar{p} + (((\lambda-1)/M-1+\lambda))p^*$, where M is the total number of subjects, $\bar{p}$ is the session average and p^* is her own initial prediction.*

The parameter λ can be estimated using subjects' revised prediction. We will refer to this as an implied variance parameter. A subject's belief about λ may represent her beliefs about the precision of other subject's initial estimates, her attitude towards ambiguity, or a number of other different factors. If we adapt the maxmin theory of ambiguity aversion (Gilboa and Schmeidler, 1989), then we may interpret $\lambda\sigma^2$ as the variance of other predictions under the worst-case prior.

The above results provide us with a theoretical framework for our experiments. We will compare our experimental results with the theoretical predictions generated using these results.

3 Experimental Design

We ran a set of laboratory experiments to empirically investigate how people incorporate independent and correlated signals in forming their beliefs about a random variable. Subjects received information about the earning per share (EPS) from a different fictitious stock in each period of a session. We elicited their beliefs about the EPS for each stock using the binarized scoring rule (BSR) proposed in Hossain and Okui (2013), which is incentive compatible independently of a subject's risk preference. Elicited beliefs allow us to directly investigate how people process information that differ in precision and correlation structure.

We ran 10 laboratory sessions under five different sets of information generation parameters. Each session consisted of 60 belief-elicitation periods. The true EPS, T, in each period was independently drawn from a normal distribution with mean 500 and variance 25,000. In each period, subjects received 10 forecasts about the EPS of a stock, where stocks across periods were independent of each other.[7] In addition to the 10 forecasts, subjects were also shown the averages of group A, group B, and all 10 forecasts. The forecasts were generated following the process described in Section 2. Subjects were completely aware of this signal generating process, including the precision level and correlation structure, used in their session. Given the parameters chosen, we had no case where the true EPS or any of the generated forecasts were negative. All subjects within a session received information about same stock in a given period. However, for each subject, we generated a completely different set of forecasts based on the above process. We clearly informed them that information received in one period was not informative about the realized values of the random variables in the other periods.

Using the notations introduced in Section 2, Table 1 presents the five sets of parameters that we used in our experiment. We ran two sessions with each of the five sets of information generation parameters. In three of the five sets of parameters, all forecasts in one of the groups, conditional on the true EPS, were independently drawn and the other group's forecasts, even after conditioning on the true EPS, were correlated. For the fourth set, forecasts from each group had a common error term and an independent error term. That is, error components of all forecasts from group A were correlated to each other and error components of all forecasts from group B were correlated to each other, although group B forecasts were independent of group A forecasts, conditional on the EPS. In the sessions with the fifth set of the parameters, all error terms (for both groups A and B) were drawn independently.

The specifications differ in the level of correlation, which we quantify by the "correlation ratio"—

[7] We will use the terms forecasts and signals interchangably.

Table 1: Set of Parameters under the Five Specifications

Specification	σ^2_{AI}	σ^2_{AC}	σ^2_{BI}	σ^2_{BC}	Characterization
1	500	0	250	250	Group A independent and group B moderately correlated
2	500	0	250	15	Group A independent and group B weakly correlated
3	500	0	15	250	Group A independent and group B strongly correlated
4	250	15	15	250	Group A weakly correlated and group B strongly correlated
5	500	0	265	0	Both groups independent

$\frac{\sigma^2_{lC}}{\sigma^2_{lA}+\sigma^2_{lC}}$. For example, forecasts from group B under Specifications 3 and 4 are considered strongly correlated as the correlation ratio is close to 0.95. On the other hand, forecasts from group B under Specification 2 and group A under Specification 4 are weakly correlated with a correlation ratio of below 0.06. We consider forecasts from group B under Specification 1, which has a correlation ratio of $\frac{1}{2}$, to be moderately correlated.

In a given period, we asked a subject to report her belief about the realization of the EPS of the stock of that period. First, we asked her to predict the EPS based on the 10 forecasts she received using the following incentive scheme: Let us denote this *initial prediction* by P. If the square of the deviation of this initial prediction from the true EPS, $(P-T)^2$, was below or equal to a random number K, generated from a uniform distribution on $[0, \bar{K}]$, she earned 100 points.[8] If $(P-T)^2$ was above K, she earned no point. We informed the subjects that it is optimal for them to report the mean of their posterior belief about T based on the 10 forecasts. We did not tell them how to optimally construct their posterior belief, which is given in Proposition 1, as learning how subjects actually form their beliefs is the main objective of this experiment.

After entering her initial prediction P, each subject was informed of the *session average* for that period, i.e., the average of the predictions for that EPS by all subjects (including herself) in that session. Then we asked them to enter a *revised prediction*, denoted by P^r, for the same EPS. We again incentivized truthful revelation of belief using the BSR with a different random number K^r, generated from a uniform distribution on $[0, \bar{K^r}]$.[9] The payoff from a revised prediction is 50 or zero points. After the subjects reported their initial and revised predictions, we reported the true EPS for that period, realizations of the relevant random numbers (K and K^r) and their income in points to the subjects. The random numbers K and K^r were redrawn in each period.

We divided the 60 periods within a session into six 10-period long blocks. In the first session for each specification, the number of forecasts from group A analysts increased, with an increment of 2, from 1 to 9 in the first five blocks of 10 periods each. In the second session, the number of forecasts from group A analysts decreased, with a deduction of 2, from 9 to 1 in the first five blocks. In the last block, i.e. periods 51 to 60, each subject could choose the number of group A analysts in each period. After all 60 periods

[8] The values of $\bar{K}$ vary across specifications and they are 40, 30, 40, 50 and 40. We chose these values to equalize simulated earnings from hypothetical, yet plausible, suboptimal strategies across specifications.

[9] The values of $\bar{K^r}$ for Specifications 1-5 are 8, 6, 8, 10, and 8, respectively. We chose these values to equalize simulated earnings across specifications.

were completed, we chose six periods, one randomly from each block of 10 consecutive periods, to determine subjects' incomes from the session.

The sessions were run between December 2015 and September 2016. The first six sessions were run in the Center for Research in Experimental Economics and Political Decision Making (CREED) at the University of Amsterdam and the other four were run at the University of Toronto. The number of subjects in a session varied between 12 and 22. In total, we had 172 subjects.[10] The experiments were programmed and conducted with the software *z-Tree* developed by Fischbacher (2007). Instructions were provided in English.[11] Subjects were sent a pdf document describing statistical concepts such as mean, variance, and uniform and normal distributions, that are relevant for the experiment, after they signed up for a session. They were asked to ensure that they understood those concepts before coming to the session. They were given a copy of the document during the experiment. Each subject had access to a calculator throughout her session. A sample instruction from the experiment and the document providing definitions of statistical concepts are provided in Appendix A.5.

4 Results

The experimental sessions have provided us with a large data set of people's beliefs about the realized EPS under a comprehensive set of information generation processes. Our main goal is to study how people value independent and correlated forecasts in coming up with their aggregate belief. Moreover, we also study how they incorporate additional information based on the actions of others. Finally, we study how subjects choose the combination of forecasts in the last 10 periods of the session.

Recall that the number of group A forecasts in the first five 10-period blocks was increasing over time in half of the sessions and decreasing in the other half of the sessions. We find no effect of this difference in ordering.[12] Hence, we pool the results from the increasing and decreasing order sessions in our empirical analysis.

4.1 Initial Prediction with Only Private Information

We analyze the data of initial predictions to examine how subjects reacted to different signal structures. We present estimates of weights on the averages of the two groups of forecasts using a linear regression model. Our main finding is that while initial predictions are unbiased, subject put suboptimally high weights on strongly correlated forecasts and suboptimally low weights on weakly correlated forecasts.

[10]Four subjects left during the middle of the session for unexpected reasons. We do not include data from those subjects in our analysis.

[11]Most of the subjects at CREED are students of the University of Amsterdam. Both English and Dutch are used as media of instruction at the University of Amsterdam. In general, their English level is very high.

[12]Results are available from the authors upon request.

Table 2: Initial Prediction, Specification 1

Dependent variable: *Initial Prediction*					
# of group A analysts	1	3	5	7	9
$\bar{X}_A$	0.263***	0.405***	0.569***	0.699***	0.820***
	(0.027)	(0.033)	(0.041)	(0.035)	(0.047)
$\bar{X}_B$	0.752***	0.577***	0.447***	0.304***	0.141***
	(0.033)	(0.030)	(0.029)	(0.024)	(0.018)
Constant	0.148	-0.809*	0.282	-0.480	-0.416
	(0.437)	(0.301)	(0.239)	(0.291)	(0.411)
Optimal weight on $\bar{X}_A$	0.357	0.632	0.75	0.824	0.9
Optimal weight on $\bar{X}_B$	0.643	0.368	0.25	0.176	0.1
F-stats:					
H_0: prediction	0.247	3.819	0.770	1.362	1.155
is unbiased	(0.782)	(0.030)	(0.469)	(0.267)	(0.325)
H_0: weights are	7.353	27.735	23.495	16.801	3.027
optimal	(0.002)	(0.000)	(0.000)	(0.000)	(0.059)
H_0: weights are	19.244	8.419	1.908	0.012	3.027
proportional to #	(0.000)	(0.001)	(0.161)	(0.988)	(0.059)
of analysts					
$\bar{R}^2$	0.757	0.745	0.747	0.688	0.516
# of observations	431	433	436	436	434

Note: Estimated by OLS. Standard errors clustered at subject level are presented below the coefficients (within parentheses) and p-values are presented below the F statistics (within parentheses). "H_0 : prediction is unbiased" corresponds to the restriction that the constant is zero and the sum of the coefficients on $\bar{X}_A$ and $\bar{X}_B$ equals 1. "H_0 : weights are optimal" corresponds to the restriction that the coefficients on $\bar{X}_A$ and $\bar{X}_B$ equal "Optimal weight on $\bar{X}_A$" and "Optimal weight on $\bar{X}_B$", respectively. "H_0 : weights are proportional to # of analysts" corresponds to the restriction that the coefficients on $\bar{X}_A$ and $\bar{X}_B$ equal $n_A/10$ and $1 - n_A/10$, respectively, where n_A is the number of group A analysts.

4.1.1 Econometric Specification and Some Hypotheses

As we have five different specifications with varying levels of correlation among forecasts, we analyze them separately when investigating how subjects process 10 forecasts to create an initial prediction about the EPS. In this analysis, we include only the first 50 periods of each session as in these periods the combination of analysts of the two groups were the same across all subjects within the session. For each subject, we have 10 observations under the same specification and analyst combination where each observation is for a different, independently drawn, stock.[13] We use these observations to estimate the prediction rules used by subjects for a given specification and forecaster combination. An initial examination of the data indicates that initial predictions can be modeled to be linear in the mean forecasts from the two groups. Specifically, neither nonlinear, such as quadratic, terms nor extreme values, such as maximum or minimum forecasts, systematically determine initial predictions. We thus focus on specifications that are linear in the means of group A and group B forecasts, which are denoted by $\bar{X}_A$ and $\bar{X}_B$, respectively.

For each specification and forecaster combination, we estimate the weights on $\bar{X}_A$ and $\bar{X}_B$ using the

[13]We exclude observations where the initial prediction is above the maximum or below the minimum of the 10 forecasts a subject received as these observations are hardly justified by any economic theory and are likely to be typos or results of simple mistakes. These exclusions do not affect our results qualitatively.

Table 3: Initial Prediction, Specification 2

Dependent variable: *Initial Prediction*					
# of group A analysts	1	3	5	7	9
$\bar{X}_A$	0.198***	0.339***	0.493***	0.647***	0.891***
	(0.030)	(0.027)	(0.025)	(0.028)	(0.037)
$\bar{X}_B$	0.884***	0.640***	0.476***	0.329***	0.104***
	(0.036)	(0.033)	(0.021)	(0.028)	(0.027)
Constant	-0.916*	-0.309	-0.247	-0.006	-0.525
	(0.351)	(0.273)	(0.300)	(0.234)	(0.364)
Optimal weight on $\bar{X}_A$	0.079	0.233	0.394	0.579	0.827
Optimal weight on $\bar{X}_B$	0.921	0.767	0.606	0.421	0.173
F-stats:					
H_0: prediction	6.428	0.998	1.121	0.261	1.095
is unbiased	(0.004)	(0.380)	(0.338)	(0.772)	(0.346)
H_0: weights are	11.193	10.007	33.405	6.778	4.281
optimal	(0.000)	(0.000)	(0.000)	(0.003)	(0.022)
H_0: weights are	8.922	1.811	0.623	1.944	0.033
proportional to #	(0.001)	(0.179)	(0.543)	(0.159)	(0.968)
of analysts					
$\bar{R}^2$	0.634	0.709	0.741	0.699	0.618
# of observations	335	337	339	337	335

Note: Estimated by OLS. Standard errors clustered at subject level are presented below the coefficients (within parentheses) and p-values are presented below the F statistics (within parentheses). "H_0 : prediction is unbiased" corresponds to the restriction that the constant is zero and the sum of the coefficients on $\bar{X}_A$ and $\bar{X}_B$ equals 1. "H_0 : weights are optimal" corresponds to the restriction that the coefficients on $\bar{X}_A$ and $\bar{X}_B$ equal "Optimal weight on $\bar{X}_A$" and "Optimal weight on $\bar{X}_B$", respectively. "H_0 : weights are proportional to # of analysts" corresponds to the restriction that the coefficients on $\bar{X}_A$ and $\bar{X}_B$ equal $n_A/10$ and $1 - n_A/10$, respectively, where n_A is the number of group A analysts.

following equation:

$$(P_{it} - T_t) = \beta_0 + \beta_1(\bar{X}_{A,it} - T_t) + \beta_2(\bar{X}_{B,it} - T_t) + u_{it},$$

where, at period t, T_t is the true value of EPS, and for subject i, P_{it} is the initial prediction, $\bar{X}_{A,it}$ and $\bar{X}_{B,it}$ are the means of groups A and B forecasts, respectively, and u_{it} is the error term. We use deviations from the EPS to estimate the weights. In our experiment, the EPSs are generated from a very diverse distribution and most of the variations in predictions and forecasts come from variations in the EPS. Using values which are not centered around the EPS thus yields a very high coefficient of determination, which might make results difficult to interpret. Using deviations is a solution to this problem.[14] Note that centering the variables around the EPS would not change the results if predictions are unbiased.[15]

Values of the coefficients, $(\beta_0, \beta_1, \beta_2)$, relate to various hypotheses.

Hypothesis 1 (Unbiased prediction): An initial prediction being unbiased corresponds to $H_0 : \beta_0 = 0$

[14] An alternative solution may be to include period fixed effects. Although not presented here, the model with uncentered variables but with period fixed effects generates similar results.

[15] Indeed, estimating the model with uncentered variables yields similar results except for the coefficients of determination and does not alter our conclusions qualitatively.

Table 4: Initial Prediction, Specification 3

Dependent variable: *Initial Prediction*					
# of group A analysts	1	3	5	7	9
$\bar{X}_A$	0.284***	0.419***	0.547***	0.662***	0.786***
	(0.036)	(0.046)	(0.048)	(0.048)	(0.056)
$\bar{X}_B$	0.758***	0.502***	0.491***	0.386***	0.163***
	(0.031)	(0.047)	(0.030)	(0.046)	(0.037)
Constant	-0.744*	-0.281	0.562	0.212	0.945**
	(0.306)	(0.399)	(0.367)	(0.289)	(0.337)
Optimal weight on $\bar{X}_A$	0.355	0.602	0.717	0.781	0.827
Optimal weight on $\bar{X}_B$	0.665	0.398	0.283	0.219	0.173
F-stats:					
H_0: prediction	4.455	3.388	1.248	1.291	3.967
is unbiased	(0.019)	(0.046)	(0.300)	(0.289)	(0.029)
H_0: weights are	4.824	8.557	24.758	6.613	0.459
optimal	(0.015)	(0.001)	(0.000)	(0.004)	(0.636)
H_0: weights are	13.282	9.332	0.480	2.054	2.486
proportional to #	(0.000)	(0.001)	(0.623)	(0.144)	(0.099)
of analysts					
$\bar{R}^2$	0.787	0.637	0.720	0.669	0.510
# of observations	309	339	332	336	333

Note: Estimated by OLS. Standard errors clustered at subject level are presented below the coefficients (within parentheses) and p-values are presented below the F statistics (within parentheses). "H_0 : prediction is unbiased" corresponds to the restriction that the constant is zero and the sum of the coefficients on $\bar{X}_A$ and $\bar{X}_B$ equals 1. "H_0 : weights are optimal" corresponds to the restriction that the coefficients on $\bar{X}_A$ and $\bar{X}_B$ equal "Optimal weight on $\bar{X}_A$" and "Optimal weight on $\bar{X}_B$", respectively. "H_0 : weights are proportional to # of analysts" corresponds to the restriction that the coefficients on $\bar{X}_A$ and $\bar{X}_B$ equal $n_A/10$ and $1 - n_A/10$, respectively, where n_A is the number of group A analysts.

and $\beta_1+\beta_2 = 1$. Note that unbiasedness is fundamental in investigating the validity of other hypotheses.

Hypothesis 2 (Reaction to correlation): Optimal initial predictions correspond to $H_0 : \beta_0 = 0$, $\beta_1 = w_A^0$ and $\beta_2 = w_B^0$, where w_A^0 and w_B^0 are the optimal weights (in Proposition 1) on the group A and group B averages, respectively. In Specifications 1-4, this hypothesis holds when subjects take the correlation structure into account perfectly. There are several ways in which this hypothesis is violated and, assuming the unbiasedness, each violation has a specific interpretation. In particular, we consider the following alternatives.

H_1 **(Correlation neglect):** $\beta_1 < w_A^0$ and $\beta_2 > w_B^0$ in Specifications 1-3 with complete correlation neglect suggesting the weights are calculated by attributing σ_{lC}^2 to σ_{lI}^2.

H_1 **(Correlation over-adjustment):** $\beta_1 > w_A^0$ and $\beta_2 < w_B^0$ in Specifications 1-3.

H_1 **(Under-adjustment for strong correlation and over-adjustment for weak correlation):** $\beta_1 < w_A^0$ and $\beta_2 > w_B^0$ in Specification 4.

Hypothesis 3 (Reaction to precision): Optimal initial predictions correspond to $H_0 : \beta_0 = 0$, $\beta_1 = w_A^0$ and $\beta_2 = w_B^0$. In Specification 5, this hypothesis holds when subjects take the degree of precision,

Table 5: Initial Prediction, Specification 4

Dependent variable: *Initial Prediction*					
# of group A analysts	1	3	5	7	9
$\bar{X}_A$	0.250***	0.350***	0.530***	0.618***	0.643***
	(0.034)	(0.043)	(0.045)	(0.047)	(0.044)
$\bar{X}_B$	0.709***	0.573***	0.490***	0.408***	0.310***
	(0.044)	(0.063)	(0.021)	(0.037)	(0.041)
Constant	0.627	0.366	0.193	-0.012	-1.271**
	(0.425)	(0.470)	(0.342)	(0.317)	(0.419)
Optimal weight on $\bar{X}_A$	0.487	0.719	0.796	0.834	0.861
Optimal weight on $\bar{X}_B$	0.513	0.281	0.206	0.166	0.139
F-stats:					
H_0: prediction	1.661	1.808	0.219	0.154	4.861
is unbiased	(0.209)	(0.183)	(0.805)	(0.858)	(0.016)
H_0: weights are	26.132	37.870	95.085	24.201	15.616
optimal	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
H_0: weights are	13.686	2.096	0.268	4.527	22.382
proportional to #	(0.000)	(0.142)	(0.767)	(0.020)	(0.000)
of analysts					
$\bar{R}^2$	0.688	0.639	0.780	0.790	0.677
# of observations	251	273	276	277	266

Note: Estimated by OLS. Standard errors clustered at subject level are presented below the coefficients (within parentheses) and p-values are presented below the F statistics (within parentheses). "H_0 : prediction is unbiased" corresponds to the restriction that the constant is zero and the sum of the coefficients on $\bar{X}_A$ and $\bar{X}_B$ equals 1. "H_0 : weights are optimal" corresponds to the restriction that the coefficients on $\bar{X}_A$ and $\bar{X}_B$ equal "Optimal weight on $\bar{X}_A$" and "Optimal weight on $\bar{X}_B$", respectively. "H_0 : weights are proportional to # of analysts" corresponds to the restriction that the coefficients on $\bar{X}_A$ and $\bar{X}_B$ equal $n_A/10$ and $1 - n_A/10$, respectively, where n_A is the number of group A analysts.

i.e. the variances, perfectly into account. Assuming unbiasedness, there are two ways in which this hypothesis is violated.

H_1 **(Precision neglect):** $\beta_1 > w_A^0$ and $\beta_2 < w_B^0$ in Specification 5, with complete precision neglect leading to $w_l^0 = n_l/N$.

H_1 **(Precision over-adjustment):** $\beta_1 < w_A^0$ and $\beta_2 > w_B^0$ in Specification 5.

In addition, comparison of the coefficient estimates from Specifications 2 and 3 and those from Specification 5 provides insights about the relative effects of precision and correlation on suboptimal predictions.

4.1.2 Detailed Discussion of Results

Tables 2 to 6 present the estimates for Specifications 1 to 5, respectively.[16] We also present a number of F-tests regarding the weights. Specifically, we test whether initial predictions can be considered unbiased,

[16] Fixed effects estimation leads to the same results, qualitatively. They are available from the authors upon request. Note that fixed effects regression model specifies that the intercept is individual specific. In our context, this means that the bias in the prediction is individual specific. It is another indication of unbiasedness of the predictions that inclusion of fixed effects does not alter the results qualitatively.

Table 6: Initial Prediction, Specification 5

Dependent variable: *Initial Prediction*					
# of group A analysts	1	3	5	7	9
$\bar{X}_A$	0.141***	0.271***	0.435***	0.503***	0.726***
	(0.024)	(0.023)	(0.027)	(0.041)	(0.044)
$\bar{X}_B$	0.820***	0.692***	0.596***	0.491***	0.187***
	(0.062)	(0.050)	(0.040)	(0.042)	(0.037)
Constant	-0.184	0.242	-0.183	-0.423	-0.166
	(0.353)	(0.334)	(0.303)	(0.350)	(0.258)
Optimal weight on $\bar{X}_A$	0.056	0.185	0.346	0.553	0.827
Optimal weight on $\bar{X}_B$	0.944	0.815	0.654	0.447	0.173
F-stats:					
H_0: prediction	0.365	1.488	0.322	0.850	3.153
is unbiased	(0.697)	(0.242)	(0.727)	(0.437)	(0.057)
H_0: weights are	6.565	7.321	5.353	0.807	3.631
optimal	(0.004)	(0.002)	(0.010)	(0.456)	(0.038)
H_0: weights are	1.666	1.142	4.060	13.563	7.887
proportional to #	(0.206)	(0.332)	(0.027)	(0.000)	(0.002)
of analysts					
$\bar{R}^2$	0.491	0.648	0.722	0.644	0.553
# of observations	316	318	317	316	317

Note: Estimated by OLS. Standard errors clustered at subject level are presented below the coefficients (within parentheses) and p-values are presented below the F statistics (within parentheses). "H_0 : prediction is unbiased" corresponds to the restriction that the constant is zero and the sum of the coefficients on $\bar{X}_A$ and $\bar{X}_B$ equals 1. "H_0 : weights are optimal" corresponds to the restriction that the coefficients on $\bar{X}_A$ and $\bar{X}_B$ equal "Optimal weight on $\bar{X}_A$" and "Optimal weight on $\bar{X}_B$", respectively. "H_0 : weights are proportional to # of analysts" corresponds to the restriction that the coefficients on $\bar{X}_A$ and $\bar{X}_B$ equal $n_A/10$ and $1 - n_A/10$, respectively, where n_A is the number of group A analysts.

whether the weights are optimal in the sense of Proposition 1, and whether the weights are proportional to the number of forecasters from the relevant group; that is, whether the weight on $\bar{X}_l$ equals $n_l/10$.

The tables provide a complete picture of how the subjects chose predictions using their private signals, which will be discussed below. First, we highlight one particular result: As seen in the literature, we find that subjects do not completely account for correlation. However, our result is more nuanced. When the correlation is moderate or strong, subjects put greater weight on correlated forecasts than optimal. On the other hand, when correlation among forecasts is weak, they put suboptimally *low* weights on the weakly correlated forecasts. In other words, subjects over-compensate for correlation when correlation is weak.

Now we discuss the main results regarding initial predictions in detail and also discuss and test a number of hypotheses. We cannot reject that the initial predictions are *unbiased* (Hypothesis 1) in most cases.[17] That is, initial prediction P can be expressed as $P = w\bar{X}_A + (1 - w)\bar{X}_B$ for some $w \in [0, 1]$. We also tested regression specifications where extreme values (maximum and/or minimum) of the forecasts (for both groups or separately) are added as explanatory variables. We rarely find any of those extreme values to be

[17]The null hypothesis is rejected in 6 and 7 out of 25 cases at the 5% and 10% levels, respectively. However, because there are 25 cases, the testing procedure should be adjusted. A simple Bonferroni correction indicates that we should reject the null when we find a p-value less than $0.05/25 = 0.002$ but there is no such case.

statistically significant. Even including them, the predictions could be characterized as a weighted average of $\bar{X}_A$ and $\bar{X}_B$.[18] This may not seem surprising as we restricted attention to observations where the initial prediction was within the minimum and maximum of the forecasts a subject received. It is nonetheless noteworthy because even simple but systematic mistakes such as over-weighting extreme values or habits such as always rounding up predictions may lead to biased predictions. Confirming that the initial predictions are unbiased is an important step and the rest of the discussions rely on this finding.

While initial predictions are unbiased, they are typically suboptimal (Hypotheses 2 and 3). Frequently, we reject the hypothesis that the weights are chosen optimally. Nonetheless, we notice a clear pattern in the departure from optimal prediction. Recall that, in Specifications 1 to 3, only group A forecasts are independent with $\sigma^2_{AI} \geq \sigma^2_{BI} + \sigma^2_{BC}$. In Specification 4, forecasts from both groups are correlated with $\sigma^2_{AI} + \sigma^2_{AC} = \sigma^2_{BI} + \sigma^2_{BC}$ and $\sigma^2_{AC} < \sigma^2_{BC}$. On the other hand, in Specification 5, forecasts from both groups are independent with $\sigma^2_{AI} > \sigma^2_{BI}$. For Specifications 1, 3, and 4, subjects chose sub-optimally high weights for group B forecasts. For example, the estimated weight on the mean of B forecasts is 0.742 under Specification 1 with one A analyst while the optimal weight is 0.643. On the other hand, for Specifications 2 and 5, they chose sub-optimally low weights for group B forecasts. For example, the estimated weight on the mean of B forecasts is 0.805 under Specification 2 with one A analyst while the optimal weight is 0.921.

These results are connected to the alternative hypotheses we have considered. When one group of forecasts was correlated, subjects overweighted correlated forecasts when the level of correlation was moderate or strong, as defined above (H_1 Correlation neglect). However, they overweighted independent forecasts when correlation was weak (H_1 Correlation over-adjustment). When both groups were correlated, following the above pattern, subjects underweighted the weakly correlated forecasts (H_1 Under-adjustment for strong correlation and over-adjustment for weak correlation). Even when both groups of forecasts were independent, subjects did not choose the weights optimally. In fact, they chose suboptimally high weights for the forecasts with higher variance (H_1 Precision neglect). Note that all of these patterns hold true for all five analyst combinations for a given specification. Thus, the nuances in the results come from differences in how subjects treat the variance structures of the forecasts.

While our results suggest that subjects cannot find optimal weights for different types of forecasts even when forecasts are independent, we can reject the hypothesis that subjects completely neglect correlation. If subjects neglected correlation completely, then they would value both groups of signals equally under Specifications 1 and 4 as $\sigma^2_{AI} + \sigma^2_{AC} = \sigma^2_{BI} + \sigma^2_{BC}$ in both cases. However, Tables 2 and 5 show that, for 6 out of the 10 combinations of analysts, we can reject that subjects chose proportional weights, i.e., the weights on $\bar{X}_l$ equals n_l/N, at the 10% level, while rejecting 4 out of the 10 at the 1 % level. Specification 4, where both groups of forecasts are correlated, illustrates that absence of complete correlation neglect is

[18]Results are available from the authors upon request.

not restricted to cases where only one set of forecasts are correlated. Moreover, complete correlation neglect would suggest that subjects use the same rule to generate predictions under Specifications 2, 3, and 5. We present pair-wise comparisons of the weights from these two specifications in the fifth, seventh, and ninth columns of Table 7. We can reject the null that subjects use the same prediction rules for Specifications 2 and 3 (for a given analyst combination) at the 2% level, in 2 out of 5 cases. In particular, when the number of A analysts is small and the optimal weights are quite different between the two specifications, we can easily reject that subjects use the same rulle.[19] Comparing each of Specifications 2 and 3 to Specification 5, we can reject the null that subjects use the same prediction rules at the 10% level, in 9 out of 10 cases.

The design of our session parameters allows us to compare how subjects incorporate the level of correlation in their predictions. Specifically we investigate the relative impact of correlation using Specifications 2, 3, and 5. Under all of these specifications, the total variance $(\sigma_{lI}^2 + \sigma_{lC}^2)$ is 500 for a group A forecast and 265 for a group B forecast. However, while group A forecasts are independent under all three specifications, the correlation ratio for group B forecasts ranges from 0 (Specification 5) to 0.94 (Specification 3). Therefore, if subjects are consistent in their weighing scheme, one would expect that, for any forecaster combination, the weight on group B forecasts will be highest under Specification 5, followed by Specification 2, and then Specification 3 (with the ranking being weak when there is only one group B forecast). Comparing Tables 3, 4, and 6 and using Table 7 for significance, we find this exact pattern (although the differences are not always statistically significant). These findings strongly suggest that although subjects do not choose weights optimally, their suboptimal behavior is not caused only by the complexity of calculating posterior distributions. Rather, they follow some formula that takes both the independent and correlated parts of the variances into account.

Recall that, under Specification 4, we can frequently reject that group A and group B signal averages receive proportional weights. This suggests that signals from the two groups are not valued the same way even though both groups were correlated and had the same total variances. This allows us to eliminate another potential alternative hypothesis about how people form beliefs using correlated signals. Suppose subjects realize that correlated signals are somewhat worse than independent signals of the same total variance, but do not understand that the optimal weight on a signal depends on the level of correlation. Such a heuristic might explain underweighting of weakly correlated signals and overweighting of strongly correlated signals. However, such a heuristic would suggest that subjects would choose the same prediction rule under both Specifications 2 and 3 and they would consider the two groups of signals under Specification 4 to be equally valuable. Our findings from the three specifications described above suggest that such a heuristic does not explain our results very well.

One may wonder if there is any evidence that subjects exhibit complete precision neglect by totally

[19] Table 7 presents F tests comparing the weights used across different specifications for all possible pairwise combinations and also for all specifications together.

Table 7: Test of Equivalence of Weights Used in Initial Predictions Across Specifications

# of A	1 vs 2	1 vs 3	1 vs 4	1 vs 5	2 vs 3	2 vs 4	2 vs 5	3 vs 4	3 vs 5	4 vs 5	All
1	3.841	0.304	0.644	11.835	4.263	4.116	3.368	1.536	11.957	5.410	5.799
	(0.022)	(0.738)	(0.526)	(0.000)	(0.014)	(0.017)	(0.035)	(0.216)	(0.000)	(0.005)	(0.000)
3	1.915	1.855	0.806	7.178	4.427	0.728	2.480	1.180	8.385	2.640	3.483
	(0.148)	(0.157)	(0.447)	(0.001)	(0.012)	(0.483)	(0.084)	(0.308)	(0.000)	(0.072)	(0.001)
5	1.973	1.056	1.067	8.346	0.755	0.542	4.225	0.050	3.933	4.221	2.482
	(0.140)	(0.348)	(0.345)	(0.000)	(0.470)	(0.582)	(0.015)	(0.952)	(0.020)	(0.015)	(0.011)
7	0.952	2.243	5.889	17.540	1.223	2.673	11.141	0.433	5.508	3.671	5.400
	(0.386)	(0.107)	(0.003)	(0.000)	(0.295)	(0.070)	(0.000)	(0.649)	(0.004)	(0.026)	(0.000)
9	1.190	0.280	12.175	1.659	2.198	17.646	5.008	6.955	0.519	4.363	5.052
	(0.305)	(0.756)	(0.000)	(0.191)	(0.112)	(0.000)	(0.007)	(0.001)	(0.595)	(0.013)	(0.000)

Note: This table presents the results of F tests for the equivalence of coefficients on $\bar{X}_A$ and $\bar{X}_B$ between specifications in the regressions presented in Tables 2 to 6. Column "All" presents the results of F tests for the null hypothesis that the weights used in the initial predictions are the same across all five signal specifications. In parentheses below F statistics are p-values.

ignoring variance. Specification 5, where both groups of forecasts are independent, offers the cleanest test. As seen in Section 4.1, subjects put suboptimally high weight on group A forecasts, which had higher variances than group B forecasts. Nonetheless, subjects frequently chose greater weights on group B forecasts than on group A forecasts, suggesting that they had noticed that the variance was higher for group A, but did not know exactly how much lower the weights on group A should be.

All these experimental evidence suggest that people are not "hyper-rational" and their "boundedly rational" behavior follows a specific pattern. Our third result is that, while subjects choose weights sub-optimally, they do not completely ignore correlation among forecasts or precision level of forecasts.

4.1.3 Directions for Future Theoretical Research

These results indicate that while subjects incorporate both variance and correlation in their estimation, they do not do so optimally. Here we suggest two possible directions to theorize our findings. This discussion is still speculative and we do not claim that they are the models that explain subject's behavior in our experiments.

The first possible direction would be to assume that subjects maximize their expected utility given their beliefs but these beliefs are incorrectly formed. Recall that the optimal prediction has the form $w_A^0 \bar{X}_A + w_B^0 \bar{X}_B$ and the weights can be written as

$$w_A^0 = \frac{V_B}{V_A + V_B} \text{ and } w_B^0 = \frac{V_A}{V_A + V_B},$$

where V_A and V_B are variances of $\bar{X}_A$ and $\bar{X}_B$ given T, respectively, such that

$$\bar{X}_A \sim N(T, V_A) = N\left(T, \frac{\sigma^2_{AI}}{n_A} + \sigma^2_{AC}\right) \text{ and } \quad \bar{X}_B \sim N(T, V_B) = N\left(T, \frac{\sigma^2_{BI}}{n_B} + \sigma^2_{BC}\right).$$

Suppose that a subject may not be able to compute the distributions correctly, but she acts as if the distributions were

$$\begin{aligned}\bar{X}_A \sim & N(T, \tilde{V}_A) = N\left(T, \left(\frac{\sigma^2_{AI}}{n_A}\right)^{\rho} + \left(\sigma^2_{AC}\right)^{\rho}\left(\frac{\sigma^2_{AI}}{\sigma^2_{AI} + \sigma^2_{AC}}\right)^{\eta}\right),\\ \bar{X}_B \sim & N(T, \tilde{V}_B) = N\left(T, \left(\frac{\sigma^2_{BI}}{n_A}\right)^{\rho} + \left(\sigma^2_{BC}\right)^{\rho}\left(\frac{\sigma^2_{BI}}{\sigma^2_{BI} + \sigma^2_{BC}}\right)^{\eta}\right),\end{aligned}$$

where ρ and η are parameters. The parameter ρ measures the degree of precision neglect or over–adjustment. The parameter η measures the degree of correlation neglect or over–adjustment. As an example, if we use $\rho = 0.8$ and $\eta = 0.3$, group A forecasts will be undervalued in Specifications 1, 3, and 4 and overvalued in Specifications 2 and 5, exactly like our experimental finding.

The second possible direction would be to develop a non-expected utility decision theoretic model to explain the findings. Levy and Razin (2018) consider a situation in which correlation structure is ambiguous. They argue that correlation neglect arises if prediction obtained by ignoring correlation is precise. Otherwise, people act as if the correlation structure was least favorable to them, which may be interpreted as correlation over-adjustment in some situations. Their setting is different from ours. In particular, in our experiments, subjects were completely informed of the correlation structure. Nonetheless, one may argue that knowing correlation structure mathematically may not be enough for them to process information correctly and this difficulty may be modeled as ambiguity in correlation.[20] Even if we follow this interpretation, the theoretical results of Levy and Razin (2018) are not fully consistent with our findings. For example, signal variance of the correlated signals is smaller in Specifications 2 and 3 than in Specification 1, but correlation neglect does not become stronger. We also find that incorporating precision neglect is needed to fit the data. It appears to be necessary to develop an alternative model that fits the setting in our experiments.

4.1.4 Robustness: Learning and Heterogeneity

Digging deeper into the data, we analyze how subject behavior changes over time under each specification. Figures 2 to 6 in Appendix A.4 present time series plots of weights on the mean of group A forecasts over periods within a 10-period block for each analyst combination. We run cross-sectional regressions of prediction on the mean of group A forecasts and the mean of group B forecasts for each period (all variables are recentered around true EPS), and use the OLS estimate of the coefficient on the mean of group A forecasts

[20] In the computer game "Guess the Correlation," http://guessthecorrelation.com/, one has to relate scatter plots to the associated correlation coefficients. The very existence of such a game may indicate the difficulty in interpreting correlation.

as "weights on A." The five panels within a figure present the time-series diagram when the number of group A analysts is 1, 3, 5, 7, and 9, respectively. There is no clear pattern or systematic indication of learning within each block. There is also no indication that subjects' behavior converge to the theoretical optimal. Overall, we do not learn much more about subject behavior by separating behavior in each period over what is presented in Tables 2 to 6.

We also analyze how the prediction generation processes vary across subjects. Specifically, we ran time series regressions of prediction on the means of groups A and B forecasts for each 10-period block for each subject separately. Comparing the weights on group A forecasts, we find that the weights across subjects are, perhaps unsurprisingly, quite dispersed and there are some outliers. Nonetheless, we do not find any clear pattern in the dispersion of weights. These results are available from the authors upon request.

4.2 Revised Prediction with Additional Public Information

We now analyze the revised predictions submitted by subjects after they learn the session average for that period—the average of initial predictions in that period by all subjects in the session. First, we ensure that the session averages of initial predictions are indeed better than individual initial predictions. We find that the session averages yield about three times higher probabilities of receiving the reward than initial predictions do in all of the sessions.[21] While not surprising, this confirms that the average judgment of the crowd is better than the judgment of the individuals who make up the crowd. We then analyze whether people actually use this aggregated information appropriately. To that end, we use a linear regression model to estimate the weights subjects put on their own initial predictions and session averages. We find that they put non-zero weights on initial predictions, which we interpret as a result of ambiguity aversion. This also suggests that even though there is wisdom in the judgment of the crowd overall, people do not fully utilize such wisdom. This result and the interpretation of ambiguity driving subjects' beliefs is consistent with De Filippis et al. (2017) who show that people overweight private information relative to social information.[22]

4.2.1 Econometric Specification and Some Hypotheses

As each subject gets a completely different set of forecasts for a given stock, the session average indirectly provides information from many more signals than the subject received before submitting the initial prediction. It is optimal to report the session average as the revised prediction when a subject believes that all subjects' initial predictions are equally precise. However, if she believes that other subjects' predictions have a higher variance due to ambiguity aversion or a belief that her own prediction process is better, she may rely on her private signals and her initial prediction in addition to the session average for this exercise.

[21] These results are available upon request.

[22] The fact that people receive social information before private information in their setting suggests that overweighting private information in our experiment cannot be attributed only to anchoring of beliefs to the information that the subjects received first.

To estimate the process subjects use to generate revised prediction, we relate a subject's revised prediction with her initial prediction and the session average. We report the results from linear regression models which are found to be reasonable after initial inspection of the data. More concretely, our estimation model is:

$$(RP_{it} - T_t) = \gamma_0 + \gamma_1(P_{it} - T_t) + \gamma_2(SA_t - T_t) + u_{it},$$

where RP_{it} and P_{it} are the revised and initial predictions, respectively, by subject i at period t, T_t is the true EPS and SA_t is the session average at period t (note that they are common to all the subjects and there is no subject indicator) and u_{it} is the error term. The theoretical result in Proposition 3 does not depend on the signal generation parameters and the combination of analysts. Hence, we can pool all the periods within a specification in these regressions.[23]

The value of the coefficients, $(\gamma_0, \gamma_1, \gamma_2)$, relate to various hypotheses.

Hypothesis 4 (Unbiased prediction): A prediction is unbiased when $\gamma_0 = 0$ and $\gamma_1 + \gamma_2 = 1$.

Hypothesis 5 (Theoretical optimal): An unbiased prediction is optimal when $H_0 : \gamma_1 = 0$ and $\gamma_2 = 1$.

We consider the following alternative,

H_1 **(Ambiguity aversion):** $\gamma_1 > 0$ and $\gamma_2 < 1$.

4.2.2 Detailed Discussions of Results

Table 8 presents the estimation results for the five specifications, one in each column. Subjects typically choose a revised prediction that is in between their initial prediction and the session average, leading to an unbiased estimate. The weights on own and session average predictions add up to one. Thus, we cannot reject Hypothesis 4. However, unlike what the theoretical optimal suggests, subjects put positive weight on their own initial prediction under all specifications. Moreover, the weights are relatively close to each other across specifications, ranging between 18.7% and 26.4%. Hence, we can reject Hypothesis 5 in support of ambiguity aversion (H_1 Ambiguity aversion) under all specifications.

Comparing these numbers across the five specifications (See Table 9), the weights on own initial predictions are statistically significantly different only between Specifications 3 and 5, where the weights on initial prediction were the lowest and the highest, respectively. While subjects' initial prediction generation processes vary quite a bit across specifications, the revised prediction generation processes do not vary as much. This is consistent with the theoretical predictions.

[23]We exclude observations where the revised prediction is above the maximum or below the minimum of the 10 forecasts she received and the session average. We also exclude periods in which at least one subject submitted an initial prediction that is above the maximum or below the minimum of the 10 forecasts she received because in such a period, the session average may not provide useful information. These exclusions do not affect our results significantly

Table 8: Estimation Results: Revised Prediction

Dependent variable: *Revised Prediction*	All	Spec. 1	Spec. 2	Spec. 3	Spec. 4	Spec. 5
Initial Prediction	0.229***	0.245***	0.241***	0.187***	0.211***	0.264***
	(0.014)	(0.027)	(0.030)	(0.022)	(0.037)	(0.030)
Session Average	0.775***	0.785***	0.739***	0.794***	0.763***	0.823***
	(0.021)	(0.047)	(0.051)	(0.035)	(0.053)	(0.039)
Constant	0.013	-0.121	0.060	-0.035	0.133	0.085
	(0.046)	(0.090)	(0.068)	(0.092)	(0.168)	(0.105)
F-stats:						
H_0: prediction	110.260	21.050	26.364	35.604	19.805	20.786
is optimal	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
H_0: prediction	0.064	1.872	0.650	0.433	0.460	2.809
is unbiased	(0.938)	(0.166)	(0.528)	(0.652)	(0.636)	(0.076)
$\bar{R}^2$	0.468	0.466	0.456	0.509	0.483	0.440
# of observations	7779	1989	1714	1351	1099	1626

Note: Estimated by OLS. Estimated by OLS. Standard errors clustered at subject level are presented below the coefficients (within parentheses) and p-values are presented below the F statistics (within parentheses). Session Average is the average of the initial predictions made by all subjects. "H_0 : prediction is optimal" corresponds to the restriction that the coefficient on Session Average is equal to one. "H_0 : prediction is unbiased" corresponds to the restriction that the constant is zero and the sum of the coefficients on Initial Prediction and Session Average equals 1.

Table 9: Test of Equivalence of Weights Used in Revised Predictions Across Specifications

	1 vs 2	1 vs 3	1 vs 4	1 vs 5	2 vs 3	2 vs 4	2 vs 5	3 vs 4	3 vs 5	4 vs 5	All
F stat	0.346	2.087	0.813	0.666	1.098	0.216	1.701	0.171	4.292	2.583	1.517
p-value	(0.708)	(0.127)	(0.445)	(0.515)	(0.336)	(0.806)	(0.186)	(0.843)	(0.015)	(0.078)	(0.154)

Note: This table presents the results of F tests for the equivalence of coefficients on Initial Prediction and Session Average in the regressions presented in Table 8. Column "All" presents the results of F tests for the null hypothesis that the weights used in the revised predictions are the same across all five signal specifications. In parentheses below the F statistics are p-values.

When a subject believes that all other subjects' initial predictions have the same variance as her initial prediction, then they should put all the weight on the session average. Nevertheless, if she believes that her initial prediction has a lower variance, this weight may be lower than one. Hence, using the weights subjects put on their initial prediction and the number of subjects in the session, we can estimate how much better a subject believes her own initial prediction is relative to the other subjects' initial predictions. If the weight a subject puts on her initial prediction implies that she believes that the variances of her own prediction and the prediction of any other participant in her session are σ^2 and $\lambda\sigma^2$ respectively, then we call λ the implied variance multiplier. As the number of subjects across sessions for a given specification may be different, Table 10 presents the estimates of implied variance multipliers for the 10 sessions separately.[24]

The implied variance multiplier varies across sessions as a session with a larger number of participants

[24]When a subject left in the middle of a session unexpectedly, we had a lab assistant enter the average of the 10 forecasts, which is an unbiased estimator of the true EPS, as that subject's initial prediction. Thus, from the viewpoint of other participants, the *session average* included initial predictions by all the subjects who had started the session initially. Hence, the number of subjects used in calculation of the implied variance multiplier from a session may be different from the number of subjects in other tables.

Table 10: Implied Variance Multipliers

Session	1	2	3	4	5	6
# of subjects	22	22	20	14	15	19
Implied variance	54.42	45.63	45.27	33.81	32.77	42.16

Session	7	8	9	10		
# of subjects	18	13	20	13		
Implied variance	40.16	29.17	44.33	27.19		

Note: These implied variances are computed from the coefficient estimate in the regression run for each session. The number of subjects in a session is that in the beginning of the session.

led to a higher implied variance multiplier. This suggests that subjects do not take the number of participants fully into account in calculating their revised predictions. Another interpretation is that a subject's worst case prior regarding the value of the session average is independent of the number of subjects in a session. Although behaving as if other subjects' predictions have a higher variance is consistent with a notion of overconfidence, the implied variance multiplier being increasing in the number of participants is more supportive of this arising from ambiguity aversion.

A subject believing that other subjects' initial predictions are less precise than her own initial prediction is similar in nature to the idea of *dismissiveness* in Eyster, Rabin, and Vayanos (2018) where traders believe that other traders' signals are less precise. As a result, in their setting, traders end up neglecting the information content of the market price. We show that even when we provide subjects with the average belief of the population, they neglect its information content. Eyster, Rabin, and Vayanos (2018) also assume that dismissive traders may mistakenly assume that other traders' signals are correlated. Similarly, our results can also be explained by subjects mistakenly believing that other subjects' initial predictions are correlated even though we explicitly mention that each subject sees an independent set of signals (conditional on the true value of the EPS).

4.2.3 Demographic Effect on Revised Prediction

Following the recent literature that explores gender effect on ambiguity aversion, we examine how weights on own prediction differ across demographic characteristics.[25] Table 11 presents the results of regressions separately run for each of four combinations of gender and whether they took a statistics course in university. Table 12 presents the results of hypothesis testings that compare the coefficients. We also examine the effects of risk aversion but we do not find any meaningful result.

It is interesting to observe that females put higher weights on own initial prediction than males did. While taking a statistics course increases the coefficient on session average, its effect is not statistically

[25] We also examine the effect of those characteristics on initial predictions. However, we do not find any systematic effect. The results are available upon request.

Table 11: Revised Prediction, Subsample Analyses across Gender and Statistical Knowledge

Dependent variable: *Revised Prediction*	Male None	Male Stat	Female None	Female Stat
Initial Prediction	0.227***	0.164***	0.282***	0.253***
	(0.035)	(0.021)	(0.026)	(0.025)
Session Average	0.784***	0.845***	0.729***	0.744***
	(0.041)	(0.032)	(0.045)	(0.046)
Constant	0.048	-0.032	0.032	0.018
	(0.093)	(0.060)	(0.128)	(0.086)
$\bar{R}^2$	0.511	0.503	0.450	0.461
# of observations	1461	2200	1421	2697

Note: Estimated by OLS. Standard errors clustered at subject level are presented below the coefficients (within parentheses). Each row presents the estimation results from the subsample specified in the row title. "Stat" stands for "have taken a statistics course" and "None" stands for "have not taken a statistics course." In parentheses below coefficient estimates are standard errors clustered at subject level. Session Average is the average of the initials predictions of all subjects in the given period.

Table 12: Test of Equivalence of Weights Used in Revised Predictions Across Gender and Statistical Knowledge

	N, M vs F	S, M vs F	M, S vs N	F, S vs N	N vs S	M vs F	F N vs M S
F stat	0.853	3.946	1.316	0.415	0.865	2.399	6.663
p-value	(0.428)	(0.021)	(0.271)	(0.661)	(0.486)	(0.052)	(0.002)

Note: This table presents the results of F tests for the equivalence of coefficients on Initial Prediction and Session Average in the regressions conducted with subsamples defined by the characteristics of the subjects. "N" stands for "have not taken a statistics course," "S" stands for "have taken a statistics course," "M" stands for "Male" and "F" stands for "Female". In each row, we present the result of the test comparing the coefficient estimates from the subsamples presented in the right and the left of "vs" in the subsample specified in the code before ",". For example, Column "N, M vs F" presents the result of the test comparing the coefficients from the subsample that consists of male subjects without having taken a statistics course and the subsample that consists of female subjects without having taken a statistics course. In parentheses below F statistics are p-values.

significant. Nonetheless, the difference between males who took a statistics course and females who have not taken a statistics course is remarkable and it seems that knowledge of statistics widens the gap between males and females.

The existing literature presents mixed evidence on the effect of ambiguity aversion. Borghans et al. (2009) find that males are more ambiguity averse while Friedl, Ring, and Schmidt (2017) obtain an opposite result in some settings. If putting high weight on one's own initial prediction is indicative of ambiguity aversion, then our results suggest that females are more ambiguity averse. One may also interpret such behavior as females exhibiting more "cautious" behavior a la Cerreia-Vioglio, Dillenberger, and Ortoleva (2015). We would like to leave this issue for future research. An alternative interpretation of our results is that weights on own prediction represent the degree of confidence a subject has on their abilities to generate better forecasts relative to those by others. Our results would suggest that female subjects are more confident on their own ability. This, however, will be in contrast with the literature on overconfidence that indicates that males are more overconfident than females (see, e.g., Barber and Odean (2001) and Coffman (2014)).

Table 13: Chosen Number of Group A Analysts

Specification	1	2	3	4	5
Optimal # of A analysts	10 or 9	3	9	9	0
Mean	6.36	4.59	7.19	4.73	4.51
Standard Deviation	3.04	3.22	2.65	2.62	3.29
# of observations	440	340	340	280	320

4.2.4 Robustness: Learning and Heterogeneity

To analyze how subjects' estimation process for revised prediction changes over time, Figure 7 in Appendix A.4 presents a time series plot of weights on session average. We run cross-section regressions of revised prediction on own initial prediction and session average for each period, and present the OLS estimate of the coefficient on the session average. Again, there is no clear pattern. Here we pool all the specifications together as the optimal weights are the same in all of the specifications. We find relatively low weights in periods 3 to 5 in this pooled picture and one might worry that this affects other results, or it indicates learning over time. However, looking at time series plots for each specification separately, we did not find any clear pattern nor an indication of learning. Moreover, the above results on the revised prediction do not change qualitatively if we drop the first 10 periods.[26]

We also analyze how heterogeneous the prediction generation processes is across subjects. Specifically, we ran time series regressions of revised prediction on initial prediction and session average for each subject. The weights across subjects are, as in the case of initial predictions, quite dispersed and there are some outliers. Nonetheless, we are unable to find any factor beyond gender and knowledge of statistics, that affects heterogeneity. These result are available from the authors upon request.

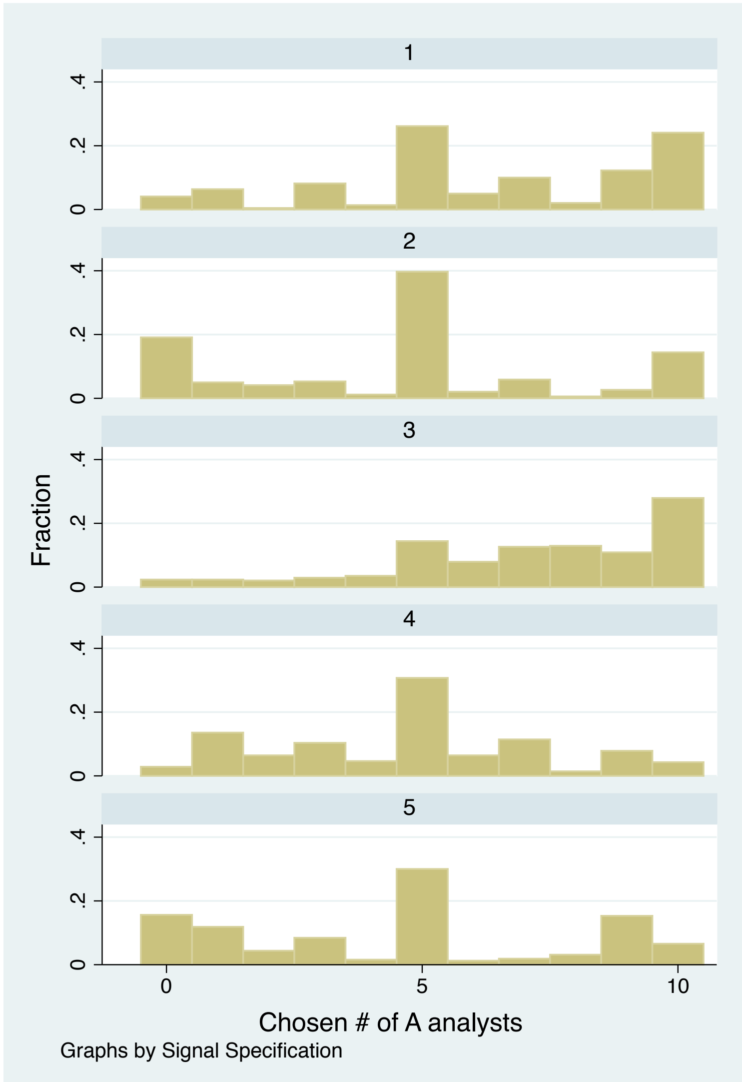
4.3 Choice of Analyst Combination

Finally, we analyze how subjects choose the combination of analysts in the last 10 periods of a session (periods 51 to 60). Table 13 presents the optimal number of group A analysts and the mean and standard deviation of subjects' choice under each specification. We can reject that the mean is equal to the optimal number in all the cases. Nonetheless, the directions of the deviation from optimal choice are consistent with our main result that people overvalue strongly correlated signals and undervalue weakly correlated signals.

Figure 1 presents the histograms of the chosen number of group A under each specification. Under each specification, relatively few subjects chose the optimal number. Moreover, there seem to be a tendency to choose equal number of group A and group B analysts. However, this tendency was not higher for Specifications 1 and 4 where the variances of the two groups would be considered equal if subjects ignored

[26]Separate time series plots for each specifications and regression results without the first 10-period data are available from the authors upon request.

Figure 1: Chosen Number of Group A Analysts, Specifications 1-5



correlation. We also see subjects choosing extreme values of 0 or 10. Under all the specifications, the extreme value closer to the optimal number of group A analysts seems to be more popular, at least nominally. Overall, we are not able to extract any deeper insight about subject behavior or the heuristic they use from their choice of the number of analysts.

5 Conclusion

We present results from a comprehensive set of experiments to ascertain how people treat correlated and independent information and information of different quality levels. We find that while people overvalue strongly correlated information, they undervalue weakly correlated information; such correlation over-adjustment is a new finding. Moreover, people seem to put suboptimally high weight on information that they directly receive than those they indirectly receive through the actions of others. This may be caused by ambiguity regarding the actions of others. Nonetheless, we find that people show some sophistication in processing information; in particular, they do not completely ignore correlation. We provide insights on the directions of bounded rationality in information processing.

Signals in this experiment are generated through a complicated process, which maybe considered a drawback. However, information in real life also follows complicated processes. Our experimental results provide insights for an "as-if" description of people's belief updating behavior. It is important to figure out how people form beliefs in a complex environment; one they may not understand completely. Our results, perhaps with the aid of more experiments, will help us in creating more realistic models of belief formation that allow people to be in between complete correlation neglect and full rationality.

Our results can be useful for many different applications. In particular, it is important to recognize that people's responses to correlated information depend on the degree of correlation. As pointed out in the literature, people can be "fooled" by financial experts' suggestions when the experts have very similar view points by design (shared information source, common analytic tools, similar incentives) even when people are aware of the similarities. But this is not the whole story. They may not appropriately appreciate suggestions by financial experts who try to obtain independent sources of information but can get only weakly correlated sources. Combined with the findings from experiments on sequential actions by subjects designed to find herding behavior, our results on revised prediction may suggest that bank runs or excessive purchase of hot stocks (leading to creation of a bubble) based on market activity maybe tempered if enough people receive contrary private signals. Our results may also suggest that in e-commerce, one may overestimate the impact of click-through rates on eventual purchase when aggregating data from similar types of consumers and underestimate it when aggregating data from dissimilar types of consumers.

Our results also have implications on marketing and public opinion. In the literature on multiple source effect, the same message from multiple media are considered more informative. While we obtain a similar implication under strong correlation and our argument provides a mechanism behind this effect, our results also imply that an opposite effect may appear under weak correlation: An effort to provide less correlated information may not be rewarded sufficiently. For example, political polls that are very weakly correlated may be "trusted" by voters much less than they are supposed to be, making them less useful to political operatives. They may instead prefer to report results from pollsters whose methodologies favor their candidate even if these polls are known to be more or less the same. Another example can be regarding public opinion about climate change. Climate scientists Anderson et al. (2013) provide an independent measure of global warming instead of using thermometer-based global surface temperature time series of 130 years which is widely used but provides correlated assessment. We would speculate that the information provided by these researches may not be appreciated by the public as much as one would expect because people may undervalue the effort to reduce correlation. More research needs to be conducted to examine how well such implications of our findings carry over to real applications.

This paper provides a framework for investigating belief formation under a large set of contexts and settings. One future direction would be to analyze how people form beliefs when information or signals arrive sequentially. One can also explore how people decide how much information to receive based on their beliefs. Another extension would be to allow for biased signals. This can be particularly relevant in the context of politics and media. There is some recent theoretical investigation on how correlation neglect may affect the way media or news outlets present information.[27] Thus, learning how people treat correlation among biased signals and form their beliefs using experiments will be useful. We can also vary the level of biases the consumers of news themselves have in their preferences and investigate how that interacts with potential misperception of correlation. In general, our experimental setup and results can open up a number of different research avenues.

References

[1] Akerlof, George A. and Robert J. Shiller (2009): *Animal Spirits: How Human Psychology Drives the Economy and Why it Matters for Global Capitalism*, Princeton University Press.

[2] Anderson, D. M., E. M. Mauk, E. R. Wahl, C. Morrill, A. J. Wagner, D. Easterling, and T. Rutishauser (2013): "Global Warming in an Independent Record of the Past 130 Years," *Geophysical Research Letters*, 40, 189-193.

[3] Au, Pack Hung and Keiichi Kawai (2012): "Media Capture and Information Monopolization in Japan," *The Japanese Economic Review*, 63(1), 131-147.

[4] Barber, Brad M. and Terrance Odean (2001): "Boys Will be Boys: Gender, Overconfidence, and Common Stock Investment," *Quarterly Journal of Economics*, 116(1), 261-292.

[5] Borghans, Lex, James J. Heckman, Bart H. Golsteyn, and Huub Meijers (2009): "Gender Differences in Risk Aversion and Ambiguity Aversion," *Journal of the European Economic Association*, 7(2-3), 649-658.

[6] Breslow, Norman. (1982): "Design and Analysis of Case Control Studies," *Annual Review of Public Health*, 3, 29–54.

[27] See, for example, Levy, Moreno de Barreda, and Razin (2017).

[7] Chang, Yuhmiin, and Esther Thorson (2004): "Television and Web Advertising Synergies," *Journal of Advertising*, 33(2), 75-84.

[8] Cerreia-Vioglio, Simone, David Dillenberger, and Pietro Ortoleva (2015): "Cautious Expected Utility and the Certainty Effect," *Econometrica*, 83(2), 693-728.

[9] Chen, Joseph S. (2000): "A Comparison of Information Usage Between Business and Leisure Travelers," *Journal of Hospitality & Leisure Marketing*, 7(2), 65-76.

[10] Coffman, Katherine Baldiga (2014): "Evidence on self-stereotyping and the contribution of ideas" *Quarterly Journal of Economics*, 129 (4), 1625-1660.

[11] De Filippis, Roberta, Antonio Guarino, Philippe Jehiel, and Toru Kitagawa (2017): "Updating Ambiguous Beliefs in a Social Learning Experiment," mimeo.

[12] Enke, Benjamin, and Florian Zimmerman (2017): "Correlation Neglect in Belief Formation," forthcoming, *Review of Economic Studies*.

[13] Ellis, Andrew, and Michele Piccione (2017): "Correlation Misperception in Choice, " *American Economic Review*, 107(4), 1264-92.

[14] Eyster, Erik, and Georg Weizsäcker (2016): "Correlation Neglect in Portfolio Choice: Lab Evidence," mimeo.

[15] Eyster, Erik, and Matthew Rabin (2014): "Extensive Imitation is Irrational and Harmful," *Quarterly Journal of Economics*, 129(4), 1861-1898.

[16] Eyster, Erik, Matthew Rabin, and Dimitri Vayanos (2018): "Financial Markets where Traders Neglect the Informational Content of Prices," forthcoming, *Journal of Finance*.

[17] Fischbacher, Urs (2007): "z-Tree - Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics*, 10(2), 171-178.

[18] Friedl, Andreas, Patrick Ring, and Ulrich Schmidt (2017): "Gender Differences in Ambiguity Aversion under Different Outcome Correlation Structures," *Theory and Decision*, 82(2), 211-219.

[19] Gilboa, Itzhak, and David Schmeidler (1989): "Maxmin Expected Utility with Non-unique Prior," *Journal of Mathematical Economics*, 18(2), 141-153.

[20] Harkins, Stephen G., and Richard E. Petty (1981a): "Effects of Source Magnification of Cognitive Efforts on Attitude: An Information-Processing View," *Journal of Personality and Social Psychology*, 52, 401–413.

[21] Harkins, Stephen G., and Richard E. Petty (1981b): "The Multiple Source Effect in Persuasion: The Effects of Distraction," *Personality and Social Psychology Bulletin*, 7(4), 627-635.

[22] Harkins, Stephen G., and Richard E. Petty (1987): "Information Utility and the Multiple Source Effect," *Journal of Personality and Social Psychology*, 52(2), 260–268.

[23] Huang, Rong, Murugappa (Murgie) Krishnan, John Shon, and Ping Zhou (2017): "Who Herds? Who Doesn't? Estimated of Analysts' Herding Propensity in Forecasting Earnings," *Contemporary Accounting Research*, 34(1), 374–399.

[24] Hossain, Tanjim, and Ryo Okui (2013): "The Binarized Scoring Rule," *Review of Economic Studies*, 80(3), 984–1001.

[25] Kallir, Ido, and Doron Sansino (2009): "The Neglect of Correlation in Allocation Decisions," *Southern Economic Journal*, 75(4), 1045-66.

[26] Kroll, Yoram, Haim Levy, and Amnon Rapoport (1988): "Experimental Tests of the Separation Theorem and the Capital Asset Pricing Model," *American Economic Review*, 78(3), 500-519.

[27] Lee, Jae Young, and Raghuram Iyengar (2016): "A Model of Consumer Search for Experience Goods," Working Paper, University of Pennsylvania.

[28] Lee, Kwan Min, and Clifford Nass (2004): "The Multiple Source Effect and Synthesized Speech," *Human Communication Research*, 30(2), 182-207.

[29] Levy, Gilat, and Ronny Razin (2015): "Correlation Neglect, Voting Behavior, and Information Aggregation," *American Economic Review*, 105(4), 1634-45.

[30] Levy, Gilat, and Ronny Razin (2018): "Combining Forecasts in the Presence of Ambiguity over Correlation Structures," mimeo.

[31] Levy, Gilat, Ines Moreno de Barreda, and Ronny Razin (2017): "Persuasion with Correlation Neglect: Media Power via Correlation of News Content," Working Paper, University of Oxford.

[32] Lim, Joon Soo, Sung Yoon Ri, Beth Donnelly Egan, and Frank A. Biocca (2015): "The Cross-platform Synergies of Digital Video Advertising: Implications for Cross-media Campaigns in Television, Internet and Mobile TV," *Computers in Human Behavior*, 48, 463-472.

[33] Luan, Shenghua, Robert D. Sorkin, and Jesse Itzkowitz (2004): "Weighting Information from Outside Sources: A Biased Process," *Journal of Behavioral Decision Making*, 17(2), 95-116.

[34] Maines, Laureen A. (1996): "An Experimental Examination of Subjective Forecast Combination," *International Journal of Forecasting*, 12(2), 223-233.

[35] Moore, David J., Richard Reardon, and William J. Qualls (1989): "Do Multiple Source Appeals Enhance the Confidence Level of Message Recipients?," *Applied Psychology*, 38(3), 237-250.

[36] Nöth, Markus, and Martin Weber (2003): "Information Aggregation with Random Ordering: Cascades and Overconfidence," *The Economic Journal*, 113(484), 166- 189.

[37] Ortoleva, Pietro, and Erik Snowberg (2015): "Overconfidence in Political Behavior," *American Economic Review*, 105(2), 504-535.

[38] Soll, Jack B. (1999): "Intuitive Theories of Information: Beliefs about the Value of Redundancy," *Cognitive Psychology*, 38, 317–346.

[39] Tang, Tang, Gregory D. Newton, and Xiaopeng Wang. (2007): "Does Synergy Work? An Examination of Cross-promotion Effects," *The International Journal on Media Management*, 9(4), 127-134.

[40] Trueman, Brett (1994): "Analyst Forecasts and Herding Behavior," *Review of Financial Studies*, 7(1), 97-124.

[41] Voorveld, Hilde A.M. (2011): "Media Multitasking and the Effectiveness of Combining Online and Radio Advertising," *Computers in Human Behavior*, 27(6), 2200-2206.

A Appendix

A.1 Proof of Proposition 1

Let $X = (X_1^A, \ldots, X_{n_A}^A, X_1^B, \ldots, X_{n_B}^B)^\top$. The likelihood function of X is propositional to

$$(X - T\iota_N)^\top \Sigma^{-1}(X - T\iota_N),$$

where

$$\Sigma = \begin{pmatrix} \sigma_{AI}^2 I_{n_A} + \sigma_{AC}^2 \iota_{n_A}\iota_{n_A}^\top & 0_{n_A \times n_B} \\ 0_{n_B \times n_A} & \sigma_{BI}^2 I_{n_B} + \sigma_{BC}^2 \iota_{n_B}\iota_{n_B}^\top \end{pmatrix},$$

ι_a is the $a \times 1$ vector of ones, I_a is the $a \times a$ identity matrix and $0_{a\times b}$ is the $a \times b$ matrix of zeros.

Because the prior is $T \sim N(\mu_p, \sigma_p^2)$, the posterior is

$$T|X \sim N(\mu^*, (\sigma^*)^2),$$

where

$$\mu^* = \left(\frac{1}{\sigma_p^2} + \iota_N^\top \Sigma^{-1}\iota_N\right)^{-1} \left(\iota_N^\top \Sigma^{-1} X + \frac{\mu_p}{\sigma_p^2}\right),$$
$$(\sigma^*)^2 = \left(\frac{1}{\sigma_p^2} + \iota_N^\top \Sigma^{-1}\iota_N\right)^{-1}.$$

By the rule $(I + AA^\top)^{-1} = I - A(I + A^\top A)^{-1}A^\top$ for an identity matrix I and a matrix A, we have

$$\Sigma^{-1} = \begin{pmatrix} \frac{1}{\sigma_{AI}^2} I_{n_A} - \frac{\sigma_{AC}^2}{\sigma_{AI}^2}\frac{1}{\sigma_{AI}^2 + \sigma_{AC}^2 n_A}\iota_{n_A}\iota_{n_A}^\top & 0_{n_A \times n_B} \\ 0_{n_B \times n_A} & \frac{1}{\sigma_{BI}^2} I_{n_B} - \frac{\sigma_{BC}^2}{\sigma_{BI}^2}\frac{1}{\sigma_{BI}^2 + \sigma_{BC}^2 n_B}\iota_{n_B}\iota_{n_B}^\top \end{pmatrix}.$$

It therefore follows that

$$\begin{aligned}
\iota_N^\top \Sigma^{-1}\iota_N =& \frac{n_A}{\sigma_{AI}^2} - \frac{\sigma_{AC}^2}{\sigma_{AI}^2}\frac{n_A^2}{\sigma_{AI}^2 + \sigma_{AC}^2 n_A} + \frac{n_B}{\sigma_{BI}^2} - \frac{\sigma_{BC}^2}{\sigma_{BI}^2}\frac{n_B^2}{\sigma_{BI}^2 + \sigma_{BC}^2 n_B} \\
=& \frac{n_A}{\sigma_{AI}^2 + \sigma_{AC}^2 n_A} + \frac{n_B}{\sigma_{BI}^2 + \sigma_{BC}^2 n_B}, \\
\iota_N^\top \Sigma^{-1} X =& \left(\frac{1}{\sigma_{AI}^2} - \frac{\sigma_{AC}^2}{\sigma_{AI}^2}\frac{n_A}{\sigma_{AI}^2 + \sigma_{AC}^2 n_A}\right)\sum_{j=1}^{n_A} X_j^A + \left(\frac{n_B}{\sigma_{BI}^2} - \frac{\sigma_{BC}^2}{\sigma_{BI}^2}\frac{n_B^2}{\sigma_{BI}^2 + \sigma_{BC}^2 n_B}\right)\sum_{j=1}^{n_B} X_j^B \\
=& \frac{n_A}{\sigma_{AI}^2 + \sigma_{AC}^2 n_A}\bar{X}_A + \frac{n_B}{\sigma_{BI}^2 + \sigma_{BC}^2 n_B}\bar{X}_B.
\end{aligned}$$

Therefore, when $\sigma_p \to \infty$ (uninformative prior), the posterior of T is a normal distribution with mean:

$$\left(\frac{n_A}{\sigma_{AI}^2+\sigma_{AC}^2 n_A}+\frac{n_B}{\sigma_{BI}^2+\sigma_{BC}^2 n_B}\right)^{-1}\left(\frac{n_A}{\sigma_{AI}^2+\sigma_{AC}^2 n_A}\bar{X}_A+\frac{n_B}{\sigma_{BI}^2+\sigma_{BC}^2 n_B}\bar{X}_B\right)$$

and variance:

$$\left(\frac{n_A}{\sigma_{AI}^2+\sigma_{AC}^2 n_A}+\frac{n_B}{\sigma_{BI}^2+\sigma_{BC}^2 n_B}\right)^{-1}.$$

The posterior mean can be written as

$$\begin{aligned}&\left(\frac{n_A}{\sigma_{AI}^2+\sigma_{AC}^2 n_A}+\frac{n_B}{\sigma_{BI}^2+\sigma_{BC}^2 n_B}\right)^{-1}\left(\frac{n_A}{\sigma_{AI}^2+\sigma_{AC}^2 n_A}\bar{X}_A+\frac{n_B}{\sigma_{BI}^2+\sigma_{BC}^2 n_B}\bar{X}_B\right)\\=&\frac{n_A(n_B\sigma_{BC}^2+\sigma_{BI}^2)}{n_A(n_B\sigma_{BC}^2+\sigma_{BI}^2)+n_B(n_A\sigma_{AC}^2+\sigma_{AI}^2)}\bar{X}_A\\&+\frac{n_B(n_A\sigma_{AC}^2+\sigma_{AI}^2)}{n_A(n_B\sigma_{BC}^2+\sigma_{BI}^2)+n_B(n_A\sigma_{AC}^2+\sigma_{AI}^2)}\bar{X}_B.\end{aligned}$$

A.2 Proof of Proposition 2

The proof of Proposition 1 demonstrates that the posterior variance under $\sigma_p^2 \to \infty$ is

$$\begin{aligned}&\left(\frac{n_A}{\sigma_{AI}^2+\sigma_{AC}^2 n_A}+\frac{n_B}{\sigma_{BI}^2+\sigma_{BC}^2 n_B}\right)^{-1}\\=&\left(\frac{n_A}{\sigma_{AI}^2+\sigma_{AC}^2 n_A}+\frac{N-n_A}{\sigma_{BI}^2+\sigma_{BC}^2 (N-n_A)}\right)^{-1}.\end{aligned}$$

Notice that our minimization problem is equivalent to maximizing the inverse of the variance. Thus we consider maximizing

$$\frac{n_A}{\sigma_{AI}^2+\sigma_{AC}^2 n_A}+\frac{N-n_A}{\sigma_{BI}^2+\sigma_{BC}^2 (N-n_A)}.$$

Note that this function is concave in n_A and the first order condition is necessary and sufficient if the solution is in $(0, N)$. The first order condition is

$$\frac{1}{\sigma_{AI}^2+\sigma_{AC}^2 n_A}-\frac{n_A}{(\sigma_{AI}^2+\sigma_{AC}^2 n_A)^2}\sigma_{AC}^2-\frac{1}{\sigma_{BI}^2+\sigma_{BC}^2 (N-n_A)}+\frac{N-n_A}{(\sigma_{BI}^2+\sigma_{BC}^2 (N-n_A))^2}\sigma_{BC}^2=0.$$

Rearranging the terms, we obtain

$$-\frac{\sigma_{AI}^2}{(\sigma_{AI}^2+\sigma_{AC}^2 n_A)^2}+\frac{\sigma_{BI}^2}{(\sigma_{BI}^2+\sigma_{BC}^2 (N-n_A))^2}=0.$$

Solving this equation, the optimal value of n_A is

$$n_A = \frac{\sigma_{AI}\sigma_{BI}^2 + \sigma_{AI}\sigma_{BC}^2 N - \sigma_{BI}\sigma_{AI}^2}{\sigma_{AI}\sigma_{BC}^2 + \sigma_{BI}\sigma_{AC}^2}.$$

A.3 Proof of Proposition 3

We note that this problem can be considers as a Bayesian problem in which the prior is the subject's belief (which is posterior obtained in the proof of Proposition 1 if the subject is truly Bayesian) and the observation is the average of predictions of all other subjects. Let $\bar{p}_- = (M\bar{p} - p^*)/(M-1)$ be the average of all other subjects' predictions. The assumption of the proposition indicates that this subject's belief is $T \sim N(p^*, \sigma^2)$, and $\bar{p}_- \sim N(T, \sigma^2/(M-1))$.

With observation $\bar{p}_-$, the posterior is updated to

$$N\left(\frac{\sigma^2}{\sigma^2 + \sigma^2/(M-1)}\bar{p}_- + \frac{\sigma^2/(M-1)}{\sigma^2 + \sigma^2/(M-1)}p^*, \frac{\sigma^2\sigma^2/(M-1)}{\sigma^2 + \sigma^2/(M-1)}\right).$$

Thus, the posterior mean is

$$\frac{\sigma^2}{\sigma^2 + \sigma^2/(M-1)}\bar{p}_- + \frac{\sigma^2/(M-1)}{\sigma^2 + \sigma^2/(M-1)}p^* = \bar{p}.$$

Next, we consider the case in which a subject believes that other subjects' estimates have variances of $\lambda\sigma^2$. In this case, $\bar{p}_- \sim N(T, \lambda\sigma^2/(M-1))$, and the posterior becomes

$$N\left(\frac{\sigma^2}{\sigma^2 + \lambda\sigma^2/(M-1)}\bar{p}_- + \frac{\lambda\sigma^2/(M-1)}{\sigma^2 + \lambda\sigma^2/(M-1)}p^*, \frac{\sigma^2\lambda\sigma^2/(M-1)}{\sigma^2 + \lambda\sigma^2/(M-1)}\right).$$

The posterior mean is

$$\frac{\sigma^2}{\sigma^2 + \lambda\sigma^2/(M-1)}\bar{p}_- + \frac{\lambda\sigma^2/(M-1)}{\sigma^2 + \lambda\sigma^2/(M-1)}p^* = \frac{M}{M-1+\lambda}\bar{p} + \frac{\lambda-1}{M-1+\lambda}p^*.$$

A.4 Additional Figures

Figure 2: Initial Prediction, Time Series of Weights on A, Specification 1

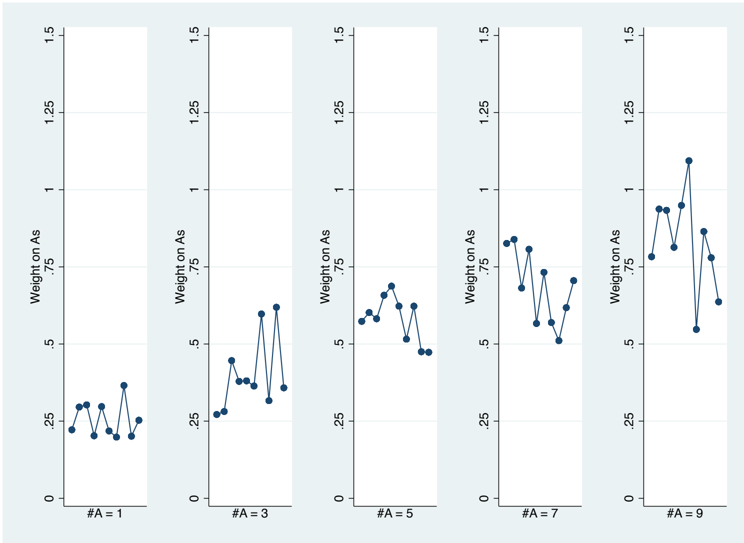


Figure 3: Initial Prediction, Time Series of Weights on A, Specification 2

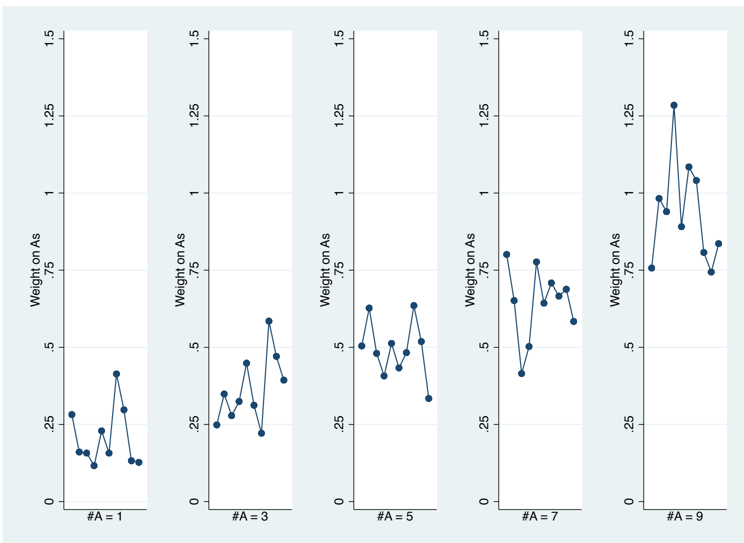


Figure 4: Initial Prediction, Time Series of Weights on A, Specification 3

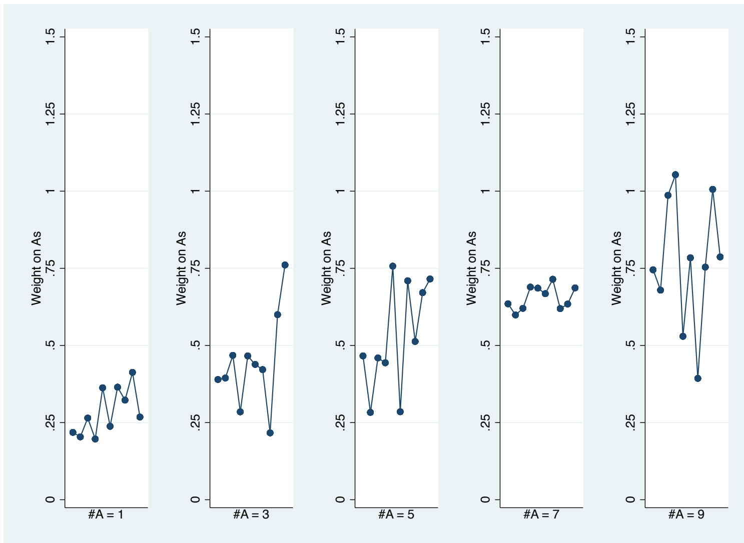


Figure 5: Initial Prediction, Time Series of Weights on A, Specification 4

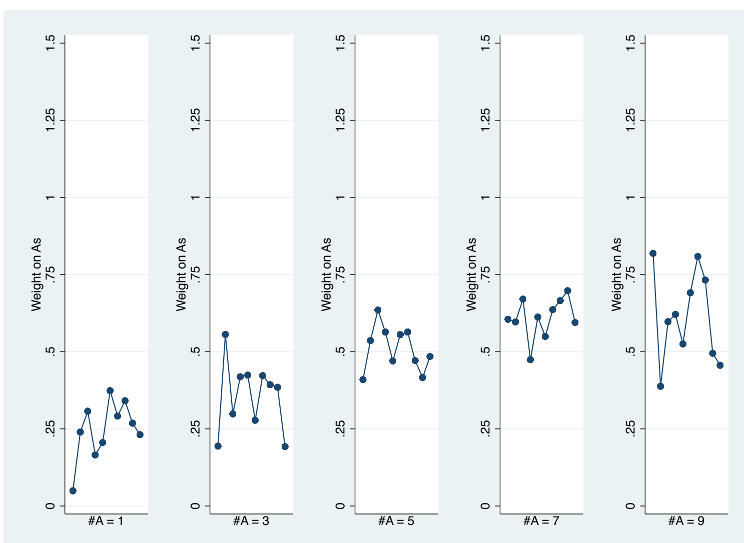


Figure 6: Initial Prediction, Time Series of Weights on A, Specification 5

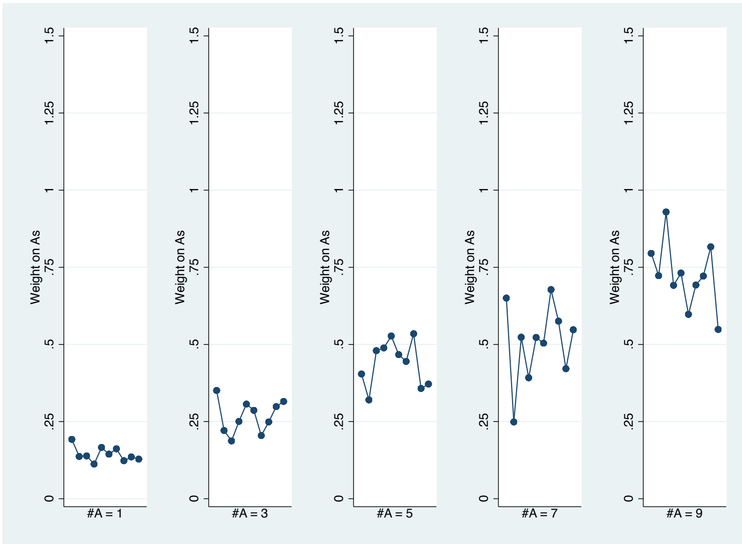
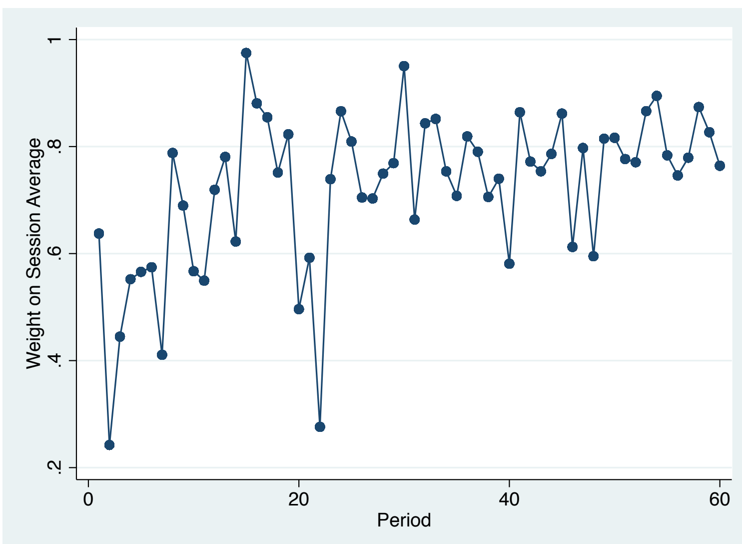


Figure 7: Revised Prediction, Time Series of Weights on Session Average, Specifications 1–5



A.5 Sample Experimental Instructions and Statistical Definitions

Sample Experimental Instructions for Specification 1

General Rules

This session is part of an experiment about how people aggregate multiple forecasts about returns from a financial asset to estimate the average return. If you follow the instructions carefully and make good decisions, you can earn a considerable amount of money. There are __ people (including you) in this room who are also participating as subjects in this session. They have all been recruited in the same way as you and are reading the same instructions as you are for the first time. It is important that you do not communicate with any other participant or discuss the details of the experiment with anyone during and after the session.

Setting

This session will consist of 60 periods. In each period, you will receive forecasts from 10 stock analysts regarding the earning per share (EPS) for the stocks of a company. In every period, the information will be provided for a new company. Thus, the EPS in each period is independent of each other. In each period, your goal is to predict the true value of the EPS for that period's stock based on the information you receive.

For each stock, you will receive 10 forecasts about the EPS. An analyst cannot perfectly forecast the true EPS of a stock. Rather, she observes the EPS along with two error terms— Division 1 error and Division 2 error. Specifically, her observation equals the true EPS plus division 1 error and Division 2 error where any error term can be positive or negative. She reports this observation as her forecast.

In each period, some of the 10 forecasts will be from *Group A* analysts and the rest will be from *Group B* analysts. For all *Group A* analysts, the errors from both divisions are independent of each other. Suppose, in some period, the true EPS for that period's stock is *T*. If analyst *i* is a *Group A* analyst, then she observes forecast F_i where $F_i = T + \varepsilon_{1i} + \varepsilon_{2i}$. In this session, Division 1 errors for *Group A* analysts are drawn from a normal distribution with mean 0 and variance of 250 and Division 2 error for *Group A* analysts are drawn from a normal distribution with mean 0 and variance of 250. On the other hand, all *Group B* analysts have the same Division 1 error term but different Division 2 error terms. That is, if analyst *j* is a *Group B* analyst, then she observes forecast F_j where $F_j = T + \varepsilon_1 + \varepsilon_{2j}$. The error ε_1 is common for all *Group B* analysts. But, error ε_{2j} is different for each *Group B* analyst. Hence, forecast errors from *Group B* analysts are correlated but not the same. Note that, in this session, the Division 1 error for *Group B* analysts are drawn from a normal distribution with mean 0 and variance of 250 and Division 2 errors for *Group B* analysts are drawn from a normal distribution with mean 0 and variance of 250.

Description of a Period

You will receive 10 forecasts regarding the EPS for the stock. The forecasts will be generated according to the system described above. Observing the forecasts, you will be asked to enter your prediction for the true EPS (see Figure 1). Let us refer to your entered prediction as *P*. Once you have entered *P*, we will calculate the squared loss, which is defined as $(P - T)^2$ where T is the true EPS. We will also independently draw a number *K* randomly from a Uniform distribution on $[0,40]$. If *K* is larger than or equal to the squared loss, you will receive 100 points. If *K* is smaller than the squared loss, you will receive no point. Thus, it is optimal for you to report what you think the true EPS is, on average, based on the 10 forecasts you receive as your prediction.

In a given period, all other participants in the room also receive 10 forecasts for the same stock as you do. However, each of them observe a completely different set of forecasts. All the participants will submit a prediction for the true EPS based on the forecasts they receive as explained above. Once all of you have reported your predictions. We will report to everyone the average of predictions from all the participants in the room (including yours). Then we will ask you to submit a revised prediction (see Figure 2). We refer to this revised prediction as P^r. After you have entered P^r, we will calculate the squared loss for this entry $(P^r - T)^2$. Note that, P^r can be same as or different from your original prediction P. We will also independently draw a number K^r randomly from a Uniform distribution on $[0,8]$. If K^r is larger than or equal to $(P^r - T)^2$, you will receive 50 points. If K^r is smaller than $(P^r - T)^2$, you will receive no point. Thus, it is optimal for you to report what you think the true EPS is, on average, based on the 10 forecasts and the average of all participants' predictions as your revised prediction.

Differences between Periods

There will be 60 periods in this session. In each period, we will consider a different company. The EPS for the stock of the company in each period will be independently drawn from a distribution with a very large variance. Thus, the true EPS across different periods are independent of each other and the EPS of the stock for one company does not provide information about the EPS of the stock for another company across different periods.

Recall that, in each period, some of the 10 forecasts will be from *Group A* analyst and the rest will be from *Group B* analysts. In periods 1 to 10, there will be 1 *Group A* and 9 *Group B* analysts. In periods 11 to 20, the number of *Group A* and *Group B* analysts will be 3 and 7, respectively. In periods 21 to 30, these numbers will be 5 and 5, respectively, and in periods 31 to 40, these numbers will be 7 and 3, respectively. In periods 41 to 50, the number of *Group A* and *Group B* analysts will be 9 and 1, respectively. That is, in the first 50 periods, all participants will have the same combination of *Group A* and *Group B* analysts, but the combination will change (for everyone) every 10 periods. However, in periods 51 to 60, every participant can choose the number of *Group A* and *Group B* forecasts they want to receive (see Figure 3). For example, if in period 51, you choose to have 6 *Group A* analysts, then you will receive 6 *Group A* and 4 *Group B* forecasts. You will be able to make this choice in each of these 10 periods. This also implies that, the combination of the two groups of analysts can be different for different participants in these periods.

In all 60 of the periods, first you will enter your prediction for the EPS of the stock in that period. Then you will receive the average of every participants' predictions and will be asked to enter a revised prediction for the EPS. After entering the prediction and then the revised prediction, you will be informed of the true EPS, the squared loss for both the prediction and revised prediction, the relevant random numbers (K and K^r) drawn from uniform distributions and your earnings (in points) from the prediction and the revised prediction.

Ending the Session

At the end of period 60, you will see a screen displaying your point earnings from each period. You will receive €10 for participating in this experiment. On top of that, you will earn an amount based on your point earnings from six randomly chosen periods, one from each 10-period block. Points earned in these six periods will be converted to money at the rate of €1 for 30 points. That is, if you earn y points in total in these six periods, your total income from the experiment will be €(10 + y/30). You will be paid this amount in cash at the end of the session.

Some Statistical Concepts

Mean (Average) is calculated by summing the observed numerical values of a variable in a set of data and then dividing the total by the number of observations involved. If we have data set (or sample) with n data points and the data points are x_1, x_2,…, x_n then the sample average, $\bar{x} = \frac{x_1+x_2+\cdots+x_n}{n}$.

Variance is the average of the squared differences between each of the observations in a set of data and the mean. Variance is used to indicate how possible values are spread around the mean. Then the sample variance, $s^2 = \frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\cdots+(x_n-\bar{x})^2}{n}$.

Example: Suppose our data set has 5 data points — $x_1 = 1$, $x_2 = 5$, $x_3 = 7$, $x_4 = 3$, $x_5 = 12$

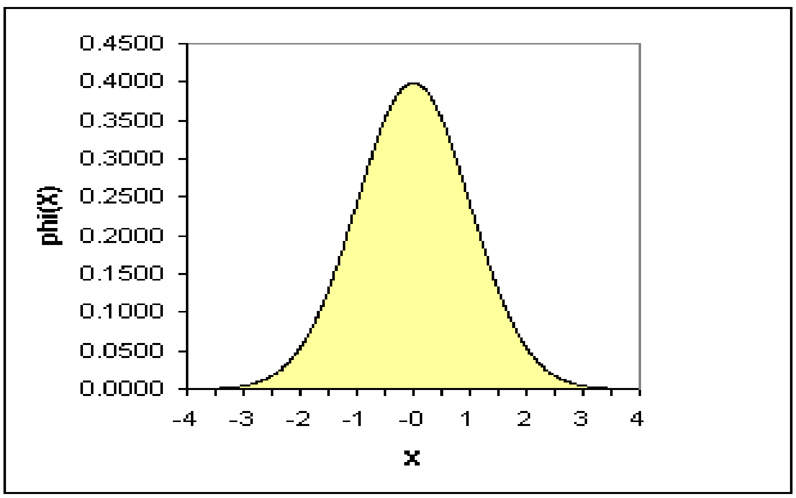
Mean: $\bar{x} = \frac{x_1+x_2+x_3+x_4+x_5}{5}$ = (1 + 5 + 7 + 3 + 12)/5 = 5.6

Variance: $s^2 = [(1 – 5.6)^2 + (5 – 5.6)^2 + (7 – 5.6)^2 + (3 – 5.6)^2 + (12 – 5.6)^2]/5 = 14.24$

Distributions

If a variable is drawn from a **Uniform Distribution** on the interval [a, b] that means that all points on [a, b] are equally likely to be drawn.

Normal Distribution is pattern for the distribution of a set of data which follows a bell shaped curve. The following example is a normal distribution with mean of 0 and variance of 1. We refer to this distribution as the standard normal distribution. A picture appears below. The data points (x) can take values from negative to positive infinity and the probability density at point x is given by $\Phi(x)$.



The probability distribution of a normal distribution with mean M and variance S^2 is as follows. That is, if the random number x is drawn from such a distribution then:

Prob($x \leq M$ – 1.28160 S) = 0.1	Prob($x \leq M$ – 0.84162 S) = 0.2	Prob($x \leq M$ – 0.52440 S) = 0.3
Prob($x \leq M$ – 0.25335 S) = 0.4	Prob($x \leq M$) = 0.5	Prob($x \leq M$ + 0.25335 S) = 0.6
Prob($x \leq M$ + 0.52440 S) = 0.7	Prob($x \leq M$ + 0.84162 S) = 0.8	Prob($x \leq M$ + 1.28160 S) = 0.9

Independent draws from the same distribution are those draws selected from the same distribution which have no effect on one another. That is, no correlation exists between the draws.