Web Supplements to

"Informational Overconfidence in Return Prediction - More Properties"

Web Supplement I: Translated Instructions (Experiment I)

Purchase and Sale Experiment on the Tel-Aviv Exchange

Hello *{participant name as received in preregistration}*

The experiment is run by researchers at the College of Management Business School in order to test purchase and sale patterns of stock traded at the Tel-Aviv Exchange.

The questionnaire is user friendly. You are requested to fill-in the form and send to our email address at exp.stock@gmail.com

The organizers would confirm the eligibility of your completed form within 4 days after submission.

Along the experiment, you would be asked to assume the role of an investment consultant of hypothetical investors that either plan to acquire or already won TASE stocks in the amounts of 100,000-200,000 NIS.

The questionnaire is composed of 5 BUY problems and 5 SELL problems.

In each BUY problem, you would be requested to select a stock for purchase, in amount of 100,000 NIS, from a list of two candidate stocks.

In each SELL problem, you would be requested to select a stock for sale, in amount of 100,000 NIS, from a list of two candidate stocks.

In addition, you would also be asked to provide an assessment regarding the range of expected returns on each of the stocks that you have selected to purchase or sale

Your selections would be tested by the relative performance of the 2 candidate stocks, in the 3 months following the delivery of your completed questionnaire. Please note, the experiment deals with stock selections for short horizon of only 3 calendar months!

By the end of the test-period (3 months after the delivery of your completed questionnaire), the experiment organizers would randomly select one of the 10 assignments and pay you a bonus depending on the relative performance of the two candidate stocks from which you were requested to choose in the respective assignment (more details would be provided in the next page)

While contemplating your choices you may use all possible information sources, including historical return data, financial statements, analysts' recommendations and more. A partial list of Web based resources would be provided – with accessible links – at the sequel.

Your participation in the experiment is private and we would not ask for details regarding your employee and exact placement.

The organizers guarantee your confidentiality and commit not to use the data collected along the experiment for any others purpose except for general analysis of the results for a research project! The results of the experiment would be distributed anonymously (using ids to identify participants) at the end of the experiment – to let you control our experimental method and the payment procedure.

_____

Part I: Preliminary Questionnaire

Before presenting the main assignments, we ask you to fill in the following questionnaire

Please provide all the requested information. Incomplete questionnaire would disqualify your questionnaire and you records would be removed from the database of the experiment!

Personal details:

Private name: _____

Family name: _____

Id:_____

Email:_____

Cellular:_____

Age:_____

Marital status: _____

{pull down options: married no children/married with children/single/divorced/other}

Gender: male/female

Please characterize your economic condition:_____

{pull down options: excellent/above average/about average/below average/poor}

Do you hold or are you an MBA student?_____

Do you have a bachelor degree in finance-related disciplines (economic, business, accounting, insurance etc)? _____ *{yes/no}*

Years of formal education (including current year where applicable): _____

{*12 for high school education, 13,14…}*

Current occupation: _____

{pull down options: professional investment industry occupation, other finance-related or money management occupation, other occupation not directly related to finance, unemployed}

In accordance with your choice of occupational category above, please describe your specific occupation: _____

Years of experience within the investment industry (where you held investments related professional jobs such as investment management, consultation, analysis, brokerage etc):

_____

Page Break

Please rank your theoretical knowledge in finance (from academic courses, professions literature, Internet resources and others)

Please check the appropriate box

1	2	3	4	5	6	7

Minimal knowledge **real expert**

Please rank your familiarity with the Israeli stock market (indices, returns, firms etc)

Please check the appropriate box

1	2	3	4	5	6	7

Minimal familiarity **expert familiarity**

Please rank your fittingness, in terms of professional and personal qualifications, to act as an investment consultant

Please check the appropriate box

1	2	3	4	5	6	7

Unfit **Perfect fit**

Please rank your level of consent with each of the following statements:

A. **The purchase of stocks is similar to lottery participation. Selecting a winning stock is a matter of luck**

B. **I am frequently able to point at stocks that would beat the TA100 index**

C. **I rarely recommend or purchase disappointing stocks**

D. **The capital market is rather efficient, it is very difficult to "beat the market"**

{ pull down menu options for each statement: strongly disagree, tend to disagree, indifferent, tend to agree, absolutely agree}

To conclude this part of the questionnaire, please give a prediction for the change in TA100 along your test period: What is your estimate for the expected change in TA100 index in the next 3 months? _____

Page break

Individual Risk Preferences

In each of the following problems please select between alternative A (a lottery) and alternative B (a certain payoff) by checking the respective box. Your choice should reflect your individual preferences. Please provide your candid choices!

Choice Problem 1

□ Alternative A

Probability	Payoff
50%	+200
50%	+50

□ Alternative B

Probability	Payoff
100%	+165

Choice Problem 2

□ Alternative A

Probability	Payoff
50%	+200
50%	+50

□ Alternative B

Probability	Payoff
100%	+125

Choice Problem 3

□ Alternative A

Probability	Payoff
50%	+200
50%	+50

□ Alternative B

Probability	Payoff
100%	+85

Choice Problem 4

□ Alternative A

Probability	Payoff
50%	+200
50%	-50

□ Alternative B

Probability	Payoff
100%	+105

Choice Problem 5

□ Alternative A

Probability	Payoff
50%	+200
50%	-50

□ Alternative B

Probability	Payoff
100%	+75

Choice Problem 6

□ Alternative A

Probability	Payoff
50%	+200
50%	-50

□ Alternative B

Probability	Payoff
100%	+45

Page break

Part II: General Explanation About the Experimental Tasks

The experiment is designed to test the ability of investments professionals and potential investors that are highly familiar with the Israeli stock market, to select stocks for purchase or sale from pairs of candidate stocks

The stocks composing each selection assignment were randomly selected from the stocks included at the TA25 index at the beginning of September 2010. These are the largest stocks in the Israeli market to that date.

Purchase Assignments

In each purchase assignment, you would be asked to selected 1 of 2 possible stocks

You are requested to assume the role of an investment consultant advising a client plaaning to invest 100,000 NIS in one of the two candidate stocks. The client is asking you to pick the stock that is expected, by your opinion, to earn higher return in the next 3 months (from the date where your deliver the completed questionnaire)

Sell Assignments

In each sell assignment, you would be asked to selected 1 of 2 possible stocks

You are requested to assume the role of an investment consultant advising a client that holds 100,000 NIS of each stock and would like to liquidate one of the two investments. The client is asking you to pick the stock that is expected, by your opinion, to show lower return in the next 3 months (from the date where your deliver the completed questionnaire)

Confidence Interval for the Return on the Stock

In parallel to selecting the better stock for purchase or sale you would be asked to provide a 90% confidence interval for the return on the stock that you selected for purchase or sale, in the 3-months test period.

The confidence interval would be delivered in 2 steps: First, you would be asked to provide a lower limit such that with probability 95%, by your assessment, the return on the stock would exceed this limit. Secondly, you would be asked to provide an upper limit such that with probability 95%, by your assessment, the return on the stock would be lower than this limit. The confidence bounds would be elicited using the following format:

Lower limit for the percentile return

In my opinion, with probability 95%, the return on the stock that I have selected would be higher than ___%

Upper limit for the percentile return

In my opinion, with probability 95%, the return on the stock that I have selected would be lower than ___%

Page break

Part III: Method of Payment

Participants that deliver eligible complete questionnaires may win up to 180 NIS for their participation

Your final check would be determined by your performance in the consultation tasks and is composed of 2 components:

A bonus for Successful Consultation

To determine your eligibility for consultation bonus we would randomly select 1 of the 10 consultation problems composing the questionnaire (henceforth: "**the selected assignment**")

Clearly, the selected assignment can either be a purchase task or a sell task with equal chances

To determine if your are entitled to the consultation bonus we would check if your selection was "correct"; i.e., if you have selected the stock that earned higher return in the 3 months test period, when the selected assignment is a purchase task, or if you have selected the stock that earned lower return in the 3 months test period, when the selected assignment is a sell task.

If your choice in the selected assignment turns out correct, you will receive a successful consultation bonus of 100 NIS. If your choice turns out incorrect, you would not receive a bonus for successful consultation.

Bonus for fair and accurate predictions

The prediction of returns by confidence intervals is considered a relatively difficult experimental task. Consider, for example, the prediction of annual 2011 returns for some stock Z (one of the stocks composing the TA25 index). If we select a very low lower-limit (say, a negative return of -90%) and a very high higher-limit (say, positive return of 150%), then our prediction would probably turn out correct, as the annual return on Z would be higher than -90% and lower than 150%. In this case, however, our prediction is imprecise and we also did not open reasonable space for 10% error (5% on each side) – as demanded in 90% confidence forecasting!

Alternatively, if we select very close lower and upper limits (say, a lower limit of +10% and an upper limit of +12%) then our prediction appears accurate, but the statement "With probability 90%, the return on stock Z would fall between 10% and 12%" seems utterly unreasonable (the range of possible returns is too tight – unless the decision maker as an economic model or other information that supports this strong assertion)

90% confidence intervals should therefore be balanced – not too narrow but not too wide. The probability of error should be 5% in each direction.

To incentivize participants to submit reasonable and accurate intervals we would pay a special bonus for fair and accurate prediction!

By the end of the experiment, the organizers would examine the intervals that you have submitted along the experiment to test their reasonability and accuracy. Participants that would meet 3 predetermined conditions would earn an additional bonus of 40-80 NIS.

The exact eligibility conditions for the additional bonus are kept confidential at this stage, but would be provided upon request at the end of the experiment!

Page break

Part IV: Link to Databases

The following links may prove useful in search for information on TA25 stocks and general market conditions:

1. http://www.globes.co.il/
2. http://www.bizportal.co.il/
3. http://www.tase.co.il/tase/
4. http://maya.tase.co.il/bursa/indeximptoday.htm
5. http://www.themarker.com

Page break

{The 10 consultation problems followed next. Each problem was presented in a separate page. About 50% of the questionnaires started with the 5 purchase problems and moved on to the 5 sell problems. Order was reversed for the other questionnaires. We illustrate the structure of the problems with the following sell assignment. All problems were presented in similar format}

Stock Sale – Consulting Problem 1

Which of the next two stocks would show lower return in the next 3-months (from the date where the questionnaire is delivered)?

Please recall that you are acting as a consultant for an investors holding 100,000 NIS of each stock.

Correct selection may pay a bonus of 100 NIS for successful consultation!

□ **OSEM** stock id 304014

□ **BAZAN** stock id 2590248

PLEASE MARK THE STOCK THAT YOU RECOMMEND SELLING (ONE STOCK ONLY)

Please provide a 90% confidence-interval for the 3-months return (starting at the questionnaire delivery date) on the stock selected for sale. For this purpose, provide an upper limit and a lower limit for the 3-months percentile return on the corresponding stock as follows:

Lower limit on return:

With probability 95%, the return on the stock that I selected to sell will be higher than ___%

Upper limit on return:

With probability 95%, the return on the stock that I selected to sell will be lower than ___%

*please consider your bounds carefully. Recall that accurate, fair prediction could pay an additional bonus of 80 NIS

Page break

To conclude the experiment we ask you to answer 2 additional short questions:

1. Please estimate your success in the assignments composing the experiment

a. Please estimate the number of stock purchase problems where your choices would turn-out correct? ____ *{pull-down menu 0,1,2,3,4,5}*

b. Please estimate the number of stock sale problems where your choices would turn-out correct? ____ *{pull-down menu 0,1,2,3,4,5}*

c. Please estimate the number of problems where the return on the stock that you selected for purchase would fall within your interval prediction ____ *{pull-down menu 0,1,2,3,4,5}*

d. Please estimate the number of problems where the return on the stock that you selected for sale would fall within your interval prediction ____ *{pull-down menu 0,1,2,3,4,5}*

2. Please characterize your stock selection method along the experiment

a. Please select the method that best describes your selection criteria in the 5 stock purchase problems

b. Please select the method that best describes your selection criteria in the 5 stock sale problems

{pull-down menu: technical analysis, economic analysis, private information, intuition, guess}

We thank you for participating in the stock selection experiment!

The organizers would follow the performance of TA25 stocks and at the end of your 3 months test-period we would check your selections and predictions and calculate your payout check as explained above

Checks would be sent to the address that you provide below after phone verification of personal details

Checks would be crossed and designated to the name that you provided at the beginning of this questionnaire

Please fill-in your exact address for check delivery:

_____

_____

_____

Thanks again for your time and effort!

_____

More details regarding the undisclosed bonus for fair and accurate predictions

Our undisclosed criteria demanded an overlap of 75% between the 2 smallest intervals in each condition (BUY and SELL) plus 9 hits across the 10 prediction intervals (perfect calibration). The overlap condition was meant to identify cases where subjects deliberately disconnect one relatively small interval from other large intervals strategically attempting to achieve the 90% hit rate. The BUY problems were separated from the SELL problems for the overlap test, to account for the possibility that predictions would be strongly affected by condition. In fact, only 2 participants were perfectly calibrated. These subjects met the overlap conditions as well and accordingly received the 80 NIS bonus.

Web Supplement II: Cross Section Probit Regressions on CCRs

To characterize the variables that significantly affected performance in the choice assignments across the sample we run regressions with model selections and test the robustness of conclusions by modifying the estimation method, the list of variables that provide the basis for the estimations and direct comparisons. The final results of a Probit analysis suggest that CCRs increased with subjective prior CCR estimates (CHOICE_CONF) while decreasing with the DIFFICULTY of the questionnaires. An indicator for successful prediction of the change in TA100 along the 3 months test period (RAT; see the discussion in the paper for more details) also shows positive significant effect. The output of an illustrative estimation is provided below:

Discrete Response Profile of CORRECT_CHOICE_RATE			
Index	**Value**	**Frequency**	**Percent**
1	0.1	1	1.08
2	0.2	5	5.38
3	0.3	18	19.35
4	0.4	23	24.73
5	0.5	12	12.90
6	0.6	18	19.35
7	0.7	11	11.83
8	0.8	4	4.30
9	0.9	1	1.08

Model Fit Summary	
Number of Endogenous Variables	1
Endogenous Variable	CORRECT_CHOICE_RATE
Number of Observations	93
Log Likelihood	-173.26632
Maximum Absolute Gradient	1.09861E-6
Number of Iterations	25
Optimization Method	Quasi-Newton
AIC	366.53265
Schwarz Criterion	391.85864

Goodness-of-Fit Measures		
Measure	**Value**	**Formula**
Likelihood Ratio (R)	4.6152	2 * (LogL - LogL0)
Upper Bound of R (U)	351.15	- 2 * LogL0
Aldrich-Nelson	0.0473	R / (R+N)
Cragg-Uhler 1	0.0484	1 - exp(-R/N)
Cragg-Uhler 2	0.0495	(1-exp(-R/N)) / (1-exp(-U/N))
Estrella	0.0487	1 - (1-R/U)^(U/N)
Adjusted Estrella	-0.176	1 - ((LogL-K)/LogL0)^(-2/N*LogL0)
McFadden's LRI	0.0131	R / U
Veall-Zimmermann	0.0598	(R * (U+N)) / (U * (R+N))
McKelvey-Zavoina	0.1647	
N = # of observations, K = # of regressors		

Algorithm converged.

Parameter Estimates					
Parameter	**DF**	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > |t|**
RAT	1	0.674111	0.284192	2.37	0.0177
CHOICE_CONF	1	0.141299	0.048572	2.91	0.0036
DIFFICULTY	1	0.126176	0.035493	3.55	0.0004
_Limit2	1	0.608685	0.252447	2.41	0.0159
_Limit3	1	1.476381	0.292043	5.06	<.0001
_Limit4	1	2.141622	0.301796	7.10	<.0001
_Limit5	1	2.478113	0.306981	8.07	<.0001
_Limit6	1	3.124195	0.325488	9.60	<.0001
_Limit7	1	3.875862	0.377192	10.28	<.0001
_Limit8	1	4.641586	0.506192	9.17	<.0001

Web Supplement III: Cross Section Probit Regressions on HIT RATEs

To characterize the variables that significantly affected performance in the prediction assignments across the sample we run regressions with model selections and test the robustness of conclusion by modifying the method, the list of variables that provide the basis for the estimations and direct comparisons. The final results of a Probit analysis suggest that HIT RATEs increased with CONTROL3 scores (see discussion in paper) and prior confidence (HIT_CONF). The average length of the prediction intervals submitted by the subject show positive effect on hit rates, while the average level of predictions (PRED) show reversed negative effect. The RAT indicator for successful prediction of the change in TA100 along the 3 months test period also shows positive significant effect on hit rates. The output of 2 illustrative estimations is provided below: First, we estimate the HIT_CONF and CONTROL3 effects separately. In the second estimation we multiply the two scores using CONTCONF=HIT_CONF*CONTROL3 to estimate the joint effect.

Illustrative separate estimations of the CONTROL3 and HIT_CONF effects:

Discrete Response Profile of TOTAL_HIT			
Index	**Value**	**Frequency**	**Percent**
1	0	17	18.28
2	10	19	20.43
3	20	14	15.05
4	30	15	16.13
5	40	10	10.75
6	50	4	4.30
7	60	5	5.38
8	70	4	4.30
9	80	1	1.08

Discrete Response Profile of TOTAL_HIT			
Index	**Value**	**Frequency**	**Percent**
10	90	2	2.15
11	100	2	2.15

Model Fit Summary	
Number of Endogenous Variables	1
Endogenous Variable	TOTAL_HIT
Number of Observations	93
Log Likelihood	-152.38743
Maximum Absolute Gradient	5.81058E-6
Number of Iterations	72
Optimization Method	Quasi-Newton
AIC	332.77486
Schwarz Criterion	368.23125

Goodness-of-Fit Measures		
Measure	**Value**	**Formula**
Likelihood Ratio (R)	85.064	2 * (LogL - LogL0)
Upper Bound of R (U)	389.84	- 2 * LogL0
Aldrich-Nelson	0.4777	R / (R+N)
Cragg-Uhler 1	0.5993	1 - exp(-R/N)
Cragg-Uhler 2	0.6085	(1-exp(-R/N)) / (1-exp(-U/N))
Estrella	0.6437	1 - (1-R/U)^(U/N)
Adjusted Estrella	0.4849	1 - ((LogL-K)/LogL0)^(-2/N*LogL0)
McFadden's LRI	0.2182	R / U
Veall-Zimmermann	0.5917	(R * (U+N)) / (U * (R+N))

Goodness-of-Fit Measures		
Measure	**Value**	**Formula**
McKelvey-Zavoina	0.632	
N = # of observations, K = # of regressors		

Algorithm converged.

Parameter Estimates					
Parameter	**DF**	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > |t|**
avg_LGTH	1	0.120281	0.015316	7.85	<.0001
avg_PRED	1	-0.181402	0.026495	-6.85	<.0001
CONTROL3	1	0.041537	0.031980	1.30	0.1940
HIT_CONF	1	0.090385	0.049441	1.83	0.0675
RAT	1	0.554358	0.290205	1.91	0.0561
_Limit2	1	0.881147	0.175898	5.01	<.0001
_Limit3	1	1.495984	0.210692	7.10	<.0001
_Limit4	1	2.212355	0.247971	8.92	<.0001
_Limit5	1	2.834498	0.288788	9.82	<.0001
_Limit6	1	3.201931	0.323713	9.89	<.0001
_Limit7	1	3.761920	0.376158	10.00	<.0001
_Limit8	1	4.287161	0.423253	10.13	<.0001
_Limit9	1	4.528853	0.469215	9.65	<.0001
_Limit10	1	5.339459	0.634549	8.41	<.0001

Illustrative estimation of the joint effect CONTCONF=CONTROL3*HIT_CONF:

Discrete Response Profile of TOTAL_HIT			
Index	**Value**	**Frequency**	**Percent**
1	0	17	18.28
2	10	19	20.43
3	20	14	15.05
4	30	15	16.13
5	40	10	10.75
6	50	4	4.30
7	60	5	5.38
8	70	4	4.30
9	80	1	1.08
10	90	2	2.15
11	100	2	2.15

Model Fit Summary	
Number of Endogenous Variables	1
Endogenous Variable	TOTAL_HIT
Number of Observations	93
Log Likelihood	-153.34358
Maximum Absolute Gradient	0.0000198
Number of Iterations	98
Optimization Method	Quasi-Newton

Model Fit Summary	
AIC	332.68715
Schwarz Criterion	365.61095

Goodness-of-Fit Measures		
Measure	**Value**	**Formula**
Likelihood Ratio (R)	83.151	2 * (LogL - LogL0)
Upper Bound of R (U)	389.84	- 2 * LogL0
Aldrich-Nelson	0.472	R / (R+N)
Cragg-Uhler 1	0.591	1 - exp(-R/N)
Cragg-Uhler 2	0.6001	(1-exp(-R/N)) / (1-exp(-U/N))
Estrella	0.6342	1 - (1-R/U)^(U/N)
Adjusted Estrella	0.4855	1 - ((LogL-K)/LogL0)^(-2/N*LogL0)
McFadden's LRI	0.2133	R / U
Veall-Zimmermann	0.5847	(R * (U+N)) / (U * (R+N))
McKelvey-Zavoina	0.6455	
N = # of observations, K = # of regressors		

Algorithm converged.

Parameter Estimates					
Parameter	**DF**	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > |t|**
avg_LGTH	1	0.119783	0.015208	7.88	<.0001
avg_PRED	1	-0.173740	0.026045	-6.67	<.0001
CONTCONF	1	0.013182	0.003098	4.25	<.0001
RAT	1	0.546369	0.290913	1.88	0.0604
_Limit2	1	0.811609	0.159737	5.08	<.0001
_Limit3	1	1.404920	0.194340	7.23	<.0001
_Limit4	1	2.108750	0.232573	9.07	<.0001
_Limit5	1	2.732399	0.275953	9.90	<.0001
_Limit6	1	3.108264	0.314433	9.89	<.0001
_Limit7	1	3.679150	0.370264	9.94	<.0001
_Limit8	1	4.205801	0.418022	10.06	<.0001
_Limit9	1	4.452456	0.466896	9.54	<.0001
_Limit10	1	5.289228	0.638537	8.28	<.0001

Web Supplement IV: Prior Confidence in Choice and Prediction Abilities

Prior confidence regarding stock-selection ability was elicited at the last page of the questionnaire, where each participant separately estimated the number of BUY and SELL problems where personal choices would turn out correct. Most respondents expected at least 3 correct choices in each condition and the median CCR estimate was 60%. Paired comparisons of prior estimates and actual CCRs confirmed that the participants were significantly too optimistic regarding their stock selection ability ($p<0.01$).[1] Across the sample, the correlation between subjective assessments and actual CCRs is positive $\rho=0.07$, but far too weak for significance. The predictive power of prior estimates however strengthens when the relatively difficult questionnaires are ignored. Comparisons for the 50% easiest questionnaires, reveal CCR of 56% to the participants that expected at least 6 correct choices, compared to 47% for the less confident respondents ($p=0.08$). When the comparison is further restricted to the easiest quartile ($N=23$), the CCR for the highly confident is 64% compared to 43% for the least confident ($p=0.02$). Probit regressions (Web supplement II) confirm that correct choice rates significantly increase with the anticipated rates when DIFFICULTY is accounted.

Similar, but stronger, results emerge for prior belief in prediction abilities. Again, subjective confidence was separately elicited for BUY and SELL at the last page of the questionnaire. On average, the participants expected almost 3 hits in each condition and the median subjective hit rate assessment was 60%. Paired comparisons reveal that the expected hit rates were significantly lower than the 90% perfect-calibration benchmark ($p<0.01$), confirming that subjects partially anticipated their low calibration rates in advance (as in Gigerenzer, 1991; Glaser et al., 2005).[2] The actual HIT RATE (27.5%), however, was still about half lower than the subjective estimate. Across the sample, prior hit rate assessments, interval lengths and actual HIT RATEs are mutually correlated (see Table III). The relatively confident participants submitted longer

[1] The high rates of overconfidence could partially follow from the separate elicitation of CCR estimates for BUY and SELL, as subjects refrained from providing unequal estimates for the two conditions. Similar rates of overconfidence, however, were observed in Törngren and Montgomery (2004) where subjects provided probabilistic assessments regarding the correctness of choice in each problem.

[2] The inconsistency between task specific confidence levels and overall success estimates was first documented in Gigerenzer (1991) who argues that the probabilistic mental model employed by the judge may vary with the task. Yaniv and Foster (1995) dual theory of interval forecasting suggests that subjects may wary about informativeness when constructing specific narrow intervals, but consider accuracy more realistically when estimating overall hit rates. Our results here are very close to Glaser et al., (2005; 2012) where respondents subjectively expected 6 hits in 10 banking and finance 90% interval predictions, while the actual number of hits was 2.2

intervals and indeed achieved higher calibration rates compared to the less confident. The predictive power of prior confidence moreover strengthens when CONTROL scores are accounted: the HIT RATE of respondents with higher than median CONTROL and prior hit rate estimate≥8 (N=15) was 45.3%.

Table III: Correlations between Subjective and Actual Hit Rates

	Actual HIT RATE	**Average LGTH**
Subjective HIT RATE estimate	0.21 (p=0.04)	0.23 (p=0.02)
Average LGTH	0.63 (p<0.01)	-

Prior confidence, in conclusion, emerges as a weak indicator for successful performance in both choice and predictions tasks. We attribute the weak results to the complexity of binary short-run stock selection and the robustness of informational overconfidence. It is still intriguing to note that the credibility of intuitive confidence significantly improves when tasks are easier or motivation increases.

Finally note that the subjects with higher CONTROL scores (as defined in Section 3.2) showed higher confidence regarding their calibration rates; .e.g., a median split by CONTROL reveals mean confidence level of 63% vs. 53%; p=0.05.

Web Supplement V:

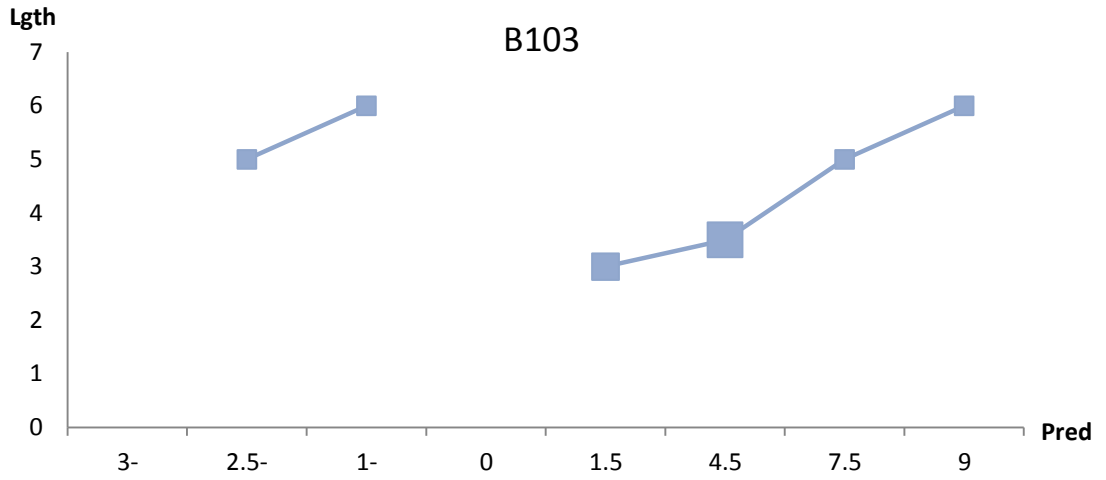
Testing the increase in LGTH with |PRED| at the individual level

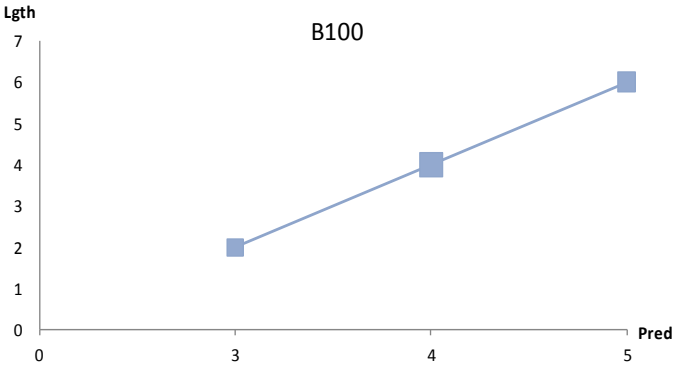
We illustrate the calculations on subject B103. The next table provides the PRED and LGTH of the 10 prediction intervals submitted by the subject (see also the scatterplot below)

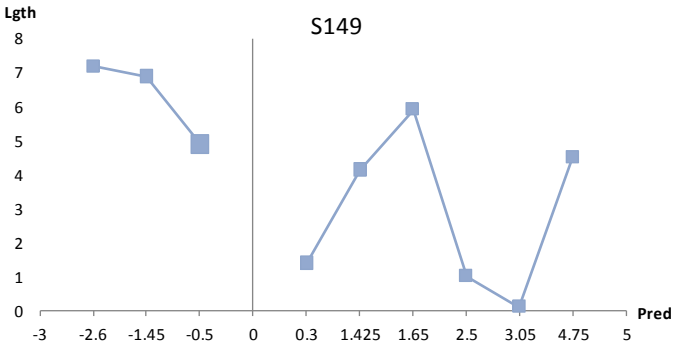
PRED	LGTH
9	6
7.5	5
4.5	3
4.5	3
4.5	3
4.5	5
1.5	3
1.5	3
-1	6
-2.5	5

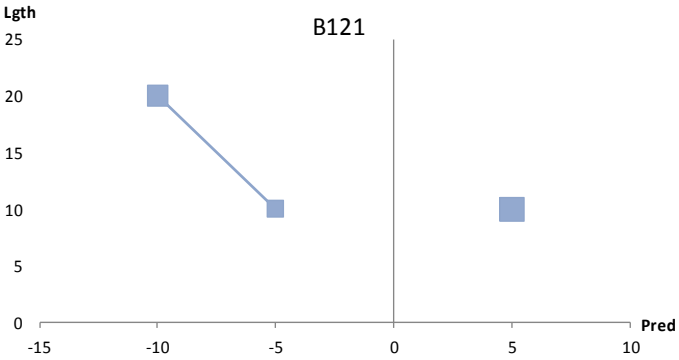
The average length of the 4 intervals centered at 4.5 is 3.5 while the average length of the 2 intervals centered at 1.5 is 3. The comparisons on the positive side therefore deals with 4 successive PRED values 1.5, 4.5,7.5 and 9 where the corresponding average LGTHs are 3, 3.5, 5 and 6. Since the average lengths increase monotonically with PRED we count 3 cases where LGTHs strongly increase in adjacent PRED values at the positive PRED domain. The comparisons for negative PRED, in this case, involve only 2 PRED values where LGTH decreases (from 6 to 5) as |PRED| increases (from 1 to 2.5), so we count one case where LGTH decreases with PRED at the negative domain. Overall, LGTH increases in adjacent |PRED| in 3 of the 4 possible comparisons by our method. (The results are robust to the specific methodology; similar proportions emerge when each of several intervals centered at the same PRED value is counted separately, so that the comparison of PRED=4.5 to PRED=7.5 would suggest that lengths increase in PRED in 3 cases while staying fixed in 1 case; or when the comparison is run between all possible PRED values in each domain so that the intervals at 1.5 are not only compared to those at 4.5 but also to the intervals at 7.5 and 9). Consider next the convexity test for B103. Since the subject has only delivered 2 intervals at negative PRED values, the convexity tests are impossible at the negative side (moreover, LGTH decreases in |PRED| along the 2 intervals centered at negative PRED, so that we would not have counted theses cases in our test of

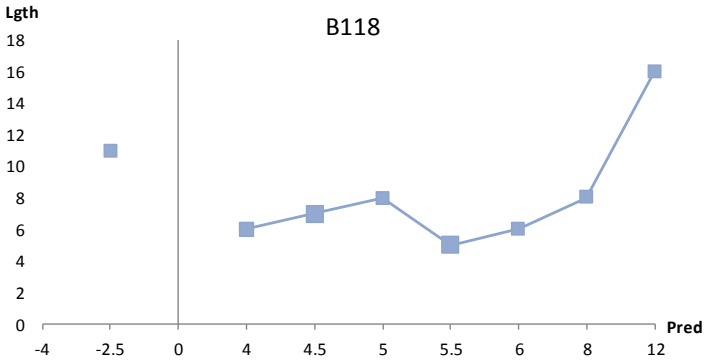
convexity even if the PRED<0 sample was larger). On the positive side, the slope of LGTH is 1/6 along the interval [1.5, 4] (see the figure of B103 below). The slope increases to 1/2 along the interval [4.5,7.5] and increases further to 2/3 over the interval [7.5,9]. Convexity is therefore supported in the 2 possible comparisons for B103. The next page provides the graphs for 7 more subjects:

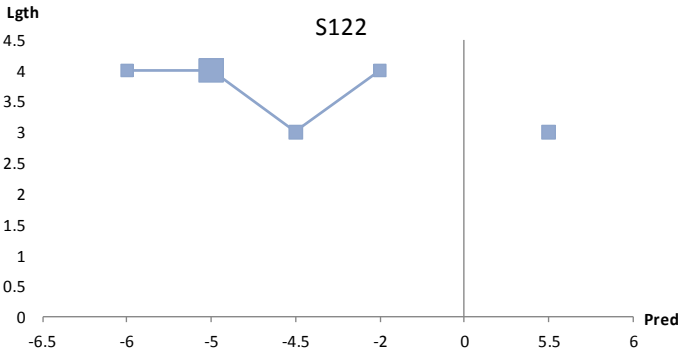


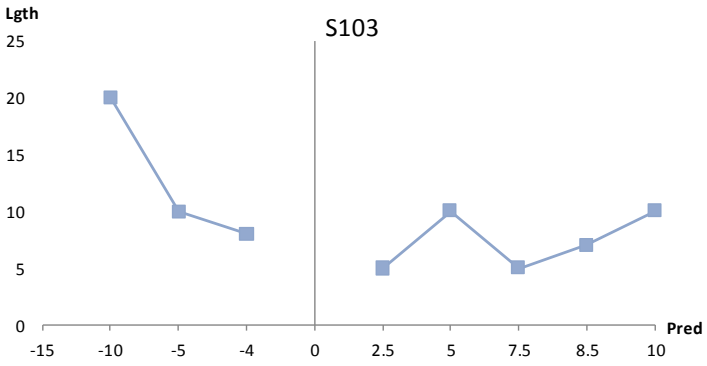

B100
Lgth
Pred
0
1
2
3
4
5
6
7
0
3
4
5

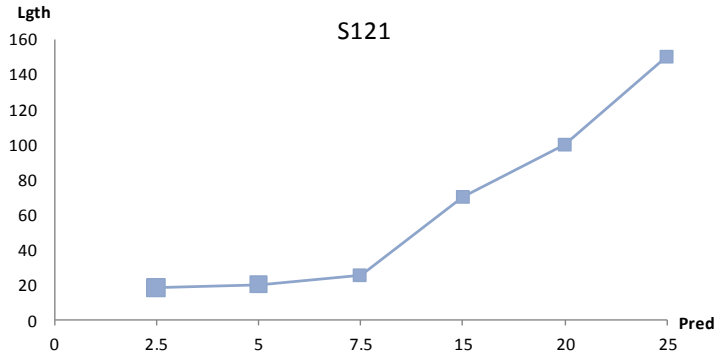

S149
Lgth
Pred
0
1
2
3
4
5
6
7
8
-3
-2.6
-1.45
-0.5
0
0.3
1.425
1.65
2.5
3.05
4.75
5


B121
Lgth
Pred
0
5
10
15
20
25
-15
-10
-5
0
5
10


B118
Lgth
Pred
0
2
4
6
8
10
12
14
16
18
-4
-2.5
0
4
4.5
5
5.5
6
8
12


S122
Lgth
Pred
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
-6.5
-6
-5
-4.5
-2
0
5.5
6


S103
Lgth
Pred
0
5
10
15
20
25
-15
-10
-5
-4
0
2.5
5
7.5
8.5
10


S121
Lgth
Pred
0
20
40
60
80
100
120
140
160
0
2.5
5
7.5
15
20
25

Web Supplement VI: Translated Questionnaire and details for Experiment II

An Experiment on Return Prediction from Historical Series

Welcome to an experiment on return prediction!

The experiment is run by researchers at the College of Management to investigate patterns of return prediction from selected historical information.

Along the experiment, you would be asked to assume the role of investment consultants of hypothetical clients that deliberate the timing of purchase or sale of given stocks

In some of the assignments (**SELL assignments**), the client is assumed to hold 100,000 NIS of the stock. The client would like to liquidate the investment and asks for your advice regarding the best timing: should he sell the stock immediately or postpone the sale by one month exactly (anticipating that the price of the stock would increase in the following month)

In the other assignments (**BUY assignments**), the client is contemplating the purchase of a stock in the amount of 100,000 NIS. The client asks for your advice regarding the best timing: should he buy the stock immediately or postpone the purchase by one month (anticipating that the price of the stock would decrease in the following month)

The experiment consists of 2 SELL assignments and 2 BUY assignments.

In each assignment, you are requested to provide a median prediction regarding the expected return on the stock in the following month and a 90% confidence interval for the return

The prediction assignments composing the questionnaire are based on real data from the American stock exchange. One of the problems, for instance, refers to the performance of a trade and services stock in the mid-seventies while one of the other problems deals with the performance of an industry stock at the first years of the new millennium. **The identity of the stocks and the specific inspection dates would not be disclosed. The experiment intends to test your predictions from partial historical information!**

Each of the consultation problems that would be presented along the questionnaire is actually based on series of 13 successive monthly returns: MON1-MON13. Month 13 is the month to which our timing tasks refer. In each of the consultation problems, you would receive information on the annual return of the stock in the 12 months preceding MON13 (MON1-MON12), the return over the last 6 months (MON7-MON12), the return in each of the last 3 months (MON10, MON11, MON12) and the standard deviation of monthly returns in months 1-12.

Please note: since the consultation problems are based on real historical data, the return for MON13 was already been realized and it is kept concealed in the our files to calibrate your

predictions! In particular, the performance of each stock in MON13 is not related to current or future trends in the stock market

The prediction of returns for month 13 would be carried out in 3 levels:

*** median prediction-** you would be asked to provide a prediction such that the probability that realized return would be lower or higher than your prediction equals 50%

*** upper 95% limit -** you would be asked to provide a prediction such that the probability that realized return would be lower than your prediction is 95%

*** lower 95% limit -** you would be asked to provide a prediction such that the probability that realized return would be higher than your prediction is 95%

Please recall that the interval between the lower 95% limit and the upper 95% limit defines a 90% confidence interval for the return in MON13. The probability that this interval would accommodate the realized return in MON13 equals 90%. With probability 5%, the return would be higher than the upper limit and with similar probability 5% the return would be lower than the low limit.

By the end of the experiment, the organizers would confront your predictions against the realized return in MON13, in order to check your eligibility for a consultation bonus. The bonus is composed of two components:

First component: A bonus for accurate prediction

To determine your bonus for accurate prediction we would randomly select 1 of the 4 consultation problems composing the questionnaire ("**the selected problem**") and subtract the realized return in MON13 from your median prediction to determine your prediction error in the selected problem. The bonus for accurate prediction would decrease with the absolute value of your prediction error. The maximal bonus (paid when the median prediction actually equals the return in MON13) would be 60 NIS.

Second component: A bonus for fair and accurate interval prediction

The prediction of returns by confidence intervals is considered a rather challenging experimental task. On one hand, the interval cannot be too long – your limit predictions should allow for 5% error in each side. On the other hand, the interval cannot be too short. Short interval would not include the predicted return in 90% of the cases, as required. To determine your eligibility for the bonus for fair and accurate prediction, the organizers would check if the intervals that you submitted along the experiment satisfy 3 predetermined conditions. Questionnaires that meet the 3 conditions would receive an additional bonus in the amount of 60 NIS.

The exact formulas for bonus derivation would be provided to the participants by request at the end of data collection. The total amount of bonuses may reach 120 NIS as explained above. The results of the experiment would

be distributed anonymously (identifying the participants by id's) at the end of the experiment – to let subjects control the experimental method and the payment calculations.

Page Break_____

Brief preliminary questionnaire

Page Break_____

{Subjects were not actually paid for experiment II. The instructors announced at beginning of the semester that payout would be distributed for only one of 2 unrelated experiments that would be run along the semester}

{The 4 prediction problems followed next. Each problem was presented in a separate page. The experiment was run in two versions. The differences across the two versions are discussed in the paper. We illustrate the structure of the problems with the following purchase assignment. All problems were presented in similar format}

Assume you serve as investment consultant of a client that deliberates the exact date of purchase of some given stock

In particular, your client is uncertain whether to purchase the stock now (end of month 12) or postpone the purchase by one month (to the end of month 13), in anticipation that the price of the stock would decrease in month 13.

To support his decision, you are requested to provide predictions regarding the anticipated performance of the stock in month 13.

The following table, provides selected information regarding the performance of the respective stock in the 12 preceding months: month1-month12. You are asked to provide your median, upper limit ad lower limits predictions at the bottom of the table.

Please note: The prediction of return from partial historical data is clearly quite challenging. The task may appear impossible. We still ask you for your serious collaboration.

Realized annual return (MON1-MON12)	-0.3%
Realized return in last 6 months (MON7-MON12)	+16.8%
Standard deviation of monthly returns (MON1-MON12)	+6.4%
Return in month 10	+3.2%
Return in month 11	-10.0%
Return in month 12	+3.6%
Predictions for month 13	
Median prediction: with probability 50% the return in MON13 will be lower or higher than:	
Upper 95% limit: with probability 95% the return in MON13 will be lower than:	
Lower 95% limit: with probability 95% the return in MON13 will be higher than:	
Would you recommend a one month delay in purchase?	*NO/YES*

More Details on experiment II

	Version V1 (N=26)				**Version V2 (N=20)**			
	Historical Data				Historical Data			
Problem	Low1	Low2	High2	High1	Low1	Low2	High2	High1
Condition	Sell	Buy	Sell	Buy	Sell	Sell	Buy	Buy
Return MON1-MON12	2.5	-0.3	+28.6	+23.5	-2.5	-0.3	+28.6	-23.5
Return MON7-MON12	-9.2	+16.8	+35.4	+42.0	+9.2	+16.8	+35.4	-42.0
STD12	+4.6	+6.4	+12.5	+18.1	+4.6	+6.4	+12.5	+18.1
Return MON10	-4.7	+3.2	-7.0	-8.5	+4.7	+3.2	-7.0	+8.5
Return MON11	+4.3	-10.0	+24.6	-21.4	-4.3	-10.0	+24.6	+21.4
Return MON12	+0.7	+3.6	-3.4	+2.3	-0.7	+3.6	-3.4	-2.3
	Results (average):				**Results (average):**			
P5	-4.2	-5.4	-6.1	-9.9	-3.1	-4.5	-5.4	-12.8
P50	0.8	1.2	4.6	1.6	1.6	2.0	3.6	-1.6
P95	6.9	7.8	16.1	14.1	4.9	6.5	15.7	14.6
LGTH	11.1	13.3	22.2	24.0	8.0	11.0	21.1	27.4

Web Supplement VII: Detailed version of the Anchoring with Noisy Monotone Adjustments Model

The model extends Tversky and Kahneman's (1974) AAT to explain the pattern of correlations between prediction lengths and absolute predictions as observed in the experiments. We start from the space S of information signals upon which agents construct the confidence intervals. For simplicity, assume S is finite using $\mu(s) > 0$ to represent the probability of each state $s \in S$. In the controlled information design of experiment II, for example, *s* may consist of the 5 statistics provided with each prediction assignment, but individual heterogeneity or cross-sample effects (e.g., range of CONTROL scores) may be accounted by extending the space.[3]

An Anchoring with Noisy Monotone Adjustments (ANMA) process on S consists of the following elements:

(1) An anchor function $A: S \to \mathcal{R}_+$ with $A(s)$ representing the anchor or point prediction from which the interval forecast is constructed when realized information is *s*. For convenience, we assume $A(s) > 0$ for every signal s, but the extension to negative anchors is straightforward.

(2) Similarity Partition $\Pi(S) = \{S_1, S_2, S_3 \ldots S_m\}$ satisfying the standard conditions $S_i \cap S_j = \emptyset$ for each $i \neq j$ and $\bigcup_i S_i = S$, with each S_i representing a set of information signals that are similar in terms of perceived volatility of the prediction target. In experiment II, for example, signals may belong to the same similarity set when STD12 is similar. The conditional standard deviation $\sigma(A|\Pi(s))$ is used to represent the volatility of the prediction environment in state s, with $\Pi(s)$ representing the similarity set containing state *s*. The environment is more volatile in *s* compared to s' when $\sigma\big(A\big|\Pi(s)\big) > \sigma\big(A\big|\Pi(s')\big)$. For the correlation analysis we assume that the anchor admits at least two distinct values in each set; i.e., for every $S_i \in \Pi(S)$, there are two states $s \in S_i, s' \in S_i$ such that $A(s) \neq A(s')$.

(3) Adjustment functions $\Delta U: S \to R_+$ and $\Delta D: S \to R_+$ that represent the expected adjustments (before noise) around the anchor for each information signal *s*. $\Delta U(s)$ denotes the expected upward adjustment from *A(s)*, while $\Delta D(s)$ is the expected downward adjustment from the anchor. The expected confidence interval for information signal s is $[A(s) - \Delta D(s), A(s) + \Delta U(s)]$ and the expected prediction length is $\Delta U(s) + \Delta D(s)$. We henceforth abbreviate $ADJ(s) = \Delta U(s) + \Delta D(s)$ to represent the total expected adjustment. The asymmetric hedging of

[3] Following the macroeconomic prediction literature (e.g., Giordani and Söderlind, 2003) we ignore incomplete information in signals. More generally, *s* may represent information sets.

forecasts is easily captured in this framework assuming $\Delta U(s) < \Delta D(s)$. In addition, following the results of experiment II, we assume that expected adjustments increase with the anchor within similarity sets, while increasing with perceived volatility between similarity sets:

3.1: Adjustments increasing with A, within similarity sets:

For every $S_i \in \Pi(S)$ *and* $s, s' \in S_i$ *such that* $A(s) > A(s')$, $ADJ(s) \geq ADJ(s')$

3.2: Adjustments increase with Volatility, between similarity sets:

For every $s, s' \in S$ *such that* $A(s) = A(s')$ *and* $\sigma\left(A\middle|\Pi(s)\right) > \sigma\left(A\middle|\Pi(s')\right)$, $ADJ(s) \geq ADJ(s')$.[4]

In words, 3.1 requires that expected adjustments increase with the anchor when volatility, as represented by $\sigma(A|S_i)$, is accounted. The assumption is motivated by the increase in prediction lengths with absolute median predictions for given STD12 in experiment II, and the increase in LGTH with PRED when predictions are separately examined for each TA25 stock in experiment I.[5] Assumption 3.2, on the other hand, relates to the increase in prediction lengths with STD12. We choose a weak version of the assumption, demanding that expected adjustments increase with volatility when the anchor is accounted.

(4) Independent Adjustment Noise Φ with mean zero and finite variance σ_Φ. The noise is fixed in distribution and independent of A. The length of predictions when noise is taken into consideration equals LGTH=ADJ+Φ. When information is $s \in S$ and realized noise is ϕ, realized LGTH would be $\Delta U(s) + \Delta D(s) + \phi = ADJ(s) + \phi$.

We now briefly demonstrate that ANMA may capture the pattern of correlation between prediction lengths and absolute predictions as observed in the experiments. First, it is easily verified that the correlation between LGTH and A, conditional on S_i, is positive:

By definition,

$$\rho(LGTH, A|S_i) = \frac{COV(LGTH, A|S_i)}{\sigma(LGTH|S_i) * \sigma(A|S_i)}$$

Since the adjustment noise Φ is independent of A,

[4] We use strong inequalities in 3.1 and 3.2, assuming that expected adjustments may be affected by other statistics/information when the anchor and perceived volatility are equal.

[5] But the perceived volatility of a given stock may vary across subjects when information is not restricted

$$\rho(LGTH, A|S_i) = \frac{COV(ADJ, A|S_i)}{\sigma(ADJ|S_i) * \sigma(A|S_i) + \sigma_\Phi * \sigma(A|S_i)}$$

Since ADJ monotonically increases with A on S_i (3.1), it follows from Schmidt (2003) that the covariance COV(ADJ, A $|S_i$) is non-negative so that $\rho(LGTH, A|S_i) \geq 0$.[6]

The basic ANMA framework however is too general to guarantee that the correlation increases as the environment becomes more volatile.[7] To explain the increased correlation without unnecessary technical complications, we simply assume that adjustments are linear in the anchor within each similarity set; i.e.,

$\Delta U(s) = \alpha_U + \beta_U(s) * A(s)$ and $\Delta D(s) = \alpha_D + \beta_D(s) * A(s)$, where $\alpha_{U,}\alpha_D$ are constants and $\beta_U: S \to R_+$ and $\beta_D: S \to R_+$ are Π(S)-measurable functions that represent the proportional response to the anchor within each similarity set. Assuming, in addition, that at least one anchor is common to each pair of similarity sets [8], 3.2 implies

(*) $[\beta_{Ui} + \beta_{Di}] > [\beta_{Uj} + \beta_{Dj}]$ whenever $\sigma(A|S_i) > \sigma(A|S_j)$,

where $[\beta_{Uk} + \beta_{Dk}]$ represent the fixed proportional adjustments on similarity set S_k.

With this additional assumption,

$$\rho(LGTH, A|S_i) = \frac{COV(ADJ, A|S_i)}{\sigma(ADJ|S_i) * \sigma(A|S_i) + \sigma_\Phi * \sigma(A|S_i)} = \frac{[\beta_{Ui} + \beta_{Di}] * \sigma(A|S_i)}{[\beta_{Ui} + \beta_{Di}] * \sigma(A|S_i) + \sigma_\Phi}$$

and (*) implies that the correlation $\rho(LGTH, A|S_i)$ increases in $\sigma(A|S_i)$. When the environment is relatively stable the noise Φ, that may represent the secondary factors that affect the length of predictions at the individual or task specific level, may strongly decrease the correlation as witnessed in experiment II.

---

[6] The more frequent violations when differences in volatility are small (e.g., comparisons of Low1 vs. Low2) compared to cases where differences are large (e.g., High1 vs. Low1) would be explained if we additionally assume that $ADJ(s1) - ADJ(s2)$increases with $\sigma(A|\Pi(s1)) - \sigma(A|\Pi(s2)) > 0$, so that larger differences in volatility trigger larger differences in expected adjustments.

[7] Adjustments may increase with conditional volatility as assumed in 3.2 while the conditional correlation between the anchor and total adjustments decreases. The intervals, for example, may be longer but almost fixed size in the high volatility environment vs. short and proportional to the anchor when volatility is low. In this case, the correlation between LGTH and A would decrease with conditional volatility.

[8] for every S_i and S_j there exist states $s_i \in S_i$, $s_j \in S_j$ such that $A(s_i) = A(s_j)$.

To conclude that average interval lengths $E(ADJ|S_i) \equiv \sum s \epsilon S_i [ADJ(s) * \mu(s|S_i)]$ increase with $\sigma(A|\Pi(S_i))$, we could further assume that the anchor increases in first order dominance when the environment gets more volatile. Assumptions 3.1 and 3.2 then guarantee that $E(ADJ|S_i)$ increases with $\sigma(A|\Pi(S_i))$. Alternatively, a stronger version of 3.2 requiring that $ADJ(s) \geq ADJ(s')$ whenever $\sigma\big(A\big|\Pi(s)\big) > \sigma\big(A\big|\Pi(s')\big)$ also gives the result.