Annals of Financial Economics
Vol. 9, No. 1 (June 2014) 1450002 (60 pages)

DOI: 10.1142/S201049521450002X

World Scientific
www.worldscientific.com

SHORT-RUN ARBITRAGE IN CRISIS MARKETS — EXPERIMENTAL EVIDENCE

DORON SONSINO* and TAL SHAVIT

School of Business
COMAS — College of Management Academic Studies
7 Rabin Blvd., P. O. B. 9017, Rishon LeZion, Israel 75190
**sonsinod@colman.ac.il*

Published 30 September 2014

The field experimental approach was utilized to collect expectations-arbitrage portfolios from competent investors in late 2008 where stock prices shrunk by 50%. Positions were closed after three months and the four-factor model was applied to characterize strategies and derive risk-adjusted returns. In line with classic judgment literature findings (Lichtenstein *et al.*, 1982), performance significantly improves with prior self-confidence, although the participants exhibit typical patterns of overconfidence. The time-series estimations reveal that the experimental arbitrageurs generally benefited from "leveraging the crisis", but the highly confident delivered positive alpha beyond loading on common premia. The experimental results are discussed in light of the literature on expertise and stock selection in crisis markets.

Keywords: Experimental arbitrage; sub-prime crisis; overconfidence; four-factor model.

JEL Classifications: C9, G1, D8

1. Introduction

More than 130 preregistered competent participants were invited to open experimental arbitrage positions on the Israeli stock market at the second half of 2008, in the midst of the sub-prime crisis. The task was introduced as "arbitrage in expectations" and subjects were instructed to concurrently select stocks for purchase and short-sale in similar volume, assuming positions would be closed approximately three months later and participation payouts would be derived from realized market returns. The paper analyzes the stock-selection and performance of the experimental arbitrageurs while controlling the unique crisis conditions under which the arbitrages were run. The Carhart (1997) four-factor model is adapted to

*Corresponding author.

the specific settings of the experiment to characterize arbitrage styles on individual basis and test the risk-adjusted performance of the arbitrageurs.

The specific type of arbitrage examined in the experiment is frequently addressed as "relative value arbitrage" (e.g., Dubil, 2004) but we have preferred the somewhat more conservative "arbitrage in expectations" (Gatev *et al.*, 2006) for the instructions.[a] The relative value arbitrageur is assumed to act upon perceived mispricing of similar stocks — concurrently short-selling the relatively overpriced stocks while purchasing the underpriced securities. If the mispricing diagnosis is correct, arbitrage positions would be closed in gain when mispricing is eliminated. The introductory pages of the experiment illustrated the mechanics of expectations arbitrage demonstrating that profits are anticipated even in sluggish markets, as the "good" stocks should decrease less steeply than similar "bad" stocks when price gaps eradicate. To emphasize the risky nature of such arbitrage, another example showed that positions might be closed in large loss if expectations are disrupted and prices keep diverging. The instructions underlined the importance of selecting similar stocks long versus short and directed the participants to an Internet chapter of William Goetzmann for details.[b]

While the parallel purchase and sell commands collected in the experiment were not actually implemented, the interface was carefully designed to elicit meaningful portfolios that skilled investors could pursue for speculative profit. Participants were instructed to build the arbitrage from the largest (most tradable) 220 stocks in the Israeli market, and the Web-based program provided direct links to databases with detailed information on stocks, companies and general market conditions. The participants could logoff the experiment and re-enter repeatedly while deliberating choices. The instructions introduced friction into the payout calculations, to turn the environment more realistic and force subjects account transaction costs.

Relative value arbitrage is frequently addressed in relation to market efficiency. Strong-form efficiency (Fama, 1970) implies that mispricing gaps should be rare in highly competitive markets. The detection of exploitable arbitrage indeed attracts vast professional resources (Beunza and Stark, 2004), while students of finance are customarily forewarned that the experts are in charge of keeping markets relatively efficient (e.g., Levy and Sarnat, 1994). Our experimental arbitrage tasks therefore

[a] Since arbitrage is out of equilibrium concept, formal definitions are problematic and existing terminology seems somewhat loose. The classic risk-free arbitrage is the easiest concept, but it is fundamentally different from the type of arbitrage explored in this paper. The term "risk-arbitrage" is frequently reserved for arbitrage transactions evolving around mergers and acquisitions (see, for instance, Mitchell and Pulvino, 2001). "Statistical arbitrage" is informally used in industry-jargon while discussing arbitrages of the type run in the current experiment (Whistler, 2004), but researchers sometimes define statistical arbitrage in terms of conditions on long-run return dynamics (Hogan *et al.*, 2004). Seeking terms that do not contradict these uses, we adopted "arbitrage in expectations" for the instructions, but henceforth also use "relative value" or "mispricing arbitrage" occasionally.

[b] http://viking.som.yale.edu/will/finman540/classnotes/class6.html.

seem especially fit for testing strong-form efficiency versus the hypothesis that skilled investors may recognize price correction opportunities in advance.

Mispricing arbitrage, moreover, is exclusively run for speculative gains. Predetermined investment policies, diversification concerns, liquidity constraints, tax considerations and other limitations that bind portfolio management in practice should play smaller role in *ad hoc* arbitrage transactions.[c] The selection of stocks for short-sale, in particular, opens a window for examining the ability of professionals to recognize overpriced stocks, independently of the long-side positions that the investor maintains. Selling decisions within portfolio management are essentially different as stocks for sale must be selected from existing investments.

Empirical data on mispricing arbitrage however is rather scarce. Some studies examine the performance of hedge funds which employ arbitrage techniques heavily (Eling, 2009). Many papers study different types of risk free arbitrage (Hao *et al.*, 2010) or specific algorithms for large-scale statistical arbitrage (Hogan *et al.*, 2004), but micro-level arbitrage transactions of professional traders are rarely exposed and empirical data cannot be extracted from available sets. Alternatively, we adopt the field-based experimental approach (Harrison and List, 2004) to test the stock-selection of skilled investors in real market arbitrage tasks.

The experiment was launched in mid May 2008 where the subprime crisis was shadowing the local economy. Stock prices showed extreme volatility in the first few months of year. The leading TA100 index decreased by 10% in January, regained 4.5% in February, but lost 13% again in March. Trends were mixed in April–August, but the crisis struck the market with full force in September–October where the main indices plunged by almost 30% (see Fig. 1). Prices slowly recovered at the first few months of 2009, but the index was still 50% lower than its early 2008 level in March 2009 where our latest portfolio (collected in 19 December 2008) was closed.

The strong uncertainty conditions under which the arbitrages were collected and run clearly shows when participants estimate the TA100 trend for their three-month arbitrage. About 59% expected recovery, while 41% expected additional decrease. Opinions did not converge towards 2009: the proportion of pessimists amongst the quartile of "latest participants" was still about 42%. Since the arbitrages were run over distinct intervals in times of extreme volatility, we control for specific market conditions throughout the analysis.[d]

---

[c] Stock selection for arbitrage is clearly affected by other constrains such as tradability, regulatory restrictions, and relatively large transaction costs.

[d] TA100 expectations were elicited in five categories: (1) more than 10% increase (2) 5–10% increase (3) up to 5% increase (4) decrease by no more than 5% (5) more than 5% decrease. On average, TA100 (TA220) decreased by 16.8% (16.5%) along the 130 arbitrage intervals. The index has increased along the arbitrage in 11 cases. The returns on these 11 portfolios (mean −10%; median +4%) were not significantly different from others. Web supplement A provides more information on arrival dates. The translated script, data, and supplements are available at http://www2.colman.ac.il/business/doron.

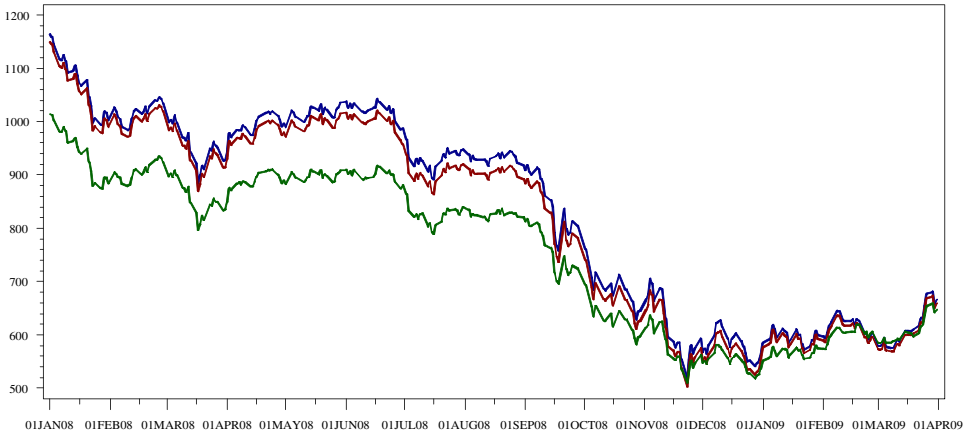


Figure 1. TASE indices from January 2008 to March 2009.

The bottom-line arbitrage results for the complete sample ($N = 130$) were positive but marginal in magnitude and significance. The mean gross three-month return (before experimental fees) was 2.7% and the proportion of portfolios showing positive result was 55%. Similar weakly positive results emerge in a variety of robustness analyzes and profitability does not show trends; e.g., close to zero ($\rho = 0.03$) correlation between eventual returns and the DATE where the portfolio was received.

Closer cross-sample analysis however exposed strongly positive results for the participants that showed high confidence in their profitability chances *ex-ante*. Individual confidence (henceforth: CONF) was elicited at the last page of the program, where each participant estimated the chances that her portfolio would show positive return in 0–100% scale. The intuitive confidence score robustly emerges as the most significant predictor of eventual performance in model-selection analyzes. The mean three-month return on the 68 portfolios with CONF $\geq$ 60% was about 8%, while the most-confident quartile earned almost 13%. The confident subjects significantly overestimated their profitability chances (e.g., mean CONF 82% versus actual profitability rate of 65% for the most confident quartile), but still delivered mean daily ALPHA of 16.4 basis points which accumulates to more than 10% return along the three months (60 trading days) arbitrage. The positive confidence-performance relation, by our results, is reminiscent of the classic "calibration curves" that manifest in dozens of psychological judgment studies (cf. Lichtenstein *et al.*, 1982 for the early findings; Klayman

et al., 1999 for later results). The paper shows that similar patterns may emerge in challenging tasks such as short-run arbitrage stock selection.

Our second main result deals with the crisis-prone arbitrage styles adopted by the experimental arbitrageurs. Comparison of Fama and French (1992) historical risk factors long versus short suggests that the participants tended to sell riskier stocks in terms of BETA, SIZE and book-to-market (BTM). The mean BTM ratio for the stocks selected for purchase was about 0.56 compared to mean BTM ratio of 0.75 for the stocks selected for sale. The stocks selected for investment were also larger (in SIZE) and slightly less aggressive (in market BETA) than the stocks selected for short sale. The crisis-sensitive CONSTRUCTION and FINANCE stocks captured (on average) about 35% of the short-sale volume, but covered only 12% of long-side investments. The participants, in general, constructed crisis-prone portfolios that may benefit from stronger decline of riskier stocks if the crisis continues.

To test if "leveraging the crisis" contributed to eventual profitability, we run four-factor regressions (Carhart, 1997) to estimate the loadings of individual portfolios on MARKET, SIZE, BTM and MOMENTUM returns. The estimations were separately run for each portfolio using the daily return-series for the three-month starting at the date where the arbitrage was received. The premia calculations were adapted to the specific settings of the experiment; e.g., taking into account the restricted list of 220 stocks. The time-series analysis robustly confirms that the participants benefited from short-selling relatively riskier stocks at the pick of the crisis. The strongest results are obtained for the BTM factor. The average three-month BTM-premia along the experiment was -7.9%, suggesting that the participants could earn almost 8% return by buying the stocks with lowest BTM ratios while short-selling the stocks with highest BTM. The mean β_{BTM} coefficient (representing the loading on daily BTM premia) was -0.11 and loadings on BTM were negative for 64% of the portfolios. The negative loadings on negative premia represent the gains from selling relatively riskier stocks in crisis times where realized premia are negative. Negative loadings were also observed for SIZE, while loadings on MARKET and MOMENTUM appeared mixed.

On top of these loading patterns, the estimations reveal average daily ALPHA of 0.03% which accumulates to about 1.8% along the three-month arbitrage. Alpha is positive and significant for 12% of the portfolios, while it is significantly negative for only 6%. ALPHA is strongly correlated with gross arbitrage returns (Pearson correlation $\rho = 0.7$), suggesting that the superior performance of selected subsamples cannot be solely attributed to increased loading on familiar premia. The estimation results are robust to the premia calculation method and reaffirm in many alternative specifications; e.g., when MARKET premia is recalculated assuming crisis-sensitive sectors are traded on one side, while less vulnerable sectors are traded oppositely.

The paper proceeds as follows: The experimental method is discussed in Sec. 2, while Sec. 3 presents the bottom-line results. Section 4 runs the comparisons of long versus short arbitrage positions and Sec. 5 discusses the four-factor estimations. Section 6 concludes.

2. The Experiment

2.1. *Participants*

Preliminary calls for participation were distributed in MBA classes, circulated to alumni and posted in selected financial forums. Subjects that emailed the requested information and met our thresholds received in response the URL of the experiment, a username, and a private password. Usernames were valid for two weeks and participants could disconnect from the program at any point and reconnect later. The instructions emphasized that all feedback regarding results and payouts would be distributed anonymously using ids to identify users. The earliest participant submitted the arbitrage in May 17th, while the last portfolio was received in December 19th. The mean age of the 130 participants was 32% and 89% were male. The sample is extremely educated with an average of 16.4 years of formal education; 57% held or currently pursued an MBA degree. Only 28 participants held investment management, consulting or analysis jobs, but 52% ($N = 67$) proclaimed past industry-experience. The average tenure of the experienced subjects was only 3.8 years, suggesting that the experiment attracted "young industry veterans" that did not face the barriers that could intimidate present professionals.

2.2. *The questionnaire*

The experiment was run on a modular receptive interface spanning over 14 pages. Participants could reexamine the instructions repeatedly but could not revise answers or choices after submission. The program measured the number of repeated logins, the total login-time and the time spent on selected pages. On average, subjects took 2.3 repeated logins to complete the experiment; only 22% completed all tasks at once. The mean total login time was about 47 min, but 56% (73 of 130) took less than 15 min to submit the questionnaire. An automatic control-mail was sent to our address whenever a portfolio was delivered. A confirmation note with additional details on payout procedures was sent in parallel to the participant.

Stock selection was restricted to the 220 largest stocks in Tel-Aviv Stock Exchange (TASE). The stocks in TA220 account for more than 90% of the stock market value and therefore include most of the actively traded stocks in the

Stock Purchase or Deposit			**Stock Sale or Loan**	
Stock Name	Amount Purchased		Stock Name	Amount Sold
▼			▼	
▼			▼	
▼			▼	
Total	1000		Total	1000

Figure 2. The arbitrage table.

exchange. The abbreviation TA220 is henceforth used to denote the list of permitted stocks, or the corresponding price index, depending on context.[e]

Figure 2 presents the table through which subjects delivered the arbitrage. In addition to selecting stocks from TA220, the participants could either deposit or borrow funds in a fixed three-month interest rate of 2%. The "DEPOSIT" and "LOAN" options were appended to TA220 and the joint list was accessible, as a drop-down menu, in each row of the table. Long-side positions were submitted through the left panel of the table. Subjects could select up to three distinct items from the menu and freely allocate the 1000 NIS arbitrage volume among the selected investments.[f] Sell-side transactions were delivered through the right-panel of the table. Again, the participants could choose 1–3 separate items but the total volume of sales and loans had to cover the volume of investments: 1000 NIS. Three examples illustrated the possibilities to construct minimal arbitrage (buying one stock and selling another), borrow funds within the arbitrage, or deposit some of the arbitrage volume. The participants were directed to three web sites, including the formal site of TASE, with vast information on available stocks and underlying companies.

Since the purchase and sale commands where not actually executed, we introduced artificial noise to turn the interface more realistic. The instructions explained that positions would be opened 1–5 days after submission and closed 87–93 days after opening, where the exact delay and duration would be randomly assigned. In addition, we deducted 2% of the arbitrage volume for "fees and commissions". The extra features were only implemented in calculating the payouts for participation. In the following sections we analyze the returns, ignoring

[e] As the 220 largest stocks are not formally indexed, we tentatively joined the stocks in TA100 (the 100 leading stocks index) and YETER-120 (the next 120 largest stocks index) to construct the list of permitted stocks. This paper focuses on the short-run arbitrage for three months. A one-year arbitrage task was run in parallel; but the samples are different and the results are analyzed in a separate paper.

[f] The exchange rate fluctuated between 3.36 and 3.74 New Israeli Shekel for 1 US$ around the experiment.

fees and costs, for the exact three calendar months starting at the date where the portfolio was received.[g]

Attempting to prohibit subjects that seek 50% chance for profits with random stock selection, we downplayed monetary incentives, emphasizing the interest that participants might find in the experiment and its results. The instructions generally explained that 100 randomly selected subjects would receive checks that increase with personal arbitrage payoffs. The exact payout formula was provided in a separate optional link. The textbox explained that the net payoff upon closing the arbitrage (using the random open and close dates and subtracting the 2% fees) would be multiplied by 4 for rough annualization, adding 200 NIS to the product to offset possible losses. If the balance was positive, the participant received a check for 50% of the amount.[h]

In addition to the main arbitrage task, subjects were requested to rank their theoretical background in finance and their familiarity with the local market in 1–7 scale; participate in a brief familiarity quiz; and mark their choices in a short risk-preference task. The Web-tailored quiz was composed of three multiple choice problems (easier, intermediate and challenging) that were randomly drawn for each participant from larger pools of 25 problems. Participants were requested to mark their choices in less than 180 s and the program proceeded automatically to the next page when time was over. The risk-preference task consisted of six standard binary choice problems, between two-outcome lotteries and a certain positive payoff. The first three problems roughly measured the inclination to take risk in gains domain, while the last three problems similarly tested for risk attitudes where losses are possible. The number of cases where the participant preferred the risk-free alternative in the first three problems is henceforth used as an indicator for RISK AVERSION; while the number of risk-free choices at the last three problems is used as tentative LOSS AVERSION measure.[i]

2.3. *Statistics and terminology*

We employ non-parametric tests throughout the paper. Sign-tests and signed-rank tests are used, depending on context, to test the significance of returns, premia and related variables, while the Pitman test is used for between samples comparisons. Small-case p denotes 1-tail significance throughout the paper, using bolded **p****

[g] The quantities purchased and sold were calculated using the closing prices for the DATE of delivery; returns were determined using the adjusted closing prices three months later.

[h] In devising the payout scheme, we did not anticipate the extreme collapse of the market. The initial balance of 200 NIS could only cover 20% annual loss, while actual losses were steeper for 46 participants. Most subjects, however, skimmed the optional payout page very briefly; the mean inspection time was 4 seconds and the maximum 33.

[i] The 3 RISK AVERSION problems asked subjects to choose between a fixed lottery paying +200 or +50 with equal probabilities and risk-free payoffs of 165, 125 and 85. The LOSS AVERSION problems similarly used a +200 or −50 lottery with risk-free payoffs of 105, 75 and 45.

where $p \leq 0.01$; plain p^{**} when $0.01 < p \leq 0.05$ and p^{*} in marginal cases where $0.05 < p < 0.1$. The discussions sometimes figuratively address the volume of "purchase" or "short-sale" although transactions were not actually executed. Arbitrage returns (if this deserves clarification) were calculated by dividing the net payoff upon closing the arbitrage positions by the 1000 NIS volume. If the participant, for example, purchased equal amounts of two stocks that decreased by 10% and 30% correspondingly while short-selling another stock that decreased by 35%, then the payoff is 150 and the return R is 15%. Some discussions distinguish the 35% positive return on the sell-side from the negative -20% return "long".

3. Results

3.1. *Arbitrage returns*

The mean (median) return on the 130 arbitrage portfolios was 2.7% (1.9%; $p = 0.08$ by Wilcoxon signed-rank test), and the proportion of portfolios with positive return (71 of 130) does not suffice for sign-test significance ($p = 0.17$). The bottom-line return data is very noisy with standard deviation 26% and individual returns ranging from a minimum of -108% to a maximum of 64%.[j] The weakly positive results, however, are very robust and reemerge with different filtrations. When the 5% extreme cases are removed on both sides, the mean return increases to 3.4% and the standard deviation decreases to 19% ($N = 116$; $p < 0.05$). Performance also strengthens when the 24 participants that invested some amount in risk-free deposit and the 21 portfolios that borrowed some of the funds (discussed separately below) are ignored: the mean return on the remaining 85 portfolios is 4.6% ($p < 0.04$) and the profitability rate slightly increases to 58%. Stock selection appears rather diverse with 63% of the stocks in TA220 ($N = 139$) selected for either purchase or sale.[k] Payoffs stay positive and marginally significant when the portfolios that selected the most popular stocks on either side of the arbitrage are ignored. The payoffs also strengthen when five participants that selected stocks that stopped trading along the arbitrage (and were later erased from TASE) are appended to the sample. Assuming that positions were closed at the last trading price, the mean return on these five additional portfolios amounts to 7.6% with median 4.8%. In the sequel however we focus on the 130 portfolios that did not trade eliminated stocks.

[j] The payoff distribution is sketched in Web supplement B.

[k] Stock selection was relatively focused on the buy side where subjects picked only 84 stocks (40% of the menu), compared to the sell side where the participants selected 112 distinct stocks. Web supplement D provides a detailed analysis of stock selection, testing if the experimental choices could be used to effectively separate TA220 into disjoint recommended BUY and SELL lists. As the results of the analysis are, again, positive but weak, we leave the details to the supplement.

3.2. *Long-side versus short-side results*

Comparison of payoffs long versus short reveals that the positive bottom-line results followed from relatively large gains that the participants obtained on their sell-side transactions. On average, the arbitrageurs made 19.8% positive return closing their short-side positions, but concurrently lost about 17.1% on their long-side investments. The proportion of positive results on the sell-side was 79%, compared to only 22% positive payoffs on the buy-side. The contrasting results long versus short clearly follow from the persistent decline of stock prices along 2008 (Fig. 1). On average, TA200 lost about 16.5% along the three-month arbitrage intervals. The traders therefore earned almost 20% from selling stocks that decreased even more steeply than the market (signed-rank test for comparing sell-side returns to $|\text{TA220}|$; $p < 0.01$), but lost most of these gains by holding stocks that decreases as sharply as the market ($p = 0.32$; N.S).

3.3. *Regressions with model selection*

To identify the moderators of performance across the sample, we ran a comprehensive regression analysis using model selections to extract the significant effects. Logistic regressions were run in parallel to locate the variables that increase the likelihood of positive results. Since the number of variables examined in the analysis is large and most attributes did not show significance, we henceforth focus on the interesting conclusions, ignoring most of the insignificant findings. Selected regression results are provided in Table 1. The list of the variables tested in the analysis is displayed in Web supplement C. To control for specific market conditions, we included the change in TA220 along each arbitrage interval in the list of possible effects. The variable, however, was persistently removed in model selections confirming that market conditions did not affect performance consistently.[1]

Column (a) of Table 1 illustrates that arbitrage returns increased with the probability that participants pre-assigned to the profitability of their portfolio (CONF), while decreasing with the time spent inspecting background material (T (INTRO)). The CONF effect is summarized in Table 2 where we split the sample into approximate quartiles by prior level of confidence. The upmost line refers to the 32 participants that assigned probability less than 50% to eventual profitability. Since even random selection has 50% chance to show positive return, CONF $<$ 50% represents a statement of pessimism or under-confidence. The mean return on the portfolios of the least-confident subjects, however, was negative -5.2% and their positions were closed in loss in 56% of the cases. On the opposite extreme,

[1]Except for the DATE effect on the likelihood of $R > 0$ by model (g) which is mentioned later.

Table 1. Results of illustrative regressions.

Dependent variable	Linear regressions					Logistic	
	Arbitrage return					Indicator on $R > 0$	
	(a)	(b)	(c)	(d)	(e)	(f)	(g)
CONF	**0.15****	**0.168****	0.171**	**0.164****	**0.208****	0.0056	0.01*
	(0.05)	**(0.047)**	(0.07)	**(0.049)**	**(0.068)**	**(0.0044)**	**(0.008)**
T(INTRO)	−0.016**	−0.015**	−0.015**	−0.015**	−0.011*	0.000	
	(0.008)	(0.007)	(0.007)	(0.007)	(0.0074)	**(0.000)**	
DEPOSIT	—	**0.023****	**0.023****	**0.023****	**0.024****	0.033**	0.033**
	—	**(0.009)**	**(0.009)**	**(0.009)**	**(0.009)**	**(0.0012)**	**(0.0022)**
LOAN	—	**−0.031****	**−0.031****	**−0.031****	**−0.030****	**−0.003****	**−0.003****
	—	**(0.007)**	**(0.007)**	**(0.007)**	**(0.007)**	**(0.0018)**	**(0.001)**
EXPECTATIONS	—	—	−0.07	—	—		
			(1.2)				
EXPERIENCE	—	—	—	0.18	—		
				(0.58)			
AGE					−0.28**		
					(0.14)		
SUP OCCU	—	—	—	—	9.2**		
					(4.1)		
PRIV INFO	—	—	—	—	9.6**		
					(5.1)		
EDU							−0.09**
							(0.04)
LOSS-AVERSE							−0.54**
							(0.25)
DATE							0.005*
							(0.003)
R^2	0.07	0.27	0.27	0.27	0.31	—	—

Table 2. Return by level of confidence.

CONF range	N	Mean R	Median R	% ($R > 0$) (%)	Mean CONF (%)
CONF < 50	32	−5.2	−1.5	44	17
$50 \leq$ CONF < 60	30	−1.2	−0.3	50	51
$60 \leq$ CONF < 75	34	3.4	4.6	59	65
CONF ≥ 75	34	12.8	15.4	65	82
All	130	2.7	1.9	55	54

the 34 participants that assigned probability $\geq 75\%$ to eventual profitability earned 12.8% with profitability rate 65%.

Individual CONF levels negatively correlated with years of industry experience ($\rho = -0.26$), while positively correlating with the number of repeated logins ($\rho = 0.19$), but EXPERIENCE and #LOGINS were consistently removed in model-selections when CONF was accounted. Interestingly, CONF shows positive significant effect on performance although subjects in the highest-confidence categories significantly overestimate their success rates; e.g., the mean CONF of the participants with CONF $\geq 75\%$ was 82%, compared to actual profitability rate of 65% ($p < 0.01$).

The positive correlation between self confidence and arbitrage returns is reminiscent of classic "calibration curves" that emerged in dozens of judgmental overconfidence studies (cf. Klayman *et al.*, 1999). The correct choice rates in multiple-choice almanac questionnaires, for example, typically increase with subjective confidence, even though subjects overestimate their success rates significantly. Clark and Friesen (2009) more recently demonstrate that positive relation between prior confidence and eventual performance may emerge (with some experience) in complicated abstract gaming tasks. Our current results interestingly suggest that gut-confidence may prove as a useful indicator in field financial decision, where formal criteria such as education, years of professional experience and other controls fail. The result is especially intriguing given the fact that performance could be summarized only three months after the delivery of confidence scores, in times of extreme market volatility.

In addition, however, the analysis shows that performance significantly deteriorated with T(INTRO), the time spent on the introductory pages where the concept of arbitrage in expectations was illustrated and discussed. The mean T(INTRO) was 269 s, with 1/3 of the participants spending more than 5 min on the introductory material.[m] The Pearson correlation between T(INTRO) and CONF was close to zero (−0.07) and the illustrative regressions in column (a) of Table 1 suggest that an additional minute on the introductory-pages decreased eventual arbitrage returns by almost 1%. When the participants with $60 \leq \text{CONF} \leq 80$ ($N = 58$) are median-split by T(INTRO), the mean return of those that went through the instructions rapidly is 10.5%, compared to marginal mean return of 0.3% on the portfolios with longer T(INTRO) ($p < 0.01$). The time spent on the instructions negatively correlates with skill-related variables (e.g., correlation with industry experience $\rho = -0.16$; correlation with proclaimed market familiarity $\rho = -0.18$), but the regressions robustly suggest that the direct skill-related

[m] Obviously, we cannot control for subjects' engagement while logged-in to various pages. We use the inspection time variables as measurable proxies.

variables do not show significance when T(INTRO) is accounted. It is still possible however that the participants that spent more time on the background pages balanced risks more carefully (as advised in the introductory examples) so that their lower returns follow from less risky positions. The hypothesis is tested in Sec. 5.4 where we test if loading styles changed with T(INTRO).

3.4. *Deposits and loans*

Since stock prices fell on average by 16.5% along the experiment, the portfolios that sold stocks to invest in deposit exhibit very strong performance. The average return on the 24 portfolios that invested some amount in risk-free interest was 16.3%, more than six times larger than the sample mean. On the opposite extreme, the 21 participants that borrowed funds to invest in stocks have missed the gains from short-selling stocks when rates drop drastically, ending with negative average return of −20.7%.[n] In column (b) of the regressions table, we add the amount of DEPOSIT or LOAN to the list of explanatory variables. The fit of the regression increases from 7% to 27%; the coefficients for DEPOSIT and LOAN are opposite in sign and significant at $p < 0.01$. Unsurprisingly, the deposit investors appear significantly more pessimistic regarding the market trends for their arbitrage (proportion of pessimists: 63% for depositors compared to 23% for borrowers), even though the dates where investors and borrowers took the experiment did not differ significantly (mean participation dates: AUG 26 versus AUG 27).[o] When EXPECTATIONS are included as explanatory variable in the illustrative regressions (column c), the coefficient is close to zero and statistically insignificant.

3.5. *Occupation*

Participants were asked to classify into 1 of 4 occupational categories: (a) investments professionals holding management, consulting, or analysis jobs at the time of participation ($N = 28$) (b) other finance related jobs ($N = 43$) (c) professions or jobs unrelated to finance ($N = 42$) (d) self-employed ($N = 17$). The 28 investment-professionals showed about three times larger than average mean return (7.2%) with profitability rate 61%. The 17 self-employed exhibited comparable superior performance with mean return 8.6% and 59% $R > 0$ rate (Pitman test for comparing the professionals and self-employed; $p = 0.4$). In closer examination, the

[n] Seven participants invested all their funds (1000 NIS) ending with mean positive return of 31.6% while 11 borrowed the complete amount, losing on average −23%.

[o] EXPECTATIONS were elicited in five categories: (1) TA100 would increase by more than 10% (2) increase of 5–10% (3) increase up to 5% (4) decrease by no more than 5% (5) decrease by more than 5%. The strong uncertainty clearly reflects in the assessments: 59% of the participants expected an increase in TA100 while 41% expected additional decrease. Opinions do not converge along the experiment: the proportion of pessimists for the quartile of "latest participants" is still about 42%.

self-employed indeed appear as young (median age 27) industry veterans (EXPERIENCE > 0 for 12 of 17) that had recently quit the industry. While sample sizes are not large enough for detailed comparisons, it is interesting to note that occupational classifications show significant effects that could not be generally attributed to related variables like EXPERIENCE, years of formal education (EDU), or proclaimed skills. The correlation between industry experience and arbitrage returns, for instance, was close to zero (0.0028) and the regressions robustly suggest that EXPERIENCE did not affect performance in general (column (d) of Table 1).

3.6. *Arbitrage method*

The final page of the experiment asked participants to select the method that best-describes their arbitrage strategy: fundamental, technical, based on private-information or "other". Those claiming to build on private information ($N = 30$) out-performed others significantly with mean return of 9.7% and 63% $R > 0$ rate. The 20 participants claiming to build their arbitrage by "other" criteria also exhibited superior performance with mean return 3.7% and 70% profitability rate. Seven of these 20 subjects specifically mentioned the economic-crisis in their verbal comments; e.g., claiming to buy stocks of companies that may gain from the economic slump, selling stocks of commercial banks etc. In column (e) of the illustrative regressions table, we add an indicator for "industry-job or self-employment" (SUP OCCU) and a disjoint indicator for "use of private-information" (PRIV INFO) (by participants that did not classify into the superior occupations) to the list of explanatory variables.[p] The estimated coefficients are about 9% each and both are significant at $p < 0.01$. In addition, regression (e) suggests that arbitrage performance marginally deteriorated with AGE (cf. Gottesman and Morey, 2006 evidence for weaker performance of older mutual fund managers). The negative AGE effect manifests clearly in direct comparisons.[q]

3.7. *Logistic regressions*

Columns (f) and (g) of Table 1 briefly summarize the logistic analysis of profitability likelihood. The results in column (f) confirm that profitability chances improved with intuitive confidence, but clearly dismiss the T(INTRO) effect that was documented in the analysis of nominal returns. Column (g) exposes some additional effects (DATE, EDU and LOSS AVERSION) that did not surface in the analysis of percentile returns. While the increase in profitability rates as the

[p] Classification into METHOD and OCCUPATION are independent by chi-square tests.

[q] For example, when the 48 participants with no industry-experience and CONF ≥ 50 are median split by AGE, the mean return of the 22 younger (AGE < 30) is 10% compared to 0% to the 26 older participants.

experiment progressed and the higher success rates of more educated participants may be sample idiosyncratic,[r] the negative results for LOSS AVERSION seem intriguing, especially in light of the fact that RISK AVERSION did not exhibit parallel effect on returns or profitability rates (the correlation between RISK AVERSION and LOSS AVERSION was far from perfect $\rho = 0.38$). The 78 relatively loss-averse participants (those that preferred the risk-free alternative at least twice in the three loss-aversion problems) showed positive returns in 49% of the cases, compared to 63% $R > 0$ rate for the 52 other participants ($p < 0.05$), and the logistic estimations suggest that an increase of 1 standard deviation in LOSS AVERSION decreased profitability chances by 7.94% on average. Intuitively, the weaker performance of the relatively loss-averse could follow from less risky arbitrage positions. In Sec. 5.4, we use the time-series estimations to test if LOSS AVERSION consistently affected arbitrage styles.

4. Comparison of Risk Long versus Short

The large profits on the sell-side of the arbitrage (Sec. 3.2) raise the hypothesis that the participants sold relatively riskier stocks to "leverage the crisis". The current section tests the hypothesis directly by comparing the stocks selected for purchase to the stocks selected for short-sale in terms of risk characteristics for the arbitrage date. The list of risk factors examined below includes the standard BETA, SIZE and BTM (Fama and French, 1992) plus a momentum factor in the spirit of Carhart (1997). In addition, we compare the sectoral affiliation of stocks long versus short to check if the crisis-sensitive CONSTRUCTION and FINANCE sectors were more common on the sell-side. The next paragraphs introduce the definitions. A short discussion follows.

- BETA is calculated from the daily return series for the calendar year ending at the date where the arbitrage was submitted; market returns are measured by TASE index.
- SIZE is the market-value of the stock, converted to US dollars, at the arbitrage date.
- BTM denotes the BTM ratio for the latest quarter preceding the arbitrage by (atleast) three months. The minimal three-month gap is imposed to ensure that the ratios are publicly available at the arbitrage date.
- PREV_3MON denotes the return on the stock in the three months preceding the arbitrage date.

[r]The $R > 0$ rate for the 62 participants that submitted the arbitrage before August 2008 was 47% compared to 62% for the 68 later arrivals ($p < 0.05$). The $R > 0$ rate for the 46 participants with 15–17 EDU was 41% compared to 61% for the $N = 47$ with EDU $\leq$ 15 ($p < 0.05$).

We calculate BETA from daily return series to match the short horizon of the three-month arbitrage. Similar BETA coefficients (based on single year histories and daily return series) are reported on regular basis in the leading financial portals that were linked to the experiment site. The coefficients therefore represent accessible data that the participants could actually consult while devising their portfolios. SIZE and BTM were similarly defined, addressing the latest information that subjects could inspect while making the arbitrage decisions.[s] The short-run nature of the experiment also motivates the use of a relatively narrow window of three months to measure momentum returns. Intuitively, the participants could check the performance of particular stocks in preceding three months before making choices for the future three months interval.

The comparison of risk exposure long versus short is summarized in Table 3. The left panel runs the comparisons between-stock; e.g., comparing the 294

Table 3. Comparison of risk factors long versus short.[a]

	Between-stock comparison			Within-portfolio comparison		
	Long	Short	P-test	Long	Short	Rank-test
Risk factor						
BETA	1.21	1.34	$p < 0.01$	1.09	1.19	$p = 0.09$
SIZE	5738	2572	$p < 0.01$	5702	3000	$p < 0.01$
BTM	0.56	0.75	$p < 0.01$	0.57	0.74	$p < 0.01$
PREV_3MON	−11.2%	−20.0%	$p < 0.01$	−12.0%	−19.4%	$p < 0.01$
Trading-sector						
Construction and real-estate	16.0%	28.6%	$p < 0.01$	6.9%	24.3%	$p < 0.01$
Financial services	8.8%	12.3%	$p < 0.05$	4.7%	11.5%	$p < 0.01$
Holding and investments	14.6%	9.7%	$p < 0.05$	7.1%	8.1%	$p = 0.39$
Industry	40.5%	28.6%	$p < 0.01$	71.3%	38.9%	$p < 0.01$
Trading and services	19.7%	19.3%	$p = 0.22$	9.8%	15.8%	$p = 0.03$

[a]The split into sectors follows the formal categorization of TASE. ENERGY stocks are omitted. Samples sizes for between-stock comparisons are: 294 (long) versus 269 (short) for BETA, SIZE, PREV_3MON and sectoral affiliations; 293 versus 264 for BTM (a few cases where BTM < 0 are ignored). Deposits and loans are ignored in the weighting of SIZE/BTM/PREV_3MON at the portfolio level. Sample sizes for within-portfolio comparisons are $N = 130$ for BETA and sectoral affiliation; $N = 112$ for SIZE, PREV_3MON and $N = 111$ for BTM. The Pitman test is employed for between-stock comparisons while the Wilcoxon signed-rank test is used for within-portfolio comparisons.

[s] If a given stock was selected in distinct dates within separate portfolios, the BETA, SIZE and PREV_3MON associated with each purchase may be different — representing a possible shift in relative risks along the volatile times of the experiment. BTM was similarly updated on quarterly basis.

stock-purchases along the experiment to the 264 short-sales in terms of BETA for the arbitrage dates. The right-panel alternatively runs the comparisons within-portfolio, taking account of the specific amounts of purchase or sale within each arbitrage. The volume-weighted BETA, for example, is separately calculated for each side of the arbitrage assuming BETA = 0 for deposits and loans, and the paired differences are used to test equality at the portfolio level ($N = 130$). When weighting SIZE, BTM and PREV_3MON within portfolio, we ignore deposits and loans. The paired comparisons are therefore restricted to the 112 participants that did not invest or deposit the entire 1000 NIS arbitrage volume.[t]

While the stocks selected for sale appear more aggressive, in terms of significantly higher BETA for the arbitrage date, the differences decrease substantially when loans and deposits are taken into consideration. Individual portfolios are almost balanced in terms of market risk, with average volume-weighted BETA 1.09 "long" versus 1.19 "short" ($p = 0.09$). The participants could deliberately attempt to balance market risk long versus short following the textbook example connected to the instructions. BETA could also be the most familiar or salient risk-factor for many subjects that translated the recommendation to balance risk into BETA-balanced portfolios, partially neglecting other risk factors. Regressions with model-selection on the difference (weighted BETA long) minus (weighted BETA short) suggest that exposure to market risk decreased with LOSS AVERSION (see Web supplement E for selected regression results). The 78 relatively loss-averse participants, in particular, almost balanced their portfolios (mean volume-weighted BETA 1.09 long versus 1.05 short, $p = 0.3$), while the others sold relatively riskier stocks compared to the stocks purchased (weighted BETA 1.09 long versus 1.38 short; $p < 0.05$). The comparison suggests that the lower profitability rates of the loss-averse could follow from more balanced arbitrage positions; the hypothesis is formally tested at the next section.

The experimental positions, however, appear extremely unbalanced in terms of SIZE and BTM. The stocks selected for purchase were about twice larger, on average, from the stocks selected for sale ($p < 0.01$, in both levels of comparison) and the average BTM ratio on the sell-side was about 0.75 compared to 0.56 on the buy-side ($p < 0.01$). The experimental arbitrageurs therefore sold riskier stocks compared to the stocks that they purchased in the sense of Fama and French (1992) historical risk factors. About 71% of the portfolios ($N = 92$) sold stocks with

[t]The weighted long-side BTM, for example, of a portfolio that invests equal amounts in two stocks with BTMs 0.7 and 0.6 while investing some positive amount in DEPOSIT is 0.65. Comparisons are run both ways as each method compensates for limitations of the other. The between-stock comparisons ignore the volumes of purchases and sales. The within-portfolio comparisons of SIZE, BTM and PREV_3MON, on the other hand, are impossible for the portfolios with1000 NIS LOAN ($N = 11$) or 1000 DEPOSIT ($N = 7$). See Web supplement E for further details and discussion of robustness.

higher weighted BTM ratio compared to the stocks purchased; 58% ($N = 76$) sold smaller SIZE stocks compared to the stocks purchased. Similar conclusions emerge in comparison of the multiplier-related factors; e.g., the proportion of negative-EPS stocks in short-side sales (about 20%) was almost twice larger than the corresponding proportion for long-side purchases (11–12%, depending on calculation method).

The PREV_3MON line in Table 3 additionally suggests that the participants tended to sell stocks that decreased more steeply (or climbed more slowly) in the three months preceding the arbitrage date. The stocks selected for short-sale plunged, on average, by almost 20% in the quarter preceding DATE, while the stocks selected for purchase decreased by 11–12% ($p < 0.01$, in both levels of comparison). The portfolio level comparisons however expose strong heterogeneity, with 45% of the participants ($N = 59$) adopting contrarian positions in the sense of buying stocks with weaker three-month performance compared to the stocks sold. Similar results emerge when historical returns are measured for the 180 days preceding the arbitrage or uniformly measured for the last three months of 2007, independently of DATE.

The comparison of sectoral affiliations finally reveals that the CONSTRUCTION and FINANCIAL sectors, that plunged most drastically at the early stages of the crisis, were actively traded on both sides of the arbitrage. Almost 41% of the stocks selected for sale classified as "REAL-ESTATE AND CONSTRUCTION" or "FINANCIAL SERVICES" by TASE formal indexing. The proportion of construction and financial stocks on the buy-side, 25%, was significantly smaller, but still relatively large. Again, the differences slightly decrease when sectoral choice is weighted within portfolio (Table 3).[u]

5. Time-Series Estimations

5.1. *Definitions*

The four-factor model (Carhart, 1997) is commonly applied to evaluate the performance of investment professionals and test if experts are able to deliver significant alpha beyond loading on familiar risk factors (for comprehensive discussion see Cochrane, 2005). A slight adaptation of the model to the specific settings of the experiment is required since the participants could only trade a restricted list of stocks and arbitrage transactions were limited to three purchases and sales. The specific risk premia that the experimental traders could collect

[u] Buyers of construction and finance stocks were significantly more optimistic. Comparisons of the DATEs where specific CONSTRUCTION and FINANCE stocks were selected for purchase to the DATEs where the same stocks were selected for sale did not reveal consistent trends.

therefore do not coincide with the premia implied by standard calculations. The adapted model still opens the possibility to test directly if stronger decline of riskier stocks contributed to profitability and estimate the risk-adjusted ALPHA on the restricted arbitrages.

Since the portfolios were collected at distinct dates and arbitrage styles could strongly differ across participants, we run separate estimations for each portfolio. Since the arbitrages were run for a limited horizon of only three months, we build on daily return series.[v] The daily returns on each portfolio were regressed on corresponding MARKET, SIZE, BTM and MOMENTUM premia (detailed definitions follow). The number of observations in the daily regressions lied between 55 and 67 depending on the specific calendar dates. Following common practice in the literature, estimations were run by GMM using the Newey and West (1987) method to correct errors for violation of standard assumptions.[w] The estimated equation is presented next:

$$R_{j,t} = \alpha_j + \beta_{\mathrm{MARKET},j}{}^{*}(\mathrm{MARKET}_{j,t}) + \beta_{\mathrm{SIZE},j}{}^{*}(\mathrm{SIZE}_{j,t}) + \beta_{\mathrm{BTM},j}{}^{*}(\mathrm{BTM}_{j,t}) + \beta_{\mathrm{MOM},j}{}^{*}(\mathrm{MOMENTUM}_{j,t}).$$

$R_{j,t}$ is the return on portfolio j at day t, while $\mathrm{MARKET}_{j,t}$, $\mathrm{SIZE}_{j,t}$, $\mathrm{BTM}_{j,t}$, and $\mathrm{MOMENTUM}_{j,t}$ denote the respective daily risk-premia.[x] The coefficients $\beta_{i,j}$ then represent the loading of portfolio j on risk-premia i, while the intercept α_j represents the excess daily returns beyond loading on the four factors. The daily returns on the arbitrage and daily risk-premia were measured in percentile form. An estimated alpha of 0.03 would therefore represent 0.03% (three basis points) daily risk-adjusted return on the arbitrage.

In calculating the daily premia, we adopt a restricted approach that takes into account the specific settings of the experiment. Consider the MARKET risk first. The most risky portfolio in terms of exposure to stock market risk could be constructed by ranking the 220 available stocks by their historical BETA; buying the three stocks with highest BETA and selling the three stocks with smallest BETA (or vice versa). Let MARKET1 denote the return on the respective portfolio

[v] For interesting examples building on daily factor estimations see Coval *et al.* (2004) and Barber *et al.* (2007).
[w] The need for GMM clearly shows in standard tests of errors' structures: homoskedasticity is rejected for 43 portfolios by White's test; zero first-order (second-order) serial correlation is rejected for 22 (20) participants by Breusch–Godfrey test; normality of residuals is rejected for 35 portfolios using the Shapiro–Wilk test.
[x] Consistently with the approach of Sec. 4, we redefine the risk loading portfolios for each arbitrage DATE (this is illustrated at the next paragraph). The risk-premia for given calendar date t may therefore vary with the DATE where portfolio j was submitted, so we index the premia by j and t concurrently. The DATE dependent portfolios represent the premia that subjects could deliberately seek by sorting with respect to updated factors. The DATE-conditional approach, in addition, bypasses the complications that could arise due to varying risk premia and possible regime shifts along the crisis.

assuming the specific amounts of purchase or sale were derived from the relative market-capitalization (SIZE) of each stock at the arbitrage date.[y] If MARKET1 is used to measure the daily market premia and some participant followed a reversed crisis-prone strategy, selling the three stocks with highest BETA while buying the three stocks with lowest beta, then the time-series regression of daily arbitrage returns on daily market premia would expose negative coefficient -1, revealing the extreme crisis-prone position adopted by the participant with respect to market risk. Since MARKET1 however is an extreme benchmark and the participants could expose to market-risk less intensely, we remove the three extreme stocks on each side of the sorted TA220 list and repeat the construction again with the truncated list of 214 items. MARKET2 is used to represent (the returns on) the portfolio that buys the 4–6 most aggressive stocks and sells the 4–6 most defensive stocks in TA220. The exercise is repeated eight additional times to construct the portfolios MARKET3-MARKET10. The simple average of the daily returns on MARKET1-MARKET10 is used to approximate the returns that could be earned by exposing to market-risk in the particular environment of the experiment, using MARKET, with no adjunct indexing, to denote the respective return.

Similar methods were applied to calculate the SIZE, BTM and MOMENTUM (henceforth abbreviated to MOM) premia. To calculate the daily SIZE premia, we sort TA220 by SIZE (as defined in Sec. 4) in descending order and construct 10 disjoint portfolios SIZE1-SIZE10 that represent different degrees of exposure to SIZE risk. The average daily return on the 10 portfolios is assumed to represent the daily premia on SIZE that the arbitrageur could collect in our specific application.

The daily premia on BTM (or MOM) were similarly calculated by sorting TA220 (in ascending order) by the BTM (PREV_3MON) as defined in Sec. 4 and constructing 10 disjoint arbitrage portfolios BTM1–BTM10 (MOM1–MOM10) that represent decreasing levels of exposure to the corresponding factor. The average daily return on the 10 portfolios is used again to represent the BTM (MOM) premia in the specific settings of the experiment.

The use of exactly 10 portfolios to calculate each premia is clearly arbitrary. Our main results however are robust and reaffirm when the number of portfolios is increased or decreased, when market returns are approximated assuming construction and finance stocks on the buy-side versus industry and trading stocks on

[y] To skip unnecessary notation we use "MARKET1" to represent the arbitrage portfolio, the daily returns and even the three-month premia. The scope of each use should be clear from the context. The premia calculations are illustrated in Web Supplement F.1. The correlations in daily risk premia are discussed in supplement F.2. The strongest correlations emerge for MARKET and BTM (mean Pearson correlation $\rho = 0.49$; median $\rho = 0.37$; $p < 0.01$ by sign-test or signed-rank test on 130 portfolios). Robustness with respect to the premia calculation method is discussed in supplement F.3.

Table 4. Three-months risk premia (percentile form).[a]

	Market	Size	BTM	Momentum
Mean/Median	**−16**/**−23****	−0.00/0.02	**−8.0**/**−7.8****	0.003/−0.006
Standard deviation	**(19)**	(0.13)	**(5.8)**	(0.13)
Minimum/maximum	−37/51	−19/44	−20/4	−36/33
Proportion negative (of $N = 130$)	81%	42%	86%	52%
Sign-test significance	$p < 0.01$	$p < 0.05$	$p < 0.01$	N.S.

[a]Asterisks mark significant premia by Wilcoxon tests. The premia calculations are illustrated in Web supplement F.1.

the sell-side, when momentum returns are ignored and various other changes are made in the premia calculation method (see Web Supplement F).[z]

5.2. *Three-months risk premia*

Before running the time-series regressions, we briefly examine the three months MARKET, SIZE, BTM and MOM premia to assess the overall returns that could be made by exposing to familiar return-generating factors in the arbitrage. The three months premia were calculated using precisely the same formulas that were introduced for the daily premia. The three-month MARKET premium, for instance, is the average three-month return on the 10 portfolios MARKET1-MARKET10 as defined above. Table 4 summarizes the three-month premia distributions; the data presented in percentile form as basis points are inconvenient for large three-month figures.

Consider the MARKET premia first. The mean three-month market premium was negative −16.3% ($p < 0.01$) and more than six times larger — in absolute value — than the average return on the arbitrages (2.7%). The large negative premium reflects again the severe crisis conditions within which the experiment was run. The relatively aggressive stocks in terms of historical BETA decreased more rapidly than the defensive stocks along the crisis and the three-month premium was negative for 105 of 130 participants.

The results for SIZE are less conclusive. The mean three-month SIZE premium was close to zero and far from significance. SIZE was negative for 42% of the

[z]Since short-run momentum-based or contrarian strategies could play significant role in the hectic times of the experiment, it is interesting in particular to note that similar ALPHAs emerge in three-factor estimations. The three-factor mean daily ALPHA was 2.7 basis points, compared to 3.02 for the four-factor model ($p \approx 0.39$ by sign-test or signed-rank test on the 130 differences). Similarly, the hypothesis that the three-factor betas equal the four-factor betas could not be rejected for MARKET, SIZE and BTM.

portfolios, but the smaller SIZE stocks exhibited stronger resistance to the crisis in 58% of the cases. In particular, the three-month SIZE premia ranged from a minimum of −19% to a maximum of 44%, underling that SIZE-risk was not an obvious factor to exploit for profitable arbitrage. Recent empirical studies indeed question the persistence of the SIZE effect and contend that the effect could have diminished since the eighties (Van Dijk, 2011). Our current examination is limited in scope but it may still be interesting to note that SIZE shows extreme volatility in three months intervals around the crisis.

The consistently negative premia reemerge for BTM. The data in the third column of Table 4 reveals that the riskier stocks in terms of highest BTM ratios decreased more steeply along the experiment. The mean three-month BTM premium was −7.96%, implying that the participants could triple their arbitrage returns by taking more risky positions with respect to BTM. The premium was negative for 86% of the portfolios.[aa] Overall, the results for BTM appear most consistent amongst the four factors examined in Table 4, suggesting that crisis prone strategies that sell relatively risky growth stocks while investing in more solid value stocks could prove profitable throughout the experiment.

The results for MOMENTUM, on the contrary, are mixed again. The mean three-month MOM premium was close to zero. The premium, ranging from −36% to 33%, was positive in 48% of the cases and negative in the remaining 52%. The statistics thus clearly suggest that MOMENTUM was not an obvious factor to exploit for profitable arbitrage in the crisis period of the experiment.[ab]

5.3. *Estimation results*

The time-series estimation results are summarized in Table 5. The most consistent results are observed for BTM. The mean β_{BTM} coefficient was −0.11 while the median was −0.13 and the proportion of portfolios with $\beta_{\text{BTM}} < 0$ was 64%. The Wilcoxon and signed-rank tests strongly confirm that the participants loaded negatively on BTM-premia along the experiment ($p < 0.01$). The estimated β_{BTM} coefficient was negative and significant (at $p < 0.1$) for 32% of the portfolios while it was positive and significant for only 18%. Recall that the average three-month BTM premia was about −8%. Multiplying the mean premia by the mean β_{BTM} (−0.11) roughly suggests that negative loading on realized BTM-premia

[aa] The premium was positive but small in magnitude (maximum 4%) for 18 portfolios that were delivered in July. Closer examinations suggest that the reversal was caused by strong decrease of some of the largest stocks in TA220 around August–September 2008.

[ab] Cooper *et al.* (2004) use the 1–3 years lagged market returns to distinguish between UP and DOWN periods and show that momentum returns are positive and significant in UP states but turn negative and lose significance in DOWN states. The current examination deals with much shorter horizons. The noisy behavior of the three-month MOM premia may follow from specific trends around the crisis.

Table 5. GMM regression of daily returns on risk premia.

	α	β_{MARKET}	β_{SIZE}	β_{BTM}	β_{MOM}	R^2
Mean	0.03	−0.04	−0.09**	**−0.11****	0.03	0.34
Median	0.02	0.02	−0.06	**−0.13**	−0.01	0.30
(Standard diviation)	(0.4)	(0.60)	(0.52)	**(0.64)**	(0.59)	(0.22)
Proportion positive	55%	52%	38%	36%	47%	—
Sign-test significance	N.S	N.S	$p < 0.01$	$p < 0.01$	N.S	—
Coefficient positive and 10% GMM significant	12%	30%	21%	18%	25%	—
Coefficient negative and 10% GMM significant	6%	32%	24%	32%	20%	—

contributed almost 0.9% return to the arbitrages.[ac] The results for BTM therefore confirm that the experimental arbitrageurs benefited from taking crisis-prone positions or "leveraging on the crisis". The portfolio with lowest β_{BTM} (user 674; $\beta_{\text{BTM}} = -2.0$; $t = -2.7$), in particular earned 42% in only 66 trading days. More generally, the mean return on the 83 portfolios with $\beta_{\text{BTM}} < 0$ was 5.1% compared to −1.7% on 47 portfolios with $\beta_{\text{BTM}} > 0$ ($p = 0.07$).

The estimations also suggest that the participants loaded negatively on SIZE premia along the experiment. The mean (median) β_{SIZE} was −0.09 (−0.06) and individual loadings on SIZE were negative in 62% of the cases. Since the realized three-month SIZE premium was positive for 58% of the portfolios and the mean premium was close to 0%, the negative loading on SIZE did not contribute (on average) and even decreased profitability. The mean return on the 81 portfolios with $\beta_{\text{SIZE}} < 0$ was only 0.3% compared to mean return of 6.7% on other portfolios ($p = 0.09$). The differences increase when the comparison is restricted to participants with significant β_{SIZE} coefficient: the mean return on the 31 portfolios with significant $\beta_{\text{SIZE}} < 0$ was −7.3% compared to mean return of 4% on the 28 portfolios with significant $\beta_{\text{SIZE}} > 0$ ($p = 0.06$).

Mixed patterns of loading were also observed for MARKET. The mean β_{MARKET} coefficient was negative −0.04 but the median was positive 0.02. The proportion of negative coefficients was close to 50%, but a split of the sample interestingly reveals that the 63 portfolios with $\beta_{\text{MARKET}} < 0$ (mean return +12.6%) significantly outperformed the 67 portfolios with $\beta_{\text{MARKET}} > 0$ (mean return −6.6%). The differences again magnify when the comparison is restricted to portfolios with significant coefficient: the mean return on the 41 arbitrages with

[ac] The approximation is validated when individual β_{BTM} estimates are multiplied by corresponding three-month BTM premia and averaged across the sample. The alternative calculation suggests that loading on BTM contributed about 1.1% average return to the arbitrages. The approximation is validated for other factors as well. Estimation results for selected portfolios are presented in Web supplement F.4.

significant negative β_{MARKET} was 16% compared to mean return of −9% on the 39 portfolios with significant positive coefficient ($p < 0.01$).

Mixed patterns of loading are also observed for momentum. The proportion of positive β_{MOM} coefficients was close to 50% and the tests could not reject the hypothesis that $\beta_{\text{MOM}} = 0$ across the sample. The data moreover does not reveal consistent effect of β_{MOM} on eventual arbitrage returns. The mean return on the 61 portfolios with $\beta_{\text{MOM}} > 0$ was 4.6% compared to 1% return on the 69 portfolios with $\beta_{\text{MOM}} < 0$ ($p = 0.21$).

The estimations in conclusion suggest that the participants benefited from selling relatively riskier stocks in terms of BTM and MARKET risk while loadings on SIZE and MOMENTUM could either improve or damage performance depending on specific DATEs and arbitrage styles. The mean intercept ALPHA in the time-series regressions was 0.03, representing daily return of 3 basis points on the arbitrage, and the proportion of ALPHA > 0 was 55% ($p = 0.13$). When the average alpha is accumulated for 60 days (the median number of days in the time-series regressions), the product (about 1.8%) covers 2/3 of the gross arbitrage return (2.7%). ALPHA was positive and 10% GMM-significant in 15 cases (about 12% of the portfolios), compared to only eight cases where it was negative and significant. The Pearson correlation between ALPHA and gross payoffs was positive $\rho = 0.70$, suggesting that successful performance in nominal terms cannot be solely attributed to differences in arbitrage styles but also represents higher loadings-adjusted return. The mean daily ALPHA of the 45 participants with $R > +10\%$ was 0.39 compared to −0.32 for the 34 subjects with $R < -10\%$ ($p < 0.01$).

5.4. *Cross-sample comparisons*

Table 6 briefly reexamine some of the effects that were discussed in Sec. 3 in light of the time-series estimation results. Rows 1.1 and 1.2 compare the 68 participants

Table 6. Cross-sample comparisons of estimation results.

	Sub-sample	N	Return	α	β_{MARKET}	β_{SIZE}	β_{BTM}	β_{MOM}	R^2
1.1	CONF ≥ 60%	68	8.1%*	0.08*	−0.09	0.03	**−0.22****	0.13*	0.37
1.2	CONF < 60%	62	−3.2%	−0.02	0.03	**−0.22****	−0.00	−0.07*	0.30
2.1	T(INTRO) > 170.5	29	0.3%	0.01	−0.02	0.06	−0.11	0.14**	0.35
2.2	T(INTRO) ≤ 170.5	29	**10.5%****	0.03	−0.13	−0.04	**−0.36****	0.16	0.36
3.1	Age > 29	66	3.1%	−0.02	−0.05	−0.08	−0.06*	0.17*	0.32
3.2	Age ≤ 29	64	2.3%	0.08*	−0.02	−0.10**	−0.17**	−0.09*	0.35
4.1	Loss Averse −	52	5.3%**	0.01	**−0.20****	**−0.17****	−0.03	−0.02	0.33
4.2	Loss Averse +	78	0.9%	0.05	0.07*	−0.04	**−0.17****	0.07	0.34

with CONF $\geq 60\%$ to the less confident participants. The confident types load negatively on BTM while the less-confident load negatively on SIZE. The mean intercept in the regressions for confident participants was 0.077 compared to mean alpha of -0.02 for the less confident ($p < 0.01$). More generally, the Pearson correlation between CONF and ALPHA is positive 0.16 while the correlation between CONF and β_{BTM} is negative -0.14. The positive CONF effect on ALPHA also shows when ALPHA is regressed on the explanatory variables tested in Sec. 3. After removal of insignificant effects, the estimations robustly suggest that 10% increase in confidence increased alpha by 3–3.5 basis points.[ad]

Rows 2.1 and 2.2 examine the second variable that showed significant effect on performance by the preliminary analysis: the time spent on the introductory pages. A median split (of the sample with $60 \leq \text{CONF} \leq 80$) by T(INTRO) reveals that the participants that spent more time on the introductory material adopted more balanced arbitrage positions (mean $\beta_{\text{BTM}} - 0.36$ for the rapid compared to -0.11 for the relatively slow; $p < 0.06$), but the risk-adjusted returns (0.03 and 0.01, respectively) were not significantly different between the sub-samples. More generally, the correlation between T(INTRO) and β_{BTM} was positive $+0.32$, while the correlation between T(INTRO) and ALPHA was close to zero ($\rho = 0.05$). The weaker performance of participants with larger T(INTRO) thus seems to follow from less risky arbitrage styles and may represent more attentive response to the instructions.

In rows 3.1 and 3.2. we split the sample by AGE. The younger participants show stronger negative loading on BTM (mean $\beta_{\text{BTM}} = -0.17$ compared to -0.06 for the older), while the older load more aggressively on momentum. The ALPHAs of younger (mean 0.08) significantly exceed the ALPHAs of older (mean -0.02) and the coefficient of correlation between AGE and ALPHA for the complete sample is -0.12. The stronger performance of younger participants also shows in regression analysis when other moderators of ALPHA, such as confidence, are accounted (Web supplement F.5). The superior performance of younger participants, in conclusion, reflects in both marginally higher profitability rates (Sec. 3.6) and significantly larger risk-adjusted returns.

Rows 4.1 and 4.2 examine another variable that proved significant in the logistic regressions: individual LOSS AVERSION. The 52 participants that displayed lower levels of loss-aversion show negative β_{MARKET} coefficients averaging at -0.20, while the 78 relatively loss-averse show mild positive β_{MARKET} coefficients averaging at 0.07 ($p < 0.01$). The Pitman tests, however, could not reject

[ad]The results of model-selection regressions on ALPHA are presented in Web supplement F.5. Regressions on the estimated BETAs did not provide additional insights and levels of fit were very low. We therefore skip this line of analysis.

equality of other betas and ALPHA across the two sub-samples. More generally, the coefficient of correlation between LOSS AVERSION and β_{MARKET} is 0.19, while the correlations with other estimated coefficients are close to zero. The estimations therefore suggest that the stronger performance of less loss-averse follows from more risky crisis-prone positions but does not extend to risk-adjusted returns.

Finally note that the professionals and self-employed that exhibited superior performance in terms of gross returns also show positive significant ALPHA around 0.08 ($p < 0.04$ by sign-test and signed-rank test for the joint group), compared to insignificant average ALPHA around 0 for other participants. The risk-adjusted return of participants claiming to build on private-information was positive in 18 of 30 cases with mean 0.01 and median 0.06 ($p = 0.18$; see Web supplement F.6 for details).

6. Discussion

The experimental approach was applied to collect arbitrage portfolios from skilled investors and test arbitrage styles and risk-adjusted returns at the pick of the crisis. Stock-selection for the arbitrages was diverse and only 20 participants (15%) restricted their sales to crisis-sensitive construction and finance stocks. Stock-selection does not appear more focused for the successful subjects. While the 130 participants that constitute the complete sample selected 112 distinct stocks on the sell-side of the arbitrage (0.86 distinct stocks for participant), the 57 participants with return $> 5\%$ sold 65 different stocks from the list (1.14 distinct stocks for participant). Only 19 of the 39 stocks that were sold but never purchased by the arbitrageurs with return $> 5\%$ affiliated as CONSTRUCTION or FINANCE companies. If the strong performance of confident or professional participants can be ascribed to some specific risk-factor beyond the factors tested in the estimations, we could not characterize the underlying risk in general. Daily ALPHA in particular still appears positive (larger than 0.03 points for more than 50% of the portfolios) when MARKET premia are approximated assuming construction and financial stocks on the buy-side versus industry, trading and services stocks on the short-side (Supplement F.3).

Anecdotally, the experimental results suggest that "consulting the most confident experts" may prove profitable, at least with relatively young, highly motivated sample of consultants, of the type employed in the current design. Moreover, the results demonstrate that skilled stock selection may bring positive risk-adjusted alpha even in crisis tumbling markets (see, for instance, Ibbotson *et al.*, 2010 analysis of hedge fund performance along the crisis). Testing similar short-run arbitrage in positive or bullish markets could be interesting in light of the current

results. If the participants in the current experiment exploited the special subprime-crisis circumstance to sell the right stocks more frequently, such opportunities may be less evident in prosperous or stable markets.

Acknowledgment

We thank participants at the Erasmus-Technion Decision and Prediction conference 2012, ESA 2010 international meetings in Copenhagen, FUR XIV 2010 conference in Newcastle, and the 2009 conference on psychology and investments at the college of management, Israel for comments on various parts of this project. We benefited from communications with Ido Kallir, Uri BenZion, Nave Eshkenazi, Ido Erev, Shirly Farkas, Shmuel Hauser, Eran Regev, Mosi Rosenboim, Eyal Shalom and Yossi Shvimer. The study was supported by the Research Unit of the School of Business Administration at the College of Management Academic Studies, Israel. We thank the research unit for funding and thank an insightful referee for helpful comments and suggestions.

References

Barber, BM, R Lehavy and B Trueman (2007). Comparing the stock recommendation performance of investment banks and independent research firms. *Journal of Financial Economics*, 85(2), 490–517.

Beunza, D and D Stark (2004). Tools of the trade: The socio-technology of arbitrage in a Wall Street trading room. *Industrial and Corporate Change*, 13(2), 369–400.

Carhart, MM (1997). On persistence in mutual-funds performance. *Journal of Finance*, 52(1), 57–82.

Clark, J and L Friesen (2009). Overconfidence in forecast of own performance: An experimental study. *The Economic Journal*, 119(534), 229–251.

Cochrane, JH (2005). *Asset Pricing*, revised edition. Princeton University Press.

Cooper, MJ, RC Gutierrez and A Hameed (2004). Market states and momentum. *The Journal of Finance*, 59(3), 1345–1365.

Coval, J, D Hirshleifer and T Shumway (2004). Can individual investors beat the market? HBS Finance Working Paper No. 04-025.

Dubil, R (2004). *An Arbitrage Guide to Financial Markets*. Wiley Finance Series.

Eling, M (2009). Does hedge fund performance persist? Overview and new empirical evidence. *European Financial Management*, 15(2), 362–401.

Fama, EF (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2), 383–417.

Fama, EF and KR French (1992). The cross-section of expected stock returns. *Journal of Finance*, 47, 427–465.

Gatev, E, WM Goetzmman and KG Rouwenhorst (2006). Pairs trading: Performance of relative value arbitrage rule. *Review of Financial Studies*, 19(3), 797–827.

Gottesman, AA and MR Morey (2006). Manager education and mutual fund performance. *Journal of Empirical Finance*, 13(2), 145–182.

Harrison, WG and JA List (2004). Field experiments. *Journal of Economic Literature*, 42 (4), 1009–1055.
Hao, J, A Kalay and S Mayhew (2010). Ex-dividend arbitrage in option markets. *Review of Financial Studies*, 23(1), 271–303.
Hogan, S, R Jarrow, M Teo and M Warachka (2004). Testing market efficiency using statistical arbitrage with applications to momentum and value strategies. *Journal of Financial Economics*, 73(3), 525–565.
Ibbotson, R, P Chen and K Zhu (2010). The ABC's of hedge funds: Alphas, betas and costs. SSRN Working Paper 1581559.
Klayman, J, JB Soll, C González-Vallejo and S Barlas (1999). Overconfidence: It depends on how, what, and whom you ask. *Organizational Behavior and Human Decision Processes*, 79, 216–247.
Levy, H and M Sarnat (1994). *Capital Investments and Financial Decision*, 5th edition. Prentice Hall.
Lichtenstein, S, B Fischhoff and LD Phillips (1982). Calibration of probabilities: The state of the art to 1980. In D Kahneman, P Slovic and A Tversky (Eds.), *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge, UK: Cambridge Univ. Press.
Mitchell, M and T Pulvino (2001). Characteristics of risk and return in risk arbitrage. *The Journal of Finance*, 56(6), 2135–2175.
Newey, WK and KD West (1987). A simple positive definite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703–708.
Van Dijk, MA (2011). Is size dead? A review of the size effect in equity returns? *Journal of Banking and Finance*, 35(12), 3263–3274.
Whistler, M (2004). *Trading Pairs: Capturing Profits and Hedging Risk with Statistical Arbitrage Strategies*. John Wiley and sons.

Supplements to Short-Run Arbitrage in Crisis Markets

Supplement A: *TA100 and TA220 data*

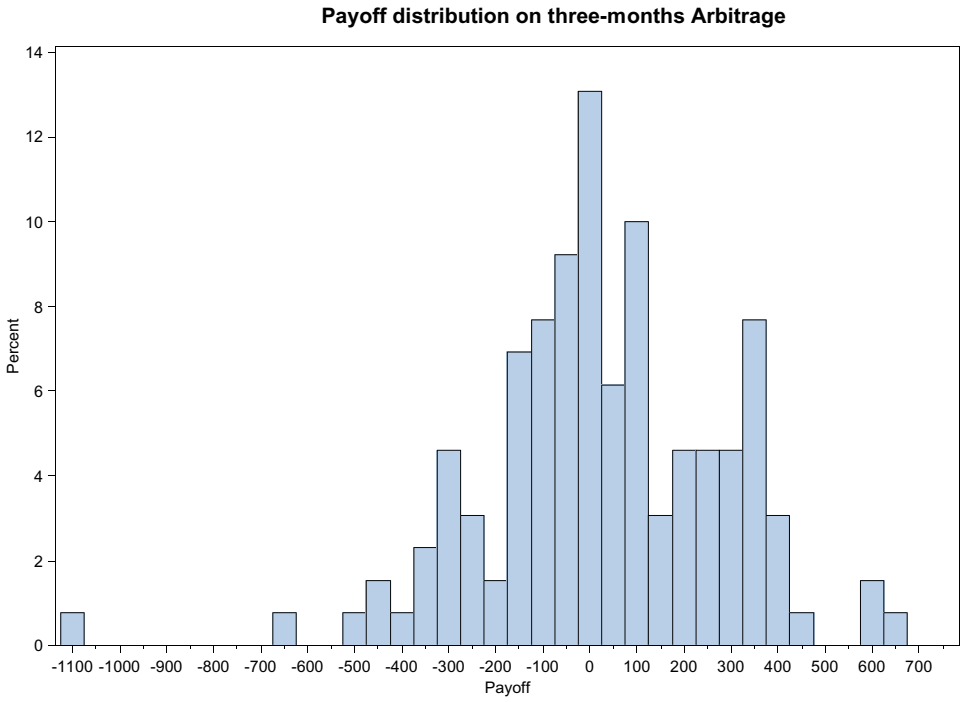
The table presents the monthly changes in TA100, TA220 and TASE from July 2007 until March 2009. The column "#Participants" provides the number of portfolios collected in each month and the column "Payoff" denotes the mean (median) payoff on the corresponding portfolios.

	TA100 (%)	TA220 (%)	TASE (%)	#Participants	Payoff
JUL 2007	−0.19	−0.55	0.18		
AUG 2007	−3.8	−4.00	−3.95		
SEP 2007	3.23	2.62	1.24		
OCT 2007	6.49	6.35	4.94		
NOV 2007	−3.66	−3.92	−4.01		
DEC 2007	1.9	1.21	1.27		
JAN 2008	−13.8	−13.8	−12.7		
FEB 2008	3.04	2.92	4.46		

(*Continued*)

	TA100 (%)	TA220 (%)	TASE (%)	#Participants	Payoff
MAR 2007	−10.2	−10.3	−9.60		
APR 2008	5.52	5.25	4.72		
MAY 2008	4.57	4.36	2.81	13	−17 (17)
JUN 2008	−4.81	−5.79	−3.18	30	30 (10)
JUL 2008	−3.99	−3.93	−4.68	19	−46 (−28)
AUG 2008	−3.14	−2.99	−2.28	6	270 (287)
SEP 2008	−11.6	−11.7	−11.0	20	59 (10)
OCT 2008	−19.7	−19.8	−17.5	13	89 (67)
NOV 2008	−8.17	−8.17	−5.83	21	−34 (12)
DEC 2008	−0.7	0.86	−1.92	8	60 (57)
JAN 2009	2.59	2.75	4.21		
FEB 2009	1.01	1.04	5.48		
MAR 2009	14.87	14.93	10.37		

Supplement B: *Payoff distribution*



Supplement C*: *Regressions on arbitrage payoffs

List of main explanatory variables tested in the regressions:

TA_220 — the change in TA220 index along the three-month arbitrage.
DATE — Participation date (in SAS numeric representation).
AGE
GENDER — 1 for male.
EXPERIENCE — years of experience holding a money-management or consulting job in the financial industry.
FAMILIARITY — subjective ranking of "familiarity with the Israeli stock market (prices, returns etc)" in 1–7 scale.
THEORY — subjective ranking of "theoretical knowledge in finance (academic courses, literature etc)" in 1–7 scale.
EDU — Years at school (including current academic year; 12 years represent standard high-school education).
MBA — Indicator for having or pursuing MBA degree.
HIGH_EDU — Indicator for EDU $\geq$ 16.
T(ARB) — Time spent on the three-month arbitrage screen (the screen where subjects submitted their transaction).
T(INTRO) — Time spent on the introductory pages 2 and 3 introducing the concept of arbitrage in expectations (we omit the other time counters that did no show significance).
TA100 — Expected return on TA100 in the three-month arbitrage period: **1** = increase of 10% or more **2** = increase of 5% to 10% **3**. Increase of 0 to 5% **4**. Decrease of less than 5% **5**. Decrease of 5% or more.
#LOGINS — Number of logins to complete the experiment: taking the values 1–4 (truncated at 4).
QUIZ — Number of correct answers minus number of wrong answers in the trivia familiarity quiz (unanswered questions are counted as 0).
CONF — subjective assessment of the probability that the selected arbitrage portfolio would earn a positive payoff.
RISK_AVERSE — Three binary choice problems are used to assess individual risk aversion in gains domain. In Problem 1 subjects are asked to choose between a lottery paying 200 or 50 NIS with equal 50% chances and a certain payoff of 155; in the second problem the certain alternative decreased to 125; in the third problem the certain alternative decreased to 95. RISK_AVERSE denotes the number of problems where the subject preferred the lottery over the certain payoff.
LOSS_AVERSE — Three binary choice problems are used to assess individual risk/loss aversion when losses are possible. In Problem 1 subjects are asked to

choose between a lottery paying +200 or −50 NIS with equal 50% chances and a certain payoff of 95; in the second problem the certain alternative decreased to 75; in the third problem the certain alternative decreased to 55. LOSS_AVERSE denotes the number of problems where the subject preferred the lottery over the certain payoff.

DEPOSIT — Amount invested in deposit.

LOAN — Amount of loan.

IND_DEPOSIT — Indicator for investing in deposit (the indicator were used in the logistics regressions while the quantities were used in the linear regressions; this improves the significance of the deposit/loan variables).

IND_LOAN — Indicator for borrowing funds by loan.

The next table presents the mean, median and standard deviation of each variable and summarizes the results of preliminary linear (logistic) regressions on the complete list of variables specified above. The maximal level of correlation between the explanatory variables included in the analysis is 0.65 — which is the correlation between "FAMILIARITY" and "THEORY". The minimal level of correlation = −0.26 is for "EXPERIENCE" and "CONF". The dependent variable in the linear regression is the return on the arbitrage in percentile form (R). The table also presents the results of the linear/logistic estimations after applying iterated backward removal of the insignificant coefficients up to the step where the all remaining variables are statistically significant at $p < 0.1$. In the logistic regressions, negative coefficients represent a decrease in the probability of positive arbitrage return.

Variable	Mean/median (STD)	Linear regression (LR)	LR with backward selection	Logistics regressions (LogR)	LogR with backward selection
TA_220	−16.5/−18.5	−0.03*		−0.07	
	(11.4)	(0.02)		(0.15)	
DATE	6/18 AUG	0.01*		0.006*	0.005*
	(65.5)	(0.04)		(0.004)	(0.003)
AGE	32/29	−0.17		0.02	
	(8.6)	(0.29)		(0.03)	
GENDER (male = *l*)	89% male	0.44		0.22	
		(6.7)		(0.69)	
EXPERIENCE	1.99/1	0.41		0.06	
	(3.2)	(0.8)		(0.09)	
FAMILIARITY	4.5/5	−0.33		−0.27	
	(1.3)	(2.3)		(0.25)	
THEORY	4.6/5	0.33		0.38*	
	(1.3)	(2.6)		(0.28)	

(*Continued*)

EDU	16.4/17	−0.67		−0.22**	−0.09**
	(1.9)	(1.0)		(0.11)	(0.04)
MBA	57%	0.2		0.37	
		(4.7)		(0.5)	
T(INTRO) (in seconds)	269/178	−0.014**	−0.016**	−0.0002	
	(258)	(0.008)	(0.007)	(−0.0009)	
T(ARB) (in seconds)	716/276	0.0002		0.0002	
	(2550)	(0.0009)		(0.00025)	
TA100	3.3/3	1.1		0.13	
	(1.4)	(1.7)		(0.16)	
# LOGINS	2.3/2	1.3		−0.07	
	(0.97)	(2.2)		(0.22)	
QUIZ	−0.7/−1	−0.8		−0.07	
	(1.55)	(1.4)		(0.15)	
CONF	54/60	**0.21****	**0.16****	0.01*	0.01*
	(25)	**(0.09)**	**(0.047)**	(0.009)	(0.0078)
RISK_AVERSION	1.3/1	−1.5		0.09	
	(0.6)	(3.7)		(0.4)	
LOSS_AVERSION	1.4/1	3.8		0.57	−0.54**
	(0.8)	(3.1)		(0.32)	(0.25)
IND(DEPOSIT)	18%	—		0.92**	0.033**
				(0.65)	(0.0019)
IND(LOAN)	16%	—		**−2.5****	**−0.003****
				(0.79)	**(0.001)**
DEPOSIT	85	0.021**	**0.023****	—	
	(236)	(0.01)	**(0.009)**		
LOAN	116	**−0.029****	**−0.0315****	—	
	(298)	**(0.075)**	**(0.007)**		
R^2	—	0.31	0.27	—	—

Supplement D: Detailed analysis of stock selection long versus short

About 63% of the stocks in TA220 (139 stocks) were selected for purchase or sale along the experiment. Stock selection was relatively focused on the buy side where subjects picked only 84 stocks, less than 40% of the menu. Larger diversity is observed on the sell side where the participants selected 112 distinct items from the list (51%). The more focused selection on the buy-side is underlined by tendency of participants to buy more stocks than they sell. The total number of stock-purchases (294) was larger than the number of short-sales (269) and only 19 participants sold more stocks than they bought.[a] Each of the 84 stocks in the

[a]The number of stocks selected (long and short) did not affect performance consistently.

purchase list was therefore selected 3.5 times (on average), while the 112 stocks in the short-sell list were only selected 2.4 times each.

When participants who sold the stock that attracted the largest volume of sale (6.4% of total sell-side volume) are removed from the sample, the mean arbitrage return increases to 3.5% and the median rises to 2.2% ($N = 115$; $p < 0.05$). When subjects that bought the stock that attracted the largest volume of purchase (7.7% of total buy-side volume) are excluded as well, the mean return increases further to 3.9% while the median climbs to 4% ($N = 96$; $p < 0.05$). Significance is maintained in various other sensitivity analyses; e.g., when portfolios that selected 1 of the 3 stocks that contributed the largest average gain or 1 of the 3 stocks that brought the heaviest average loss are ignored, the sample size moderately decreases to $N = 120$, the mean arbitrage return increases to 3.4% and the median rises slightly to 2%.[b]

To test if the accumulated volumes of purchase or sale could be used to identify purchase or sale opportunities, we examine the correlation between volumes of trading and eventual returns. First, we compute the correlation between net volume of purchase (total volume of purchase minus total volume of sale) and volume-weighted buy-side return, for the 84 stocks selected for purchase. The Pearson coefficient of correlation is positive 0.22 and the Spearman rank correlation is 0.32. The mean (buy-side) weighted-return on the 20 stocks that attracted the largest net-volume of purchase was -17.3% compared to mean weighted-return of -34.1% on the 20 stocks that drew the lowest net-volume of purchase ($p < 0.01$).[c] The stocks that attracted larger volume of net-purchase therefore showed stronger resistance to the crisis. The five stocks that drew the largest net volume of purchase, in particular, declined by a relatively modest rate of 8.8% along the corresponding arbitrages. Next, we run similar analysis on the 112 stocks that were selected for short-sale. The Pearson correlation between the net volume of sales (total sales minus total purchases) and weighted sell-side return is 0.09 and the Spearman rank correlation is 0.21. Again the data confirms that larger net sale volumes are associated with larger returns. The mean (sell-side) weighted-return on the 20 stocks that drew the largest net-volume of sale was $+34.5\%$, compared to mean return of $+17.2\%$ on the 20 stocks that obtained the lowest net-volume of sale ($p < 0.01$).

[b]The average return on each stock is separately calculated for the buy-side and sell-side. The average buy-side return on a given stock is the volume-weighted average of the returns earned by participants that purchased the stock. Sell-side returns are calculated similarly. The sample of 120 portfolios is obtained when portfolios that (bought 1 of the 3 strongest or 1 of the 3 weakest stocks in the buy-side) or (sold 1 of the 3 strongest or 1 of the 3 weakest stocks in the sell-side) are removed from the sample.

[c]The mean weighted-return for several stocks is calculated using a simple average, but the results do not change much when weighted averages (of weighted returns) are employed instead.

Table D.1. Separating stocks by type of transactions.[a]

	Long-side comparisons		Short-side comparisons	
	BUY-ONLY	BUY-and-SELL	BUY-and-SELL	SELL-ONLY
Complete sample ($N = 130$)				
Number of stocks	27	57	57	55
% (trading-volume)	11%	81%	60%	29%
Mean weighted-return	−18%	−23%	17%	18%
Mean % ($R > 0$)	25%	17%	78%	88%
Return > 0 ($N = 71$)				
Number of stocks	23	32	32	43
% (trading-volume)	36%	50%	51%	46%
Mean weighted-return	−10%	−19%	22%	29%
Mean % ($R > 0$)	34%	18%	84%	92%
Return $> 5\%$ ($N = 57$)				
Number of stocks	22	26	26	39
% (trading-volume)	41%	44%	48%	49%
Mean weighted-return	−5%	−17.5%	21%	32%
Mean % ($R > 0$)	37%	18%	80%	94%

[a]% (trading-volume) denotes the proportion of trading in each category (proportions do not sum to 100% because of loans and deposits). Mean weighted-return is the simple average of the volume-weighted average return on the stocks in each list. Mean % ($R > 0$) denotes the average ratio (volume of profitable transactions/total volume of transactions excluding deposits and loans) for the stocks in each category.

Table D.1 splits the collection of stocks selected along the experiment into three distinct categories. The "BUY-and-SELL" list includes the stocks that were traded on both sides of the arbitrage (bought and sold at least once by different participants). The "BUY-ONLY" list includes the stocks that were selected long but never selected short. The "SELL-ONLY" list similarly refers to the stock that were sold-short but never purchased. The table compares subjects' performance with the stocks that were only traded one-side to their performance with the stocks that were traded on both sides. The comparison is separately run for long-side (left-panel) and short-side (right panel) transactions. The upmost panel summarizes the results for the complete sample ($N = 130$). The % (trading-volume) figures interestingly reveal that most of the experimental trading (81% of buy-side volume and 60% of short-side volume) involved the 57 stocks composing the BUY-and-SELL list. Stock-selection along the experiment therefore did not separate the "stocks for purchase" from "stocks for sale"; only 20% of the trading was performed with the stocks composing the BUY-ONLY or SELL-ONLY lists. The stocks that were only selected on one side of the arbitrage exhibit stronger performance relatively to the

stocks that were traded on both sides, but most of the differences are insignificant. The mean volume-weighted loss on the 27 stocks composing the BUY-ONLY list, for example, was -18% compared to mean weighted-loss of -23% on (buy-side transactions with) the 57 stocks composing the BUY-AND-SELL list ($p = 0.18$; N.S.).

The relatively stronger performance of stocks that were only traded in one side of the arbitrage, however, strengthens when the analysis is restricted to the participants with positive eventual returns. The intermediate panel of Table D.1 discloses the results for the 71 subjects with $R > 0$. The volume of trading in stocks that were either sold or purchased but did not attract mixed trading (amongst these 71 participants) is 41%, compared to the 20% proportion for the complete sample. Moreover, the weighted-returns on the stocks that were selected one side only now significantly exceed the weighted-returns on the stocks composing the BUY-and-SELL list (see table).

Finally, at the lowest panel of Table D.1 we run the comparison again restricting the calculations to the 57 participants with arbitrage return larger than 5%. The results here are strongest: the mean weighted-return on the 39 stocks composing the SELL-ONLY list is 32% compared to mean weighted (sell-side) return of 21% on the 26 stocks composing the BUY-and-SELL list ($p < 0.02$). The mean weighted-loss on the 22 stocks composing the BUY-ONLY list was -5% compared to mean (buy-side) weighted-loss of -17.5% on the stocks traded on both sides of the arbitrage ($p = 0.05$). The large difference between the average SELL-ONLY returns (32%) and BUY-ONLY returns (-5%) suggests, in rough interpretation, that the collective selection patterns of profitable arbitrageurs could prove valuable to outside observers, but the sample size and length of the current experiment are insufficient for testing persistence and specific trailing strategies.

Finally, note that similar (slightly weaker) results emerged when transactions were classified as "BUY-ONLY" or "SELL-ONLY" when no opposite transactions were recorded within three months from the DATE of delivery. The comparisons for the 57 participants with $R > 5\%$ for example revealed mean weighted return of 30% in 53 cases where stocks were sold but never purchased within three months, compared to mean weighted return of 17% in 12 cases where opposite transactions occurred within three months ($p = 0.03$). The corresponding long-side results are: -11% average loss in 34 cases where stocks were selected long but never short-sold within three months, compared to -15% average loss in 14 cases where mixed-transactions occurred within 90 days (a given stock may classify into both categories when three-month screening is applied; in fact, none of the stocks appeared in both categories).

Supplement E: Comparison of risk factors long versus short (Sec. 4)

E.1. Complete version of the upper panel of Table 3-including MLP-related factors

	Between-stock comparison			Within-portfolio comparison		
Risk factor	Long	Short	*P*-test	Long	Short	Rank-test
BETA	1.21	1.34	$p < 0.01$	1.09	1.19	$p = 0.09$
SIZE	5738	2572	$p < 0.01$	5702	3000	$p < 0.01$
BTM	0.56	0.75	$p < 0.01$	0.57	0.74	$p < 0.01$
MLP	21.7	17.3	$p < 0.01$	22.1	17.9	$p < 0.05$
1[EPS < 0]	9.9%	17.1%	$p < 0.01$	10.1%	19.5%	$p < 0.01$

E.2. Measuring risk factors at the level of arbitrage portfolios

To compare risk-taking long versus short at the level of individual portfolios, we calculate a weighted average of the risk factors of stocks purchased and sold. The risk-factors on long-side purchases and short-side sales are weighted by the volume of purchase or sale correspondingly. Deposits and loans are ignored in the calculation of all weighted factors except BETA. In addition, we ignore stocks with negative book-equity in the comparison of BTM and stocks with non-positive EPS in the comparison of MLP. The weighting method is illustrated by a hypothetical example: Assume the participant purchased stock X in 600 NIS and deposited the remaining 400 NIS, while selling 600 NIS of stock Y and 400 NIS of stock Z. The 4 leftmost columns of the next table provide the assumed risk factors (at the arbitrage date) for each of the three selected stocks. The symbol "—" is used where the risk factor is undefined for the corresponding stock/deposit (e.g., MLP is undefined when EPS $\leq$ 0).

	Stock-level risk				Weighted risk	
	X	Deposit	Y	Z	Long	Short
Amount invested (sold)	600	400	(600)	(400)	1000	(1000)
BETA	1.5	0	1.5	1.2	0.9	1.38
SIZE	1000	—	2000	800	1000	1520
BTM	0.4	—	0.6	0.8	0.4	0.68
MLP	8	—	9	—	8	9
1[EPS < 0]	0	—	0	1	0	0.4
PREV (3MON)	−5%	—	−10%	−20%	−5%	−14%

The weighted BETA long (short) is directly calculated as a volume-weighted average of the betas of stocks purchased or sold assuming BETA = 0 for deposits and loans. Weighted BETA long (short) is accordingly equal to 0.6*1.5 + 0.4*0 = 0.9 (0.6*1.5 + 0.4*1.2 = 1.38). The weighted SIZE factor long is 1000 the size of the only stock purchased by the participant.[d] The weighted SIZE short is 0.6*2000 + 0.4*800 = 1520. The weighted BTM and PREV(3M0N) are similarly calculated but the weighting of MLP in short-side transactions ignores stock Z which is assumed to have negative EPS. Weighted MLP long therefore equals 8 (the MLP of stock X) while weighted MLP short equals 9 (the MLP of stock Y). The proportion of negative-EPS stocks in long side transactions is 0 while the proportion of negative EPS stocks in short side transactions is $400/1000 = 0.4$. (The short-side weighted $1[\text{EPS} < 0]$ would not change if the participants was borrowing 500 NIS; selling A in 300 NIS and selling B in 200 NIS, since loans and deposits are ignored throughout). The separate consideration of MLP in cases of positive EPS and 1[EPS] in cases where EPS $\leq$ 0 follows the approach of Fama and French (1992).

E.3. Robustness of results in Table 3 (upper panel)

BETA still appears to be more aggressive on the sell-side when measured from the daily return series for 2007, independently of the specific arbitrage date. The mean 2007-BETA of the stocks selected long was 1.14 compared to mean beta of 1.28 for the stocks sold short ($p < 0.01$); the weighting of 2007-BETAs at the portfolio level again lessens the differences (mean weighted betas 1.03 versus 1.12; $p = 0.09$). BETA also appears slightly more aggressive on the sell-side when measured contemporaneously along the three months arbitrage intervals but the differences diminish: the mean contemporary BETA on stocks selected long was 1.30 compared to mean BETA of 1.39 on stocks sold short ($p = 0.11$).[e] The tendency to sell smaller stocks still shows clearly when SIZE is measured from the market values at the end of 2007 (mean SIZE "long" 5901 versus mean SIZE "short" 2741). The uniform measurement of BTM ratios at the end of December 2007 similarly reconfirms the hypothesis that sell-side portfolios are significantly

[d] Weighting the SIZE of stock X by 0.6 (the proportion of total volume) is misleading. If SIZE is weighted by 0.6 the weighted SIZE factor long would be 600 while other participants that chose to buy the same stock X in 1000 NIS would show weighted SIZE long of 1000. The lower weighted SIZE of the former participant might suggest that risk-taking in terms of SIZE is stronger.

[e] The contemporaneous risk-factors (BETA for the three months arbitrage-interval and the BTM ratio from the most recent quarterly statements) were formally unavailable at the date where the arbitrage was submitted but may represent expectations regarding the updated levels of risk.

riskier with respect to BTM (mean BTM ratios: 0.47 long versus 0.59 short; $p < 0.01$). Relatively larger BTM ratios on the sell-side alternatively emerge when BTM is measured from the most recent quarterly statements (skipping the three months gap): the average buy-side BTM is 0.65 compared to average BTM of 0.87 on the sell-side ($p < 0.01$). (BTM ratios increased along 2008 because of the crisis). The tendency to buy stocks with stronger historical performance still shows when one week gap is enforced between the historical measurement and actual arbitrage date. The momentum effect is also established when histories are measured for a backwards-window of 6 months (instead of three months) or alternatively measured for the last quarter of 2007, independently of the specific arbitrage date (mean October–December 2007 return on the 294 stocks selected long 4.3% compared to mean return of -0.5% on the 269 stocks sold short; $p < 0.01$).

E.4. Regressions on net weighted risk-factors

The next four models were obtained by regressing the net weighted risk factors (long minus short) on the explanatory variables introduced in Supplement C and applying iterated backwards removal.

Model I: Variables that affect net weighted BETA (NW(BETA)) after removal of insignificant coefficients

$$\mathrm{NW(BETA)} = b\text{*LOSS\_AVERSE} + c\text{*FAMILIARITY} + d\text{*DEPOSIT} + e\text{*LOAN}.$$

Model II: Variables that affect net weighted SIZE (NW(SIZE)) after removal of insignificant coefficients

$$\mathrm{NW(SIZE)} = b\text{*FAMILIARITY} + c\text{*CONF} + d\text{*MBA} + e\text{*LOAN}.$$

Model III: Variables that affect net weighted BTM (NW(BTM)) after removal of insignificant coefficients

$$\mathrm{NW(BTM)} = a + b\text{*DATE} + c\text{*TA100} + d\text{*RISK\_AVERSE}.$$

Model IV: Variables that affect (NW(PREV 3MON)) after removal of insignificant coefficients (none of the variables shows significance at 5%)

$$\mathrm{NW(PREV\_3MON)} = a.$$

	Model I	Model II	Model III	Model IV
Dependent variable	NW(BETA)	NW(SIZE)	NW(BTM)	NW(PREV_3MON)
Intercept (a)			**1.1****	**6.9****
			(0.3)	**(2.1)**
LOSS_AVERSE	**0.20****	—		
	(0.05)			
RISK_AVERSE	—	—	−0.21**	
			(0.10)	
FAMILIAR	0.04**	−1125**		
	(0.02)	(400)		
LOAN	**0.0013****	**15.6****		
	(0.00015)	**(6.47)**		
DEPOSIT	**−0.0017****	—		
	(0.00019)			
CONF		**88.9****		
		(29.8)		
MBA		**4486****		
		(1725)		
TA100			**−0.14****	
			(0.05)	
DATE			**−0.002****	
			(0.001)	
R^2	0.60	0.21	0.15	0.00

Supplement F: GMM estimations (Sec. 5)

F.1. Example to the calculation of daily risk-premia

Consider the participant that submitted the arbitrage in MAY17.

To calculate the daily market-premia we first sort the stocks in TA220 in ascending order by their BETA for the arbitrage date MAY17.

The ids of the six stocks with largest and smallest BETA are disclosed at the next table where we also present the BETA (for MAY17), SIZE (for MAY17; rounded to nearest US$) and daily return in MAY18 for each of the six stocks.

Long side of MARKET1-Buying the three stocks with highest MAY17 BETA (among TA220)

Stock id	BETA(MAY17)	SIZE(MAY17)	MARKET1	MAY18 return
526012	3.14	113	+45%	−0.0142
434019	2.39	60	+24%	+0.0126
366013	2.34	78	+31%	−0.0168

Short side of MARKET1-Selling the three stocks with lowest MAY17 BETA (among TA220)				
Stock id	BETA(MAY17)	SIZE(MAY17)	MARKET1	MAY18 return
1082510	0.52	430	11%	−0.0238
1085182	0.52	27	1%	−0.0336
1101534	0.55	3358	88%	+0.00336

The relative SIZE of each stock determined the (fixed) weights of each stock in the MARKET1 arbitrage portfolio (see the column titled MARKET1). The MAY18 return on MARKET1 is then calculated directly as a SIZE-weighted average of the MAY18 returns on the stocks that composed the portfolio:

The weighted return on long-side purchases is:

$$0.45*(-0.0142\%) + 0.24*(0.0126\%) + 0.31*(-0.0168\%) = -0.086\%.$$

The weighted return on short-side sales (return signs are reversed for sell-side transactions) is:

$$0.11*(0.0238\%) + 0.01*(0.0336\%) + 0.88*(-0.00336\%) = -0.000\%.$$

The MAY18 weighted return on MARKET1 is therefore −0.086% (the sum of returns long and short).

The returns for the next trading days (MAY19, MAY20, etc.) are calculated similarly. Weekends and holidays where trade did not take place are ignored. The portfolios MARKET2–MARKET10 are similarly defined from the sorted TA220 list by BETA for MAY17. The (simple) average MAY18 return on the 10 portfolios MARKET1–MARKET10 is used to represent the daily market premia for the corresponding date. The daily premia for next few days are calculated similarly using the same fixed MARKET1–MARKET10 weights. The stocks in each portfolio and the weights however change with the specific DATE where each arbitrage was delivered.

F.2. Correlation in daily risk premia

Table F.2 examines the correlation in daily risk-premia along the experiment. The Pearson coefficients of correlation were separately calculated for each participant from the 55–67 daily observations. The mean, median and standard deviation are presented in the table.

F.3. Robustness of GMM estimation results

To discuss robustness, we outline four main conclusions that emerge from the GMM estimations: (a) portfolios load negatively on BTM and SIZE but loading

Table F.2. Correlations in daily risk-premia.

	SIZE	BTM	MOM
MARKET	**−0.39/−0.33****	**0.49/0.37****	**−0.31/−0.30****
	(0.22)	(0.26)	(0.46)
SIZE		−0.01/−0.14	0.11/0.23**
		(0.26)	(0.50)
BTM			**−0.23/−0.26****
			(0.24)

patterns are mixed for MARKET and MOMENTUM (b) positive average ALPHA; ALPHA > 0 for the majority of portfolios (c) positive 0.7 correlation between ALPHA and arbitrage returns (d) positive ALPHAs for confident participants, industry-professionals, and part of the portfolios built on technical methods. In the next paragraph, we briefly argue that these results are robust and reflect in various alternative specifications.

First, we test robustness with respect to the risk premia calculations. When the number of stocks in the risk-riding portfolios (MARKET1–MARKET10, SIZE1–SIZE10, etc.) is increased, for instance, from three to six stocks on each side of the arbitrage, the mean three-month BTM premia increases to −5.6% (compared to −8% in Table 4) reflecting the less extreme position of larger portfolios with respect to BTM risk. The average β_{BTM} in the alternative estimations is −0.20 (compared to −0.11); the coefficient is still negative for 74% of the portfolios. The average ALPHA in the estimations with six-stock portfolios is smaller 0.013 (compared to 0.03 above) but the median stays robust at 0.02. The average fit of the regressions does not change (mean $R^2 = 0.34$) but the left tail of the $R2$ distribution is longer revealing low fit levels for more portfolios.

Results (a)–(d) also sustain when the number of stocks in the risk-riding portfolios are decreased. When daily risk premia are calculated from five portfolios (BETA1–BETA5; SIZE1–SIZE5, etc.) instead of 10, β_{BTM} is negative and significant in 40 cases while it is positive and significant for only 21 portfolios. The risk-adjusted ALPHA is positive for 55% of the portfolios; the mean ALPHA for the 28 professionals is 0.07 compared to 0.02 for other participants.

To test if sectoral selection of stocks for purchase and sale captures the loading on MARKET-risk more effectively in the specific conditions of the experiment, we recalculate the MARKET premia assuming that only construction and financial-industry stocks are purchased within the MARKET1–MARKET10 portfolios while "INDUSTRY", "TRADING and SERVICES" and "INVESTMENT" stocks are sold-short (keeping the riskier sectors on the buy-side as in preceding definitions). The modified daily MARKET premia turns out strongly correlated with the

preceding premia measure (mean correlation level 0.96) and the average estimation results are very similar to the results of Table 5. The "sectoral" premia however captures the loading styles of several portfolios more effectively and ALPHA appears negative and statistically significant in only three cases compared to eight such cases in Table 5 (see Table F.3.1 for details). Opposite patterns emerge when daily risk premia are calculated directly from TA220 (using the daily changes in TA220 to calculate daily market premia, six standard SIZE/BTM sorted portfolios to calculate the SIZE and BTM premia and 30% of the TA220 list to calculate momentum returns): although the estimations results are basically similar, the revised premia are less effective in capturing the loading-styles of 47% of the portfolios (in terms of lower R^2) and reveal negative significant ALPHA in 13 cases compared to 8 such cases in Table 5 (see Table F.3.2). Results (a)–(d) also reemerged when we iteratively removed the factors that did not show significant effect on portfolio returns (at $p < 10\%$). The coefficient β_{BTM} survived the iterated removal in 79 cases; the coefficient was negative in 63% of these cases. The estimations with backward removal moreover revealed positive significant ALPHA for 17 portfolios compared to only 9 cases where the intercept was significantly negative.

The factor model was also run on weekly data (regressing the 13 weekly returns on each portfolios on the corresponding weekly risk-premia) to test the sensitivity of the results to the daily-level estimations. The results of the weekly estimations are somewhat noisier and less informative, possibly because of the small number of 13 observations per portfolio. The participants still appear to load negatively on BTM premia (mean β_{BTM} -0.03; median -0.15; $\beta_{\text{BTM}} < 0$ for 55% of the portfolios) but the negative loading on SIZE disappears (mean β_{SIZE} 0.02; median 0.01; $\beta_{\text{size}} < 0$ for only 49% of the portfolios). The mean intercept ALPHA by the weekly regression was 0.14 which appears compatible in magnitude with the 0.03 average ALPHA in the daily regressions, but the median was close to zero. ALPHA was positive and statistically significant for 15% of the portfolios while it was negative and significant in only 8% of the cases. An alternative method to address the high volatility of daily data is to add the one-day lagged premia to the list of explanatory variables in the regressions (as, for example, in Coval *et al.*, 2004. The results of estimations under this approach also appear consistent with conclusions (a)–(d) as outlined above. The intercept ALPHA was positive and 10% significant for 16 portfolios while it was negative and significant for only 10 portfolios. The correlation between ALPHA and arbitrage returns is almost 0.5; the mean β_{BTM} coefficients are -0.115 (for current premia) and -0.018 (for 1-day lagged premia).

In the estimations summarized next, MARKET premia was calculated assuming that only "REAL ESTATE and CONSTRUCTION" or "FINANCIAL SERVICES" stocks are purchased within the MARKET1–MARKET10 portfolios while only

"INDUSTRY", "TRADING and SERVICES" or "INVESTMENT" stocks are sold within these portfolios. Technically, the list of "REAL ESTATE and CONSTRUCTION" and "FINANCIAL SERVICES" stocks was sorted by BETA and the three stocks with highest BETAs were selected for purchase within MARKET1, the next three stocks were selected for MARKET2, etc. Similarly, the list of "INDUSTRY", "TRADING and SERVICES" and "INVESTMENT" stocks was sorted by BETA and the three stocks with lowest BETAs were selected for sale within MARKET1, the next three stocks were selected for sale in MARKET2, etc. The calculation of other premia was not modified. The mean (median) three-month MARKET premia with the revised definitions are −15.1% (−20.4%) and the correlation with the preceding MARKET premia (the one used for the estimations in Table 5) was 0.96. The correlation between ALPHA and arbitrage returns by these estimations is about 0.73.

Table F.3.1. Time-series estimation results with sectoral selection of stocks for calculating MARKET premia.

Time-series estimation results with revised MARKET premia ($N = 130$)						
	α	β_M	β_S	β_{BTM}	β_{MOM}	R^2
Mean	0.03	−0.05	−0.07*	**−0.12****	0.04	0.34
Median	0.03	0.06	−0.04	**−0.10**	−0.00	0.30
(Standard deviation)	(0.47)	(0.70)	(0.53)	**(0.67)**	(0.58)	(0.22)
Proportion positive	55%	52%	44%	38%	50%	—
Sign-test significance	N.S	N.S	$p < 0.1$	$p < 0.01$	N.S	—
Coefficient positive and 10% GMM significant	13%	39%	23%	20%	20%	—
Coefficient negative and 10% GMM significant	2%	37%	22%	27%	23%	—

In the estimations summarized next, daily MARKET premia is calculated by subtracting the daily experimental interest rate from TA200 daily return. To calculate the premia on SIZE and BTM, we split TA220 by SIZE and then sort each half by BTM to obtain six SIZE/BTM portfolios in the spirit of Fama and French (1993). The daily SIZE and BTM premia are then calculated in the standard manner. Following Carhart (1997) 30% criterion, we subtract the mean return on the 70 stocks (about 30% of TA220) with weakest historical performance in the 90 days preceding the arbitrage from the mean return on the 70 stocks with strongest performance in the 90 days preceding the arbitrage — to obtain the daily MOMENTM premia. The results of the estimations are summarized in the table. The average fit with the revised premia (0.36) is slightly higher than the average fit in the main-text regressions (0.34) but the left tail of the R^2 distribution is longer (e.g., worst decile 0.06 in the current specification compared to 0.08 in the preceding estimations). The correlation between ALPHA and arbitrage return is 0.62.

Table F.3.2. Time-series estimations using Fama and French (1993) six portfolios to calculate SIZE and BTM premia and Carhart (1997) 30% criterion to calculate MOMENTUM returns.

Time-series estimation results using Fama and French (1993) six portfolios to calculate SIZE, and BTM premia and Carhart (1997) 30% criterion to calculate MOM returns from recent 90 days histories ($N = 130$).

	A	β_M	β_S	β_{BTM}	β_{MOM}	R^2
Mean	0.02	0.01	−0.09*	**−0.21****	0.04	0.36
Median	0.03	−0.00	−0.15	−0.14	0.05	0.32
(Standard deviation)	(0.46)	(0.72)	(0.69)	(0.94)	(0.53)	(0.24)
Proportion positive	53%	49%	40%	43%	55%	—
Sign-test significance	N.S	N.S	$p < 0.02$	$p < 0.07$	N.S	—
Coefficient positive and 10% GMM significant	12%	28%	14%	19%	29%	—
Coefficient negative and 10% GMM significant	10%	30%	28%	34%	19%	—

F.4. Factor estimation results for selected portfolios

	Return	α	β_{MARKET}	β_{SIZE}	β_{BTM}	β_{MOM}	R^2
Lowest on BTM/MOM	+42%	0.9	**−1.3****	**0.4****	**−2.0**	**−2.0**	0.28
(user 674)		(1.1)	**(0.4)**	**(0.1)**	**(0.7)**	**(0.7)**	
Lowest on Size (user 369)	−1%	−0.4	−0.5	**−2.0****	**1.7****	−0.48**	0.47
		(0.6)	(0.4)	**(0.6)**	**(0.7)**	(0.26)	
Lowest on MARKET	+61%	−1.0	**−1.8****	−0.8*	−0.7	−0.07	0.68
(user 88)		(1.0)	**(0.3)**	(0.6)	(0.55)	(0.4)	
Second lowest on MOM	−43%	−0.2	0.6**	**−0.8****	0.8**	**−1.1****	0.75
(user 86)		(0.5)	(0.3)	**(0.2)**	(0.3)	**(0.4)**	
Highest alpha (user 569)	+64%	1.3**	**−1.1****	0.7*	0.0	0.3	0.76
		(0.7)	**(0.3)**	(0.5)	(0.4)	(0.4)	
Lowest alpha (user 445)	−28%	−0.8**	0.05	−0.2**	**0.95****	0.9	0.30
		(0.45)	(0.3)	(0.7)	**(0.35)**	(0.7)	

F.5. Regressions on estimated ALPHA

To detect the factors that affect the regression intercept ALPHA, we run regressions with model selection testing the explanatory variables that were introduced in supplementary Appendix C. The following table presents the final estimation results after removal of insignificant affects. To obtain model I we regressed ALPHA on all the explanatory variables that were introduced in appendix C except for DEPOSIT and LOAN. Only CONF and EDU showed significance. Model II

was obtained after including DEPOSIT and LOAN in the list of variables for model selection.

Regressions on loading coefficients did not contribute additional insights, we therefore omit the details.

	Model I	Model II
Dependent variable	ALPHA	ALPHA
CONF	**0.0034****	**0.003****
	(0.0014)	**(0.001)**
EDU	−0.0088**	—
	(0.0052)	
AGE		−0.004**
		(0.002)
Deposits		**0.0005****
		(0.0002)
Loans		−0.0002**
		(0.0001)
R^2	0.05	0.15

F.6. Factor estimation results by occupation and arbitrage method

Sub-sample	N	Return	α	β_{MARKET}	β_{SIZE}	β_{BTM}	β_{MOM}	R^2
Occupational categories								
Investment professionals	28	7.2%*	0.07*	0.07	**−0.21****	0.02	0.03	0.32
Self-employed	16	8.6%	0.08**	−0.20	−0.11	−0.17**	0.16**	0.34
Finance jobs	43	−1.7%	−0.05	−0.11	−0.03	−0.15*	−0.01	0.35
Other jobs	42	1.8%	0.07*	0.04	−0.06	−0.14**	0.03	0.34
Arbitrage method								
Private info	30	9.7%**	0.01	−0.05	−0.02	**−0.28****	0.13	0.28
Other methods	20	3.7%	−0.05	−0.13	**−0.27****	−0.12	0.12	0.31
Fundamental	51	−0.3%	−0.02	−0.08	−0.03	−0.12**	0.07	0.37
Technical	29	0.1%	**0.19****	0.12	**−0.15****	0.07	**−0.19****	0.35
Technicians with payoff > 0	14	**22.2%****	**0.52****	−0.03	−0.09	−0.10	−0.05	0.34
Technicians with payoff < 0	15	−20.5%**	−0.12**	0.25*	−0.20**	0.22*	−0.31**	0.36

Discussion of the results for subjects claiming to build the arbitrage by technical considerations: Interestingly, the estimations reveal positive significant intercept for the subjects claiming to build on technical methods. The mean ALPHA for this

group (0.19) seems puzzling in light of their relatively weak performance (mean gross return 0.1%). In the two bottom lines of the table, we separate the technicians with positive returns from those with negative payoffs. Positive ALPHAs (averaging at 0.52%) emerge for the 14 technicians with positive returns that in addition appear to load negatively on MARKET, SIZE and BTM. The results for the 15 technicians with $R < 0$, on the other hand, reveal negative intercept (mean -0.12) and positive loading on MARKET and BTM. The majority of these later participants (11 of 15) were optimistic, expecting the market to rise along their arbitrage.

Translated Script of the Experiment

Page 1

Welcome to the Arbitrage Experiment of the College of Management!
The experiment is run by researchers at the college of management in order to examine various hypotheses regarding the Israeli stock market and aspects of financial decision on stock purchase and sale.

Participation is not restricted in time — you may complete the experiment in your own pace with no disruptions.

First, we ask you to fill-in some personal details
Private Name:_____
Family Name:_____
I.D. number:_____
Email:_____
Age:_____
Gender: Male_____/Female_____

Education:
Years in school (including current academic year):_____
MBA or Student pursuing the MBA degree: Yes/No

Current Profession: (mark the most suitable option)

A. Professional job within the "financial investment" industry (consulting, investment management, brokerage, etc.)
B. Other professional job related to money management or finance
C. Other occupation. Please explain:_____
D. Unemployed

Experience:
Number of years where you held a professional job within the "financial investment" industry (consulting, investment management, brokerage, etc.).

Marital Condition:

A. Married with children
B. Married without children
C. Single
D. Other

Net Monthly Income:

A. Less than 4000 NIS
B. Between 4000 NIS and 8000 NIS
C. Between 8000 NIS and 12000 NIS
D. More then 12000 NIS

Please rank in 1–7 scale your theoretical knowledge in finance? (academic courses, financial literature, etc.)

1 — minimal knowledge
7 — real expert

_____

Please rank in 1–7 scale your familiarity with the Israeli financial market? (prices, returns, companies, etc.)

1 — minimal knowledge
7 — real expert

_____

In my opinion, in the next three months the TA100 index would:

A. Rise by 10% or more
B. Rise by more than 5% but less than 10%
C. Rise by less than 5%
D. Decrease by no more than 5%
E. Decrease by more than 5%

Page 2

Arbitrage Transactions: Synopsis

In the experiment, you would be requested to buy and sell simultaneously stocks that are traded in the Tel-Aviv stock exchange in total amount of 1000 NIS.

Using common professional terminology, our main experimental task is an "arbitrage transaction": parallel purchase and sale of securities intended to produce positive return with minimal risk.

The purchase of stocks within the arbitrage transaction is standard: The arbitrageur is supposed to buy stocks that should subsequently increase in order to profit from the holding.

The sale of stocks within the arbitrage transaction however is a "short sale": The arbitrageur sells stocks that she does not currently hold, intending to cover the short subsequently — by purchasing the stocks that she sold. If possible, the arbitrageur would clearly prefer to short stocks that would decrease subsequently in order to gain the difference between the expensive selling price and the lower buying price.

Here is a simple example to profitable arbitrage: Suppose the investor has purchased in 1.1.2007 some stock A at the amount of 1000 NIS and — in parallel — has sold stock B in similar amount. The investor has closed the arbitrage transaction at 9.30.2007. Assume that the price of stock A has increased to 1200 while the price of stock B has decreased to 950 by this date. The arbitrage in this scenario turned out quite profitable. The investor has purchased a stock that gained 200 NIS while shorting a stock that lost 50 NIS. The total arbitrage gains are therefore 250: 1200 that the investor collects from selling A minus 950 paid for purchasing stock B. The cash flow from the arbitrage is summarized in the following table (positive figures refer to revenues from selling stocks while negative figures represent payments for purchasing stocks):

	1/1/2007	9/30/2007
Stock A (purchase)	−1000	+1200
Stock B (short sale)	+1000	−950
Net revenue	0	+250

Clearly, it is easy to imagine other scenarios where the arbitrage would actually fail. If for instance the price of stock A went down from 1000 to 900 while the price of stock B increased from 1000 to 1150 then the arbitrageur would lose an amount of 250: the 900 NIS that would be received in selling stock A would not cover the

amount that should be paid to "cover the short" and purchase stock B (1150). The next table summarizes the outcome of the arbitrage in this alternative scenario:

	1/1/2007	9/30/2007
Stock A (purchase)	−1000	+900
Stock B (short sale)	+1000	−1150
Net revenue	0	−250

Page 3

The Rational Behind Such Arbitrage Transactions

Comment: This page discussed very briefly some aspects of risk management in arbitrage transactions. In case you are familiar with the topic and short in participation time — you may skip this material and continue to the next page immediately.

Arbitrage transactions are especially common today in the hedging-funds industry.

Hedging funds enjoy tremendous gains from successful arbitrage:
http://optionarmageddon.blogspot.com/2008/01/other-paulson.html

But also suffer immense losses when their arbitrage gambles fail:
http://www.post-gazette.com/pg/06263/723449-28.stm

What is the rational for such cyclic transactions?

Assuming that the risk level of stocks purchased within the arbitrage is similar to the risk-level of the stocks sold short, the cyclic transactions protects the trader from positive or negative market trends.

Assume for example that the arbitrage-trader (ARB henceforth) has purchased stock A while shorting stock B.

Assume the two stocks are "similar"; e.g., two moderate size textile-industry stocks (just for illustration).

Obviously, ARB believes that A is traded in more attractive prices than B.

Consider the outcome of the arbitrage in two difference scenarios:

Case A:

Suppose the textile index rises along the arbitrage period. ARB's expectations turn true and stock A rises more steeply than stock B. Assume specifically that A gains 300 NIS while B increases by only 100 NIS. In this case, ARB gains 300 on holding A while losing 100 in closing the short-position on B. The net gain from the arbitrage is therefore 200.

Case B:

Suppose alternatively the textile index declines along the arbitrage period. ARB's expectations however regarding the under-pricing of stock A turn true and stock A decreases less steeply than stock B. Assume specifically that A losses 100 NIS while B decreases by 300 NIS. In this case, ARB losses 100 on holding A while gaining 300 when closing the short-position on B. The net gain from the arbitrage is therefore 200.

We have demonstrated that the purchase and sale of similar stocks (e.g., similar stocks from the same industry) protects the trader from industry specific trends; the trader enjoys the profit from buying the better stock while selling the worse stock — independently of specific market conditions!

When the stocks purchased in the arbitrage are riskier or less risky than the stocks sold, it is possible to adjust risk levels by investing funds in deposits or borrowing funds in risk-free interest. If stock A for instance is more risky than stock B, we may sell 1000 NIS of stock B, purchase A at the amount of 600 and invest the remaining 400 in risk-free deposit, to "adjust risk levels".

Interested in more details? See the discussion in William Goetzmann Internet chapter:
http://viking.som.yale.edu/will/finman540/classnotes/class6.html.

Page 4

Outline of the Experiment

{We omit the paragraphs dealing with the long-run arbitrage which will be discussed in a subsequent paper}.

In the experiment, you will be asked to submit two arbitrage portfolios on leading stocks from the Israeli market.

The first arbitrage portfolio is for three months. That is, the arbitrage "positions" that you would submit along the experiment would be closed after three months (approximately) to calculate the gains or losses from the arbitrage. Clearly, in this case, you would prefer to purchase stocks that seem as a relatively attractive investment for the next quarter while shorting stocks that seem less attractive for such short-run investment.

The experiment organizers would calculate the gains/losses from the arbitrage (as illustrated above) and the results would be reported at the of arbitrage period. The experimental process would therefore end in more than one year from today. Still,

the results for the short-run arbitrage for three months would be reported about three months after closing the experiment.

The reports that would be distributed at the end of each phase would let you compare your won payoff with the payoff of all other participants. Participants would be only identified by id numbers to preserve anonymity.

In addition, we would pay to 100 of the participants in the experiment "real" monetary prizes. The identity of real-prize winners would be randomly determined using random number generation. The payoff, in check, would be sent to the address that you would provide in the last page of the experiment. The actual payment would increase with the gains from the arbitrage. Participants that ended-up with larger arbitrage gains (or smaller losses) would receive larger payments than participants with weaker results. The id numbers of the participants that were selected for real-payment and the amount paid to each of these participants would be disclosed by email at the end of the corresponding phase of the experiment. More details regarding the payment method would be disclosed in subsequent pages.

Please note that we guarantee your anonymity! Your private name, family name or other identifying details (except your id number) would not be published in the emails that we distribute or any other public reports regarding the experiment.

The fairness of the experimental procedure and the payment method is guaranteed!

Upon request, we will let subjects examine the complete result files at the end of the experiment — to let you control the process. The experiment is run for academic research. We ask for your serious candid cooperation!

Page 5

More Details on the Experimental Arbitrage

The next table presents the specific format that would be used to collect your arbitrage portfolios:

Stock Purchase			**Stock Sale**	
Stock Name	Amount Purchased		Stock Name	Amount Sold
Total	**1000**		**Total**	**1000**

* You may choose stocks from the TA100 list (the 100 largest stocks in the market) or the TA YETER120 list (the next 120 largest stocks in the market). The list of 220 available stocks would be presented — together with links to updated information regarding these stocks — in the page where you would be asked to deliver your arbitrage.

* In addition to selecting stocks from the list, you would be permitted to deposit funds or borrow funds in fixed three-month interest rate of 2%. In case you decide to deposit funds in such risk-free interest for three-month, write "DEPOSIT" instead of providing a specific stock name in the stock purchase column. In case you decide to borrow funds in the risk-free interest-rate, write "LOAN" in the stock sale column.

* The number of stocks that may be purchased in the arbitrage and the number of stocks that may be sold in the arbitrage is limited to 3. The arbitrage table therefore includes only three lines in each panel. Deposits and loans however are considered stock purchase and sale correspondingly. The maximal number of stocks purchased (or sold) therefore decreases to 2 in case you decide to deposit (or borrow) funds in risk-free interest.

* The next page provides some arbitrary example to the type of arbitrage portfolios that may be submitted.

Page 6

Examples

This page presents three examples to possible arbitrage transactions. The stocks for each transaction were selected in alphabetic order from the list of stocks traded in the Tel-Aviv exchange.

Example 1:

Stock Purchase			**Stock Sale**	
Stock Name	Amount Purchased		Stock Name	Amount Sold
AAORA	**1000**		**ABGOL**	**1000**
Total	**1000**		**Total**	**1000**

The investor in example 1 believes that AAORA is traded in attractive prices while ABGOL is overpriced (we ignore the risk aspects for these technical examples).

The investor therefore constructs a simple arbitrage portfolio with only one stock in each side of the table. Similarly, it is possible to construct other portfolios where the investor buys 1–3 stocks from the list and sells 1–3 stocks from the list — with no investing in deposit or borrowing by loans. The number of stocks purchased in the arbitrage may be different from the number of stocks sold. You are permitted, for example, to buy 1 stock while selling three stocks, etc.

Example 2:

Stock Purchase			**Stock Sale**	
Stock Name	Amount Purchased		Stock Name	Amount Sold
AVNER Y	**600**		**LOAN**	**1000**
IGOD	**400**			
Total	**1000**		**Total**	**1000**

In example 2, the investor believes that AVNER Y and IGOD are attractive for purchase but she hesitates on choosing stocks suitable for sale. She therefore chooses to borrow the arbitrage funds (1000 NIS) in risk-free interest rate instead of selling stocks short. If the investment in AVNER Y and IGOD indeed turn profitable, we would subtract the interest on the loan of 1000 (2%) to calculate the final gain from the arbitrage.

Example 3:

Stock Purchase			**Stock Sale**	
Stock Name	Amount Purchased		Stock Name	Amount Sold
AGARI	**500**		**ELDAR**	**100**
EDGAR	**200**		**O.R.T**	**400**
DEPOSIT	**300**		**AUDIOCODES**	**500**
Total	**1000**		**Total**	**1000**

In example 3, the investor believes that ELDAR, O.R.T, and AUDIOCODES are relatively "weak" stocks compared to AGARI and EDGAR and therefore sells the first 3 stocks while buying the last 2 stocks in the proportions provided in the table. The investor also chooses to invest 300 NIS in risk-free interest to decrease the risk on long-side transactions.

Page 7

More Details on the Experimental Method

In reality, arbitrage transactions are clearly more complicated. The purchase of given stocks and short sale of other stocks might affect market prices and transaction costs arise and may turn the gain from skillful selection of stocks long and short into net losses. To import this complexity into the experiment, we take the following measures:

(1) The buying or selling price of the stocks that you selected would be determined by the market price 1–5 days after the date where your arbitrage transactions was delivered. The exact date would be randomly determined by the program of the experiment. For example, if your portfolio was sent in 4/4/2008 then your buying and selling prices would be randomly selected for one of the days between 4/5/2008 and 4/9/2008.
(2) The exact arbitrage period would also be randomly determined. The arbitrage positions (long and short) would be closed after 87–93 days — where the exact length would be randomly determined for each participant.
(3) An amount of 20 NIS would be subtracted from your arbitrage gain (or added to your arbitrage loss) to represent transaction costs and fees.

Page 8

Actual Payment Method

At the end of the experiment, we would randomly select 100 participants (henceforth: "the selected winners") that would receive an actual payment depending on the results of their arbitrage.

Each selected winner would be randomly assigned to one of the two major experimental tasks: the three months arbitrage or the long-run arbitrage (henceforth: "the selected task").

The payment to "selected winners" would depend on the result of the selected task. The actual payment may even exceed 200 NIS!

The complete list of participants marking the 100 "selected winners" and the type of "selected task" for each selected winner would be distributed by email at the end of the experiment — to let you control the experimental process!

The selected winners would be identified by their id numbers to preserve anonymity.

The exact payment method is disclosed in the box below:

> If the selected task is the three months arbitrage, then we multiply the gain or loss from the arbitrage by 4 (to annualize the payoffs) and add or subtract the product from an initial balance of 200 NIS. Your actual payout would be equal to 0.5 of the resulting balance.
>
> For example, if your arbitrage has produced a gain of 21 NIS, your final balance is 284 (200 + 21*4) so your check would be 142 NIS. If, on the other hand, your arbitrage was closed in a loss of 15 NIS, then your final balance is 200 − 4*15 = 140, so your payoff for participation would be 70 NIS.

Questions to the experiment-organizers?
Don't hesitate. Email NISSOY@GMAIL.COM

Page 9

Final Remarks Before Experimental Assignments

* You may retrieve detailed information on each of the stocks in the following sites:

> The Tel-Aviv Stock Exchange
>
> **(link to http://www.tase.co.il/TASE/MarketData/Stocks/MarketData/)**
>
> Bizportal
>
> **(link to http://www.bizportal.co.il/shukhahon/sh_rezef.shtml)**
>
> The Marker
>
> **(link to http://finance.themarker.com/)**

* The links to the three databases would be also available at the main page where you submit the arbitrage portfolio.

* The experimental results (anonymous; identifying subjects only through their social security numbers) would be distributed by email at the end of each phase. The three month arbitrage results would be distributed approximately three months after we close the experiment for further participation. The email would provide details on the specific arbitrage selected by each participant; the gain or loss from the selected arbitrage portfolio and the actual payment to "selected winners".

* We can now proceed to the experiment itself! The experiment consists of four pages. The two first two pages are introductory task where we first attempt to characterize your level of familiarity with the Israeli market and then try to characterize your risk preferences. The final two pages ask you to submit your arbitrage portfolios.

* Good luck!

Task 1: Familiarity Quiz

This page presents three multiple-choice problems regarding the performance of the Israeli stock market over the last calendar year.

In each problem, please mark the option that seems most appropriate.

You have 3 min to complete the assignment.

The program would automatically proceed to the next phase after 180 s.

{The three problems were randomly selected by the experiment program for each participant. Each problem was randomly selected from a pool of 25 problems of similar level of difficulty; the first problem was meant to be easiest and the last one most difficult. We present one of the problems in the easiest category for illustration:

Which of the following indices earned the lowest return in 2007?

1. TA DIV20
2. TA TECH 15
3. TA100
4. TA25

Task 2: Personal Risk Preferences

In this page, we ask you to mark your choices (A or B) in six distinct binary choice problems.

In each problem, option A represents a binary lottery — paying 2 distinct prizes with equal chances 50% for each prize. In some cases, one of the prizes is negative so the lottery may result in a loss rather than paying a positive prize. Option B, on the other hand, always represents a given (certain) amount. We ask you to choose between option A (the lottery) and option B (the certain amount), in each of the six problems.

Please submit your candid choices!
Clearly, the choice between lotteries and risk-less prizes depends on the personal tastes of the decision-maker and does not admit a unique solution.

Actual Payment: At the end of the experiment, we will randomly select six participants — one participant for each of the six problems and pay each of these participants a cash prize according to the choices he made in the corresponding problem. More details on the exact payment method can be found in the next link.

Link for more explanations the payment method

Your actual payment would be derived from your choice between options A and B. If you selected option A, we would randomly draw the outcome of the lottery and pay you the corresponding outcome + 60 NIS. If you selected option B, your payoff would be equal to the amount stipulated in option B + 60 NIS. Consider for instance the first choice problem 1. If you have selected option A and the outcome of the lottery (which would be randomly determined by the computer) would be 200, then you would receive 260 NIS. If on the other hand, the outcome of the lottery would be 50, then your check will be for 110 NIS. If however you selected option B, your payment would be 215 NIS. Please note again that only 6 participants (one participant for each problem) – would actually be paid for the choice assignments.

Choice Problem 1:

• Option A1: (9%)

Probability	NIS
50%	200
50%	50

• Option B1: (91%)

Probability	NIS
100%	165

Choice Problem 2:

• Option A2: (30%)

Probability	NIS
50%	200
50%	50

• Option B2: (70%)

Probability	NIS
100%	125

Choice Problem 3:

• Option 3A: (93%)

Probability	NIS
50%	200
50%	50

• Option 3B: (7%)

Probability	NIS
100%	85

Choice Problem 4:

• Option A4: (12%)

Probability	NIS
50%	200
50%	−50

• Option B4: (78%)

Probability	NIS
100%	105

Choice Problem 5:

• Option A5: (37.5%)

Probability	NIS
50%	200
50%	−50

• Option B5: (62.5%)

Probability	NIS
100%	75

Choice Problem 6:

• Option A6: (90%)

Probability	NIS
50%	200
50%	−50

• Option B6: (10%)

Probability	NIS
100%	45

Task 3: Three Months Arbitrage

You may now select the arbitrage portfolio for three months:
The following databases may help your decision:

The Tel-Aviv Stock Exchange
(link to http://www.tase.co.il/TASE/MarketData/Stocks/MarketData/)
Bizportal
(link to http://www.bizportal.co.il/shukhahon/sh rezef.shtml)
The Marker
(link to http://finance.themarker.com/)

* To reexamine the list of 220 stocks that can be selected for the arbitrage click here:

List of Available Stocks

* Please recall that you may also deposit amounts in risk-free interest rate of 2% for the next three months. To invest funds in risk-free interest select "DEPOSIT" in the "Stock Purchase" panel in the table below. Alternatively, you may borrow funds at the risk-free interest rate of 2%. To borrow funds please select "LOAN" in the "Stock Sell" panel in the table below.

* Please submit you transaction through the following table as demonstrated in preceding examples.

To Reexamine the Examples Click Here

Stock Purchase			**Stock Sale**	
Stock Name	Amount Purchased		Stock Name	Amount Sold
Total	**1000**		**Total**	**1000**

* Please recall that you may logoff the experiment and reenter later successively with no limitations.

* Questions to the experiment organizers?

Email NISSOY@GMAIL.COM and we'll get back to you asap.

Final Page

Thank you for taking part in the arbitrage experiment of the college of Management!

Few final questions to conclude the experiment:

* Could you specify the criteria that you have used to select stocks for the three-month arbitrage? (mark the most appropriate box)

_____ Fundamental (economic) analysis
_____ Technical analysis
_____ Private information
_____ others?
Please explain _____

* Please estimate the chances that the arbitrage portfolio that you have constructed would generate positive payoff:

The chances — in my opinion — that the three-month arbitrage would earn positive return are_____%

* Please provide us a valid address for the next 14 months — for distributing the cash prizes.

Name:_____

Street and number:_____

City:_____

Zip code:_____

* Please write down our email address for future inquires:

NISS0Y@GMAIL.COM

Thanks again for your cooperation!