Internet Appendix to "Who Gambles In The Stock Market?"*

In this appendix, I present background material and results from additional tests to further support the main results reported in the paper.

A. Profile of Casino Players

While the demographic characteristics of lottery players exhibit stable and well-defined patterns, other forms of gambling such as casino gambling do not show such distinct patterns. Casino gamblers consist of a growing number of recreational players who play occasionally and a stable group of regular players (e.g., Hinch and Walker (2005)). The characteristics of regular casino players are similar to regular lottery players, but recreational casino players have an opposite profile (e.g., Campbell and Ponting (1984)). In particular, recreational casino gamblers are mostly tourists, belong to a higher socioeconomic group, are older (age between 51 and 65), and are more likely to be female, hold professional jobs, and earn higher incomes. Since recreational and regular casino gamblers have opposite characteristics and the size of the recreational group is growing, a distinct and stable demographic pattern among casino gamblers at the aggregate level is not noticeable.

B. Conditional Participation Rates

Using the five lottery-type stock preference measures, I find that about 34 to 39% of investors hold lottery-type stocks at least once during the sample period. Across the income and wealth categories, the lottery-type stock participation rate does not vary significantly. The participation rate is about 35% in all income and wealth deciles.

Conditional upon participation, the portfolio weight allocated to lottery-type stocks decreases with both income and wealth. For instance, using the $LP^{(1)}$ measure, I find that investors with an annual income below $25,000 allocate a weight of 13.53% to lottery-type stocks, while investors with an annual income over $100,000 allocate a weight of 9.04% to lottery-type stocks. Furthermore, the mean $LP^{(1)}$ measures for the lowest and the highest wealth deciles are 11.03% and 7.75%, respectively. This evidence is similar to the negative lottery expenditure-income relation identified in previous state lottery studies.

*Citation format: Kumar, Alok, 2009, Internet Appendix to "Who Gambles in the Stock Market?", *Journal of Finance* 64, 1889-1933, http://www.afajof.org/IA/2009. Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the authors of the article.

C. Salient Lottery Characteristics

The three lottery-stock characteristics can be combined in seven different ways to define lottery-type stocks. To determine which lottery characteristic or combination of characteristics is more important for explaining investors' gambling preferences, I conduct two additional tests. First, for different definitions of lottery-type stocks, I examine the average difference between the actual weight and the expected weight in lottery-type stocks in the aggregate individual investor portfolio. The combination of lottery characteristics that has the greatest explanatory power should generate the largest average weight differential. Second, I examine the adjusted R^2 in the investor-level cross-sectional regression for different definitions of lottery-type stocks. The adjusted R^2 should be highest for the combination that has the greatest explanatory power.

When I use only idiosyncratic volatility, idiosyncratic skewness, or stock price to define lottery-type stocks, the average weight differentials are 0.95%, 1.25%, and 1.73%, respectively, and the adjusted R^2 estimates in the investor-level cross-sectional regressions are 0.007, 0.013, and 0.014, respectively. When I use pairs of lottery characteristics, the average weight differential and the adjusted R^2 estimates for the volatility-skewness, volatility-price, and skewness-price pairs are (1.31%, 0.026), (1.49%, 0.020), and (1.83%, 0.029), respectively. Among the three pairs, the skewness-price pair exhibits the strongest explanatory ability. In comparison, when all three characteristics are used to define lottery-type stocks, the average weight differential and adjusted R^2 estimates are 2.49% and 0.043, respectively. These estimates indicate that stock price is the most important lottery characteristic, followed by idiosyncratic skewness and then idiosyncratic volatility.

D. Evidence of Overconfidence Mechanically Induced?

One might argue that to a large extent, the positive relation between excess investment in lottery-type stocks and investor overconfidence reported in Section V.D is hard-wired because stock price is used to define both the dependent and the independent variables. To eliminate the effects of this potential hard-wiring, I re-estimate the cross-sectional regression with two alternative measures of preference for lottery-type stocks. In the first case, the dependent variable is the first lottery-type stock preference measure, where the average is computed using only the portfolio weights during the first three months of the sample period (January 1991 to March 1991). This approach allows me to use non-overlapping time periods to measure the lottery-type stock preference of an investor and her overconfidence. In the second instance, the dependent variable is the average portfolio weight in lottery-type stocks obtained using the weights during the first three months, but stock price is excluded from the definition of lottery-type stocks.

In untabulated results, I find that the coefficient estimate on *Overconfidence Dummy* is positive and strongly significant in both instances. The other coefficient estimates in the regression specification are very similar to the estimates reported in Table IV. This evidence indicates that the positive relation between preference for lottery-type stocks and overconfidence is robust and unlikely to be mechanically induced.

E. Performance of Lottery-type Stocks

I estimate Fama and MacBeth (1973) cross-sectional regressions to examine the performance of lottery-type stocks. I consider a regression specification that is almost identical to the model used in Ang, Hodrick, Xing, and Zhang (2008) to examine the relation between idiosyncratic volatility and stock returns. The dependent variable is the monthly stock return (raw or characteristic-adjusted), and the main independent variables are a lottery-type stock dummy that is set to one for stocks that have lottery features in the previous month and the three lottery characteristics (idiosyncratic volatility, idiosyncratic skewness, and stock price). Other independent variables are the three factor exposures (market, small-minus-big (SMB), and high-minus-low (HML) betas) and various firm characteristics (firm size, book-to-market ratio, past six-month return, and monthly turnover). The factor exposures are measured contemporaneously, while firm size, monthly turnover, and six-month returns are measured in the previous month and the book-to-market ratio is from six months ago.

The cross-sectional regression estimates (Table IA.I) reinforce the evidence from the time-series regressions reported in Table VIII of the main text. When the dependent variable is the raw monthly stock return, the estimation period is from 1980 to 2005, and the full regression specification is used (see column (3)), *Lottery Stock Dummy* has a significantly negative coefficient estimate (estimate = -0.113, t-statistic = -3.97). In economic terms, the estimate of -0.113 translates into an annual underperformance of 1.36%. When I obtain the estimates for the 1991 to 1996 time period (see column (4)), the coefficient on *Lottery Stock Dummy* is similar (estimate = -0.100, t-statistic = -2.48). For additional robustness, when I use characteristic-adjusted stock return as the dependent variable, I find very similar results (see columns (5)-(8)). Overall, the Fama-MacBeth cross-sectional regression estimates indicate that lottery-type stocks earn lower average risk-adjusted returns.

REFERENCES

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2008, High idiosyncratic volatility and low returns: International and further U.S. evidence, *Journal of Financial Economics* Forthcoming.

Campbell, Colin C., and J. Rick Ponting, 1984, The evolution of casino gambling in Alberta, *Canadian Public Policy* 10, 142–155.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Hinch, Thomas, and Gordon J. Walker, 2005, Casino markets: A study of tourist and local patrons, *Tourism and Hospitality Research* 6, 72–87.

Pontiff, Jeffrey, 1996, Costly arbitrage: Evidence from closed-end funds, *Quarterly Journal of Economics* 111, 1135–1151.

Table IA.I

Performance of Lottery-type Stocks: Fama-MacBeth Cross-sectional Regression Estimates

This table reports the estimates from monthly Fama and MacBeth (1973) cross-sectional regressions, where the monthly stock return is the dependent variable. The main independent variable is the lottery-type stock indicator, defined at the end of the previous month. The idiosyncratic volatility in month t is defined as the standard deviation of the residual from the factor model, where daily returns from month t are used to estimate the model. Other independent variables include three factor exposures (market, small-minus-big (SMB), and high-minus-low (HML) betas) and four firm characteristics (firm size, book-to-market ratio, past six-month return, and monthly turnover). The factor exposures are measured contemporaneously, while firm size, six-month returns, and turnover are measured in the previous month and the book-to-market measure is from six months ago. I use the Pontiff (1996) method to correct the Fama-MacBeth standard errors for potential serial correlation. In specifications (1)-(3) and (5)-(7), the sample period is from January 1980 to December 2005. In specifications (4) and (8), the sample period is from January 1991 to December 1996. The t-statistics for the coefficient estimates are shown in smaller font below the estimates. I winsorize all variables at their 0.5 and 99.5 percentile levels and the independent variables have been standardized. Only stocks with CRSP share code 10 and 11 are included in the analysis.

Table IA.I (Continued)

Performance of Lottery-type Stocks: Fama-MacBeth Cross-sectional Regression Estimates

	Raw Return				Characteristic-Adjusted Return			
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	1.243	1.265	1.377	1.828	0.059	0.061	0.095	0.123
	3.97	4.03	4.29	4.13	0.94	0.86	1.22	1.77
Lottery Stock Dummy	−0.105	−0.108	−0.113	−0.100	−0.102	−0.107	−0.116	−0.104
	−2.99	−2.83	−3.97	−2.48	−3.01	−3.06	−4.21	−2.73
Idiosyncratic Volatility		−0.191	−0.301	−0.181		−0.143	−0.257	−0.175
		−3.69	−3.41	−3.11		−3.59	−3.31	−3.16
Idiosyncratic Skewness		−0.103	−0.143	−0.160		−0.153	−0.139	−0.096
		−3.70	−5.27	−3.76		−5.09	−5.32	−2.80
Stock Price		−0.054	0.367	0.016		−0.072	0.253	0.082
		−0.64	5.23	1.17		−1.39	4.37	1.87
Market Beta			1.128	0.958			1.134	0.934
			6.78	4.37			7.47	4.31
SMB Beta			0.037	0.336			0.091	0.296
			0.41	2.64			1.08	2.86
HML Beta			−0.483	−0.363			−0.529	−0.386
			−3.38	−2.04			−4.25	−2.34
Log(Firm Size)			−0.602	−0.407			−0.399	−0.207
			−5.71	−2.89			−5.40	−2.99
Book-To-Market			0.247	0.307			0.092	0.162
			5.23	4.08			2.67	2.77
Past Six-Month Return			−0.387	−0.336			−0.466	−0.436
			−3.34	−2.53			−4.43	−3.73
Monthly Turnover			−0.238	−0.213			−0.190	−0.159
			−3.38	−1.94			−3.31	−1.69
Average Number of Stocks	5,890	5,658	4,370	4,947	5,249	5,111	4,029	4,592
Average Adjusted R^2	0.007	0.021	0.060	0.039	0.004	0.013	0.038	0.029