MANAGEMENT SCIENCE
DOI 10.1287/mnsc.1080.0962ec
pp. ec1–ec13

e-companion
ONLY AVAILABLE IN ELECTRONIC FORM

Electronic Companion—"Trading as Entertainment?" by Daniel Dorn and Paul Sengmueller, *Management Science*, DOI 10.1287/mnsc.1080.0962.

Online Supplement to “Trading as Entertainment?”

EC.1. Nonresponse Bias

Given that only a minority of clients invited to participate in the survey chooses to do so, one might be concerned that the analysis suffers from a nonresponse bias.

In the paper, we specify a model of trading intensity, measured by the logarithm of average turnover, that is linear in investor and portfolio attributes:

$$Y = \boldsymbol{\beta}'\mathbf{X} + \epsilon \tag{EC.1}$$

where Y is the logarithm of turnover and $\mathbf{X}$ is a vector of attributes thought to affect trading intensity. Because turnover is calculated using transactions data and not survey data, it can be calculated for all clients invited to participate in the survey – not just the respondents. The subjective attributes used to construct proxies for the entertainment attributes, are only available for the survey participants, however. Estimating equation EC.1 for the selected sample might lead to biased coefficient estimates. The two-step procedure suggested by Heckman (1979) offers a way to address the resulting specification issue. Consider the model

$$Y = \boldsymbol{\beta_1}'\mathbf{X_1} + \beta_2 X^* + \epsilon_1 \tag{EC.2}$$

where $\mathbf{X_1}$ are investor attributes that are always observed and X^* is an investor trait elicited by the survey. Let us assume that the self-reported survey attribute X is a valid proxy for the true underlying attribute X^* and that

$$X^* = X + u, \text{where } E[u|\mathbf{X_1}, X, \Delta = 1] = 0 \tag{EC.3}$$

where Δ is one if the investor participates in the survey and zero otherwise. This assumption implies that the self-reported attribute X in the participant sample and the corresponding latent construct in the non-participant sample are not subject to differential measurement error as proxies for the true attribute.[8] Under this assumption, taking conditional expectations of equation EC.2 yields

$$E[Y|\mathbf{X_1}, X, I > 0] = \boldsymbol{\beta_1}'\mathbf{X_1} + \beta_2 X + \beta_2 E[u|\mathbf{X_1}, X, \Delta = 1] + E[\epsilon_1|\mathbf{X_1}, X, \Delta = 1] \qquad \text{(EC.4)}$$

$$= \boldsymbol{\beta_1}'\mathbf{X_1} + \beta_2 X + 0 + \lambda H(\boldsymbol{\beta}'\mathbf{Z}) \qquad \text{(EC.5)}$$

where $H(\cdot)$ is the inverse Mills ratio. The coefficients $\boldsymbol{\beta_1}$ and β_2 in equation EC.5 can then be consistently estimated by regressing the logarithm of average turnover on investor and portfolio attributes that are always observed ($\mathbf{X_1}$), the selectively observed attribute (X), and the inverse Mills ratio that can be estimated from the following model of survey participation:

$$\Delta = 1(\boldsymbol{\beta}'\mathbf{Z} + \epsilon) \qquad \text{(EC.6)}$$

where $\mathbf{Z}$ are investor or account attributes that are always observed. Note that, in order to identify the model, Z contains not only $\mathbf{X_1}$, but also variables that affect the participation decision, but not the trading decision.

In a first step, we estimate a probit model for survey participation. In this model, a participation dummy variable is regressed on variables that are thought to affect the decision to respond to the invitation and that are available for investors whether they choose to participate in the survey or not: gender and age of the main account holder, the time between the account opening date and the end of our sample period, the distance between respondent and broker headquarters, the account status (one if the account is active, zero if the account is closed), and the survey response channel (one if the investors filled out the online survey and zero if he filled out the paper survey). The results of this regression are reported in Table EC.1. Other things equal, male and older investors are more likely to participate in the survey, as are longer-time clients, active as opposed to former clients, and clients who received the survey by mail as opposed to by email.

[8] We thank Wei Jiang for pointing out this simplifying assumption.

In a second step, we regress the logarithm of excess turnover on the same attributes as in Table 3 in the paper as well as the inverted Mills ratio obtained from the first step. The regression results are tabulated in Column (1) of Table EC.2. To identify the model, we impose three exclusion restrictions: the distance between respondent and broker headquarters, the account status, and the survey response channel are regressors in the first-stage model, but excluded from the second-stage model.

These exclusion restrictions are, in part, driven by the data since the attributes that improve identification need to be available for the full client sample. However, they appear to be valid in the sense that there is no a priori reason why the corresponding variables should affect investor behavior. Indeed, none of the pairwise correlation coefficients between the three variables and average excess turnover is significantly different from zero.

EC.2. Unimportant Accounts

Another possibility is that the responses to the entertainment statements reflect the importance of the account to the investor. In other words, the accounts of respondents classified as entertainment driven may be small relative to their total (financial) wealth.

Survey respondents provide an estimate of their net worth as well as the allocation of their net worth across different asset classes. We use these responses to construct two proxies for the importance of the observed accounts. First, we divide the value of the investor's observed portfolio at the end of the sample period by the net worth estimate. Second, we divide the value of the observed portfolio by the estimated value of the investor's brokerage assets; brokerage assets are the part of net worth that could, in principle, be held in a brokerage account – for example, stocks, funds, options, but not real estate or cash value life insurance – though not necessarily in the observed account. The typical account represents about one third of self-reported wealth and one half of self-reported brokerage assets.

Columns (2) and (3) of Table EC.2 below report the results of regressions with the two proxies for account importance as additional control variables. Both proxies are significantly negatively related with excess turnover, but the explanatory power of the entertainment attributes is unchanged.

Survey respondents also indicate whether they hold any brokerage accounts other than the observed account. In another robustness check reported in Column (4) of Table EC.2, we focus on roughly the one third of the sample for whom the observed account is the only brokerage account. Again, the explanatory power of the entertainment attributes is similar to that in the baseline regression.

EC.3. Past Performance

It is possible that the enjoyment of investing is merely an artifact of past performance. If so, one would expect the respondents classified as entertainment investors to have outperformed their peers during the sample period.

To check the merits of this story we compute, for a given investor and month, portfolio returns based on the beginning-of-the-month holdings of individual stocks and stock funds (assuming that the positions are held throughout the month). To calculate monthly benchmark returns for a given investor portfolio of individual stocks and stock funds, we create a value-weighted benchmark based on the investor's beginning-of-the-month holdings as follows. To each German stock, we assign an equally-weighted portfolio of German stocks with the same Datastream industry designation and in the same market capitalization tercile based on the beginning-of-the-month market cap (the size terciles are calculated separately for every month-industry combination of German stocks). To each foreign stock, we assign an equally-weighted portfolio of foreign stocks that have the same Datastream industry designation and are in the same market cap tercile. To each stock fund, we assign an equally-weighted portfolio of stock funds that have the same investment objective (large-cap US stocks, for example). The investor's monthly excess return is the difference between the actual portfolio return during the month and the return of the benchmark portfolio.

To assess the effects of trading cost on performance, we consider trading commissions, bid-ask spreads, intra-day returns, and mutual fund loads as follows. If an investor bought 200 shares of an individual stock at a price of DEM 50 per share (this is the actual transaction price, that is, it reflects the bid-ask spread and any price impact), paid a commission of DEM 90, and the Datastream closing price for the stock on the trading date were 49, then the associated trading costs would be DEM 290 (90+200*(50-49)). When trading stock funds, the sample investors do not pay any commissions. However, they often face front loads which may be discounted by the broker. If an investor bought 200 shares of a fund with a net asset value per share of DEM 50, a front load of 5%, and a load rebate of 50% offered by the broker, the trading costs would be DEM 250 (50*200*0.025). Across all transactions, trading costs average 1.2% of transaction value; by themselves, trading commissions average 0.9% of transaction value. To calculate monthly excess returns after trading costs, we sum the trading costs across all transactions of a given investor and month, divide this sum by the average actual portfolio value during the month, and subtract this ratio from monthly excess returns.

Table EC.3 reports monthly raw returns, excess returns, and excess returns after trading costs for investors grouped by their responses to the four entertainment statements, assuming that each group's returns are independent across time; the group return for a given month is the equally-weighted average return across the members of the group.[9] If anything, the raw returns of investors classified as entertainment-driven are below those of their peers. The pattern is

[9] The average returns reported for investors grouped by their responses to the statement "I enjoy investing" are higher than the returns averages for the other statements because there is no investor who disagrees with "I enjoy investing" during the first eight months of the sample. The missing returns and the fact that returns are relatively low at the beginning of the sample period explain the differences in averages.

clearer still for excess returns net of costs, reflecting the higher turnover of entertainment-driven investors. However, return differences between entertainment-driven investors and their peers are not significantly different from zero at conventional levels of significance – owing in part to the small number of respondents that form a particular group at the beginning of the sample period.

In another robustness check (not reported), we include the interaction between past returns and the self-reported enjoyment of investing as an additional explanatory variable in the regression reported in Column (1) of Table 4 in the paper. The interaction coefficient is statistically indistinguishable from zero.

EC.4. Contrasting the Effects of Leisure, Gambling, and Overconfidence

Figure EC.1 below contrasts the excess turnover of investors sorted on their self-professed enjoyment of both investing and gambling. This double-sort produces four investor groups: those who enjoy both investing and gambling (they agree or strongly agree with the two statements "I enjoy investing" and "Games are only fun when money is involved"), those who enjoy neither investing nor gambling (they disagree or strongly disagree with both statements), and those who enjoy one, but not the other. We combine "agree" and "strongly agree" (and "disagree" and "strongly disagree") responses to ensure that each of the four groups contains more than a handful of investors.

Excess turnover increases both with the enjoyment of investing and the enjoyment of gambling; it is highest among those who enjoy investing and gambling. It appears that both leisure and gambling considerations motivate trading.

Similarly, Figure EC.2 below contrasts the excess turnover of investors sorted on their self-professed enjoyment of investing and their knowledge relative to the average investor. Consistent with the results reported in the paper, those who enjoy investing trade much more aggressively than their peers. In contrast, variation in relative knowledge *conditional on enjoyment of investing* is not associated with significant differences in excess turnover.

Finally, Figure EC.3 below contrasts the excess turnover of investors sorted on their affinity for gambling as well as relative knowledge. Excess turnover does not vary significantly as a function of relative knowledge among investors who dislike gambling. However, investors who like gambling and appear overconfident (because they perceive themselves to be much more knowledgable than the average investors) trade more than gamblers who do not appear overconfident.

Figure EC.1: Excess turnover of investors sorted by enjoyment of investing and gambling

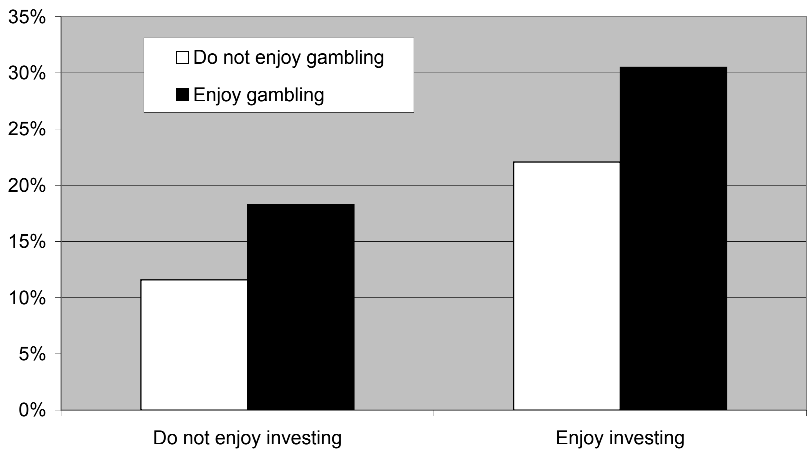

Figure EC.2: Excess turnover of investors sorted by enjoyment of investing and relative knowledge

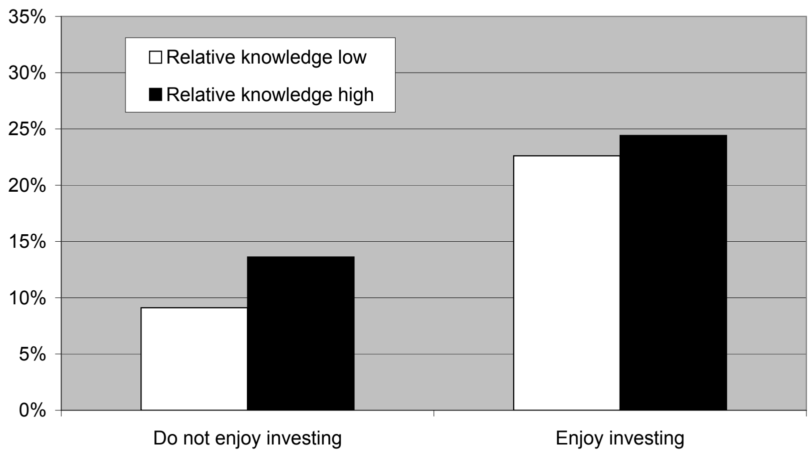

Figure EC.3: Excess turnover of investors sorted by enjoyment of gambling and relative knowledge

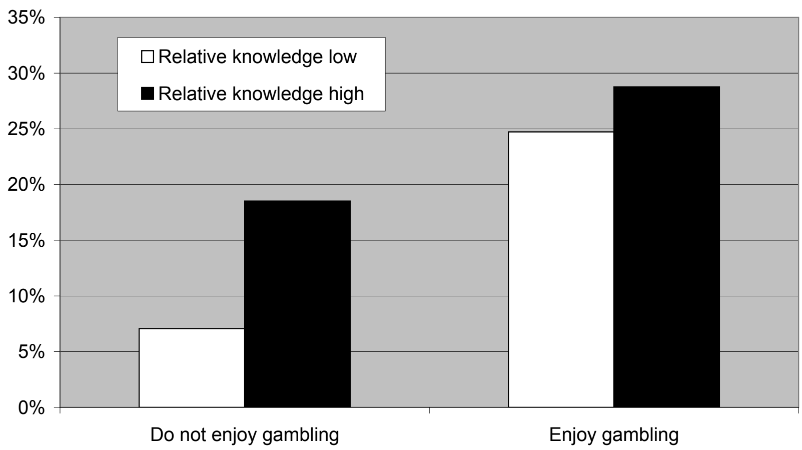

Table EC.1 First-stage probit model for survey participation

The dependent variable is a dummy variable that is one if the invited participant chose to respond and zero otherwise. The number of observations is smaller than the number of investors invited to participate in the survey because one or more of the regressors are missing for several investors invited to participate in the survey; for example, account holders may or may not report their age at account opening. "Account tenure" is the time between account opening and May 31, 2000 or account closing. "Distance to broker" is the distance in kilometers between client and brokerage headquarters, estimated from geographic latitudes and longitudes of client and broker addresses. If a client has a foreign address, distance is set to 1,000 km. "Account active" is a dummy that is one if the account is active as of May 31, 2000 and zero if it was closed before that date. The online survey dummy is one if the investor was invited to participate in the online survey and zero if the investor was invited to participate in the paper survey. *ln* denotes the natural logarithm. Standard errors are in parentheses. Note: ***/**/* indicate that the coefficient estimates are significantly different from zero at the 1%/5%/10% level.

Dependent variable:	Response dummy
Constant	-1.861***
	(0.193)
Gender	0.342***
	(0.051)
Age	0.002
	(0.001)
ln(Distance to broker)	0.002
	(0.005)
Account active	0.633***
	(0.181)
ln(Account tenure)	0.177***
	(0.044)
Online survey dummy	-1.101***
	(0.038)
Ancillary statistics	
Number of observations	14975
Pseudo R^2	10.8%

Table EC.2 Robustness Checks

All regressors are defined as in Table 3 in the paper. The inverted Mills ratio is obtained from the first-stage participation model in Table EC.1 above. “Account as a fraction of wealth” is the portfolio value at the end of the sample period divided by the investor’s net worth as reported in the questionnaire. “Account as a fraction of brokerage wealth” is the portfolio value at the end of the sample period divided by the portion of the investor’s net worth that could be held at a broker, in principle. The regression reported in Column (4) is estimated only for survey respondents who report that the observed account is their only brokerage account. The standard errors in parentheses in Column (1) are the asymptotic standard errors derived in Heckman (1979). The standard errors reported in parentheses in Columns (2)-(4) are corrected for heteroskedasticity as suggested by White (1980). Note: ***/**/* indicate that the coefficient estimates are significantly different from zero at the 1%/5%/10% level.

	(1)	(2)	(3)	(4)
	Logarithm of (1+Excess Turnover)			
Constant	0.279***	0.349***	0.340***	0.296***
	(0.073)	(0.065)	(0.063)	(0.113)
Inverted Mills ratio	0.003			
	(0.016)			
Account as a fraction of wealth		-0.088***		
		(0.024)		
Account as a fraction of brokerage wealth			-0.090***	
			(0.025)	
“I enjoy investing.”				
Tend to agree	0.033	0.030	0.030	0.047*
	(0.032)	(0.023)	(0.023)	(0.026)
Strongly agree	0.082***	0.081***	0.080***	0.095***
	(0.031)	(0.023)	(0.023)	(0.029)
“I enjoy risky propositions.”				
Tend to disagree	-0.017	-0.022	-0.021	-0.020
	(0.024)	(0.019)	(0.019)	(0.033)
Tend to agree	0.036	0.032	0.031	0.013
	(0.026)	(0.024)	(0.024)	(0.039)
Strongly agree	0.158***	0.162***	0.162***	0.117
	(0.037)	(0.062)	(0.062)	(0.105)
“Games are only fun when money is involved.”				
Tend to disagree	0.014	0.013	0.014	0.033
	(0.018)	(0.016)	(0.016)	(0.025)
Tend to agree	0.022	0.015	0.015	-0.003
	(0.022)	(0.021)	(0.021)	(0.037)
Strongly agree	0.057	0.050	0.054	0.035
	(0.048)	(0.049)	(0.048)	(0.071)
“In gambling, the fascination increases with the size of the bet.”				
Tend to disagree	0.002	0.003	0.002	-0.001
	(0.017)	(0.017)	(0.017)	(0.022)
Tend to agree or strongly agree	0.046**	0.048	0.048	0.132**
	(0.023)	(0.029)	(0.029)	(0.062)
Ancillary information				
Number of observations	1054	1031	1031	378
Control variables	Investor attributes as in Table 3 in the paper			
R^2	13.0%	14.4%	14.7%	19.2%

Table EC.3 Past Returns of Investors Grouped by Entertainment Attributes

Groups of investors are formed based on the investors' entertainment attributes. The panels report time-series averages of the groups' returns where the group return for a given month is the equally-weighted average return across the group's member returns that month. Excess returns are measured relative to an investor-specific benchmark that is based on the investor's holdings at the beginning of a given month. The benchmark matches each stock in the investor's portfolio at the beginning of a month with an equally-weighted portfolio of stocks with the same Datastream industry designation and with a similar market capitalization; the benchmark matches each stock fund with an equally-weighted portfolio of stock funds with the same investment objective. To obtain excess returns after fees, we subtract the ratio of trading costs incurred during a given month to average portfolio value from an investor's excess return that month; the trading costs include trading commissions, spreads, price impact, and fund loads where applicable. ***/**/* indicate that the time-series averages of excess returns are significantly different from zero at the 1%/5%/10% level.

	(1) Raw returns	(2) Excess returns	(3) Excess returns net of costs
"I enjoy investing."			
Tend to disagree or strongly disagree	3.3%	0.2%	-0.1%
Tend to agree	2.6%	-0.3%	-0.5%**
Strongly agree	2.8%	-0.1%	-0.4%**
"I enjoy risky propositions."			
Strongly disagree	2.5%	0.0%	-0.2%
Tend to disagree	2.5%	-0.1%	-0.3%**
Tend to agree	2.5%	-0.1%	-0.4%*
Strongly agree	2.2%	-0.3%	-0.6%**
"Games are only fun when money is involved."			
Strongly disagree	2.6%	0.0%	-0.2%
Tend to disagree	2.4%	-0.2%	-0.4%**
Tend to agree	2.2%	-0.4%**	-0.6%***
Strongly agree	2.4%	-0.1%	-0.3%
"In gambling, the fascination increases with the size of the bet."			
Strongly disagree	2.6%	0.0%	-0.3%*
Tend to disagree	2.2%	-0.2%	-0.4%**
Tend to agree or strongly agree	2.4%	-0.3%	-0.6%**