Rudy De Winne

Measuring the disposition effect

LFIN
LOUVAIN FINANCE

LFIN

Measuring the Disposition Effect

Rudy De Winne[*]

December 28, 2020

Abstract

Despite hundreds of papers confirming the existence of the disposition effect, too little attention has been devoted to the prevailing arguments on the choice of a given method to measure it. This paper fills this gap and compares different measurement approaches. First, based on empirical and simulation-based data, I show how results may differ across measures depending on market trends but, more importantly, on the frequency at which investors make their decisions. Second, the pitfalls in analyzing cross-sectional differences in the disposition effect are illustrated and discussed. Finally, I clearly show that hazard models are quite appropriate to measuring the disposition effect of any investor, be it a day trader or a typical retail investor who monitors his portfolio infrequently.

Keywords: Disposition effect, Individual investors, Behavioral finance, Financial Literacy, Trading activity

JEL Classification: G11, G40

1　Introduction

The disposition effect (DE hereafter) refers to investors' stronger propensity to sell winning assets than to sell at a loss. Since Shefrin & Statman's (1985) seminal paper, an extensive literature has confirmed the DE both empirically (e.g., Dhar & Zhu (2006), Barber et al. (2007)) and experimentally (e.g., Weber & Camerer (1998), Chang et al. (2016)).

Two papers pioneered the measurement of the DE. While Weber & Camerer (1998) measure it as the difference between an investor's sales of winner and loser stocks normalized by the total number of his sales, Odean (1998) counts realized gains, realized losses, paper gains and paper losses observed during the sample period (or on specific days of this sample period) before computing the difference in the proportions of realized gains and losses. We classify these methods as ratio-based approaches. They are very easy to implement and the resulting measures are easy to interpret. This probably explains their huge popularity in the last twenty years.

In addition to these ratio-based approaches, econometric techniques have been proposed to analyze the DE and provide interesting alternatives, especially when the goal is to understand which factors (such as financial literacy or portfolio size) may affect the individual DE or when there is a need to control for such factors when comparing individuals. Frydman & Wang (2020) use ordinary least squares regressions to check whether the salience of a stock purchase price affects the disposition effect in a natural experiment. Logistic regression models allow Grinblatt & Keloharju (2001) to check whether the DE interacts with past returns to modify the propensity to sell. Feng & Seasholes (2005) analyze the effects of investor experience and sophistication on the DE. They propose using survival analysis to account for more information than regression analysis can incorporate. In particular, hazard models are well suited to checking whether the holding period of an asset in the portfolio typically observed for an individual investor is affected by changes in a given factor.

An impressive number of papers deal with DE estimation and use one of the approaches mentioned above. Surprisingly, many of them do not give any argument for the choice of the implemented approach. Sometimes there is no clear information on how a given approach has been actually applied

(i.e., we don't know on which days paper gains and losses are counted). This is rather puzzling since not only the chosen method but also its implementation may lead to very different results and conclusions.

This article presents the various approaches to measuring the disposition effect in an exhaustive way and provides some methodological insights. Based on both empirical and simulation-based data, the main contribution is a comparison of results that shows how these different approaches are sensitive to several factors. In particular, the paper shows (1) that Odean's (1998) approach may lead to very different results depending on how paper gains and losses are counted and (2) that hazard models provide more robust results because they account for the typical holding period of stock positions by a given investor better than other models. Another contribution consists in (partially) replicating Dhar & Zhu's (2006) study in order to show how some methodological choices may influence the results. This study analyzes how the DE is affected by financial literacy and trading activity.

The paper is organized as follows. Section 2 presents the alternative approaches and the way that they are implemented in the empirical literature. Using a sample of 17,364 investors, Section 3 analyzes alternative DE measures empirically estimated at the individual level. It shows how these measures are sometimes poorly correlated and how the results are affected by the frequency at which paper gains and losses are counted. In Section 4, the pitfalls to be avoided in analyzing cross-sectional differences in the disposition effect are illustrated through a partial replication of Dhar & Zhu's (2006) study. Section 5 uses simulation to compare the sensitivity of different measures to market trends and to the frequency of investor decisions. Section 6 concludes the paper.

2   Alternative measures of the disposition effect

Whether the analysis is at the individual or group investor level, the DE is assessed by observing trading behavior (actual data or data gathered through an experimental investment setting) or through a set of questions as in Janssen et al. (2020) and Ploner (2017). This paper provides an exhaustive review[1] of the measurement approaches relying on the observed trading behaviors. Taken together, these approaches have been used in hundreds of studies, whether empirical or experimental. Table

---

[1]The only exception to this exhaustive review is the method proposed by Schlarbaum et al. (1978), who analyze the difference in duration of winning and losing roundtrip trades to measure the DE. To my knowledge, only Shapira & Venezia (2001) have applied this approach since.

1 provides a very incomplete list of papers structured according to the approach used for the DE measurement. These approaches may be classified into two families: ratio-based approaches (Panels A & B) and econometric approaches (Panels C, D & E).

2.1 Ratio-based approaches

Weber & Camerer (1998) and Odean (1998) pioneered DE measurement based on trading behaviors. Weber & Camerer's (1998) measure is probably the easiest to compute since it relies on trades only. Using the sale prices and the average purchase prices for the different positions held by an investor, they compute the difference between the number of realized gains (NRG_i) and the number of realized losses (NRL_i) by one investor normalized by the total number of sales by that investor:

$$DE_i = (NRG_i - NRL_i)/(NRG_i + NRL_i) \tag{1}$$

This measure was first used by Weber & Camerer (1998) in an experimental paper, and most of the studies applying this methodology are also experimental. However Dhar & Zhu (2006) and Goetzmann & Massa (2008) use this measure in empirical studies. The latter empirically analyze how the DE affects asset price dynamics and compute a ratio based on Equation 1, except for the sign of the ratio. Dhar & Zhu (2006) point out that Odean's (1998) measure could be sensitive to trading frequency or portfolio size and use Weber & Camerer's (1998) measure for a robustness check. The potential drawbacks of ratio-based approaches will be further discussed in Section 4.

Odean (1998) uses a measure based not only on trading decisions (purchases and sales) but also on investors' portfolio holdings. For each sale, the price is compared with the average purchase price to determine whether it is a realized gain or loss, as in Weber & Camerer (1998). In addition, Odean counts the number of paper gains and losses, which requires additional data about daily high and low prices. On a given day, every position that is not sold is considered a paper gain or loss depending on how the average purchase price of the stock position compares with the highest and lowest prices on that day. If the average purchase price is lower than the lowest price of the day, it is a paper gain, and if it is higher than the highest price, it is a paper loss. If the average purchase price lies between these highest and lowest prices, neither a paper gain nor a paper loss is counted. Based on this counting,

Table 1: Overview of different papers involving a DE measure

References	Main Topic	Data	Comment
Panel A: Weber & Camerer's (1998) measure			
Weber & Camerer (1998)	Potential causes of the DE	X	
Chui (2001)	DE and locus of control	X	
Dhar & Zhu (2006)	DE and investor sophistication	E	
Goetzmann & Massa (2008)	Impact of DE on prices	E	Minus W&C's measure
Da Costa et al. (2013)	DE and experience	X	
Trejos et al. (2019)	DE and overconfidence	X	
Panel B: Odean's (1998) measure			
Odean (1998)	Potential causes of the DE	E	Days with sales
Barber & Odean (1999)	DE in Taiwan ??? market	E	Days with sales
Barber et al. (2007)	Potential causes of the DE	E	All days
Frydman & Rangel (2014)	DE and salience of purchase price	X	All days
Rau (2015)	DE and team trading decisions	E	Days with sales
Fischbacher et al. (2017)	DE and stop orders	X	All days
Panel C: OLS regression			
Chang et al. (2016)	DE and cognitive dissonance	$X+E$	
Frydman & Wang (2020)	DE and salience of capital gains	E	
Panel D: Logit regression			
Grinblatt & Keloharju (2001)	DE and past returns	E	Days with sales
Kaustia (2010)	DE and Prospect Theory	E	Days with sales
Lehenkari (2012)	Potential causes of the DE	E	Days with sales
Birru (2015)	DE and momentum	E	Days with sales
Dierick et al. (2019)	DE and financial attention	E	Days with sales
Panel E: Hazard models			
Feng & Seasholes (2005)	DE and sophistication	E	Weibull
Seru et al. (2010)	DE and the effect of learning	E	Cox
Barber & Odean (2013)	DE and past returns	E	Cox
Richards et al. (2018)	DE and System 1 or 2	E	Weibull
Vaarmets et al. (2019)	DE and learning	E	Cox
Frydman & Wang (2020)	DE and salience of capital gains	E	Cox

Notes: This table shows some of the many papers in which at least one DE measure is used. The list is organized according to this DE measure. Panels A and B refer to ratio-based approaches while the last three panels report papers using an econometric approach. *Data* indicates whether the paper relies on empirical data (E) or on data gathered through an experiment (X). In the last column, I provide some additional comments, such as the distribution used for proportional hazards models and which days are taken into account for the approaches based on Odean (1998) or on a logit regression.

the DE of investor i (or of a group of investors) is measured as the difference between proportion of gains realized (PGR_i) and proportion of losses realized (PLR_i):

$$PGR_i = NRG_i/(NRG_i + NPG_i) \tag{2}$$

$$PLR_i = NRL_i/(NRL_i + NPL_i) \tag{3}$$

$$DE_i = PGR_i - PLR_i \tag{4}$$

In the equations above, NPG_i and NPL_i refer to the number of paper gains and paper losses respectively. In Odean (1998) and Barber & Odean (1999), paper gains and losses are computed "every day that a sale took place in a portfolio of two or more stocks", while Barber et al. (2007) compute these items for every day. According to the latter, the sale-based frequency used in the former papers was imposed by computing resource constraints. They pretend that taking a daily frequency instead allows comparisons across investors holding portfolio of different sizes. In this paper, I compute the measures using both frequencies and show that this choice is not incidental. While seemingly innocuous, different frequency choices can produce different results, depending largely on investors' characteristics such as portfolio size or trading frequency.

These ratio-based approaches are by far the most popular in the literature. The main reason for this is probably that they are simple to compute and to interpret. The ratios range from -1 to +1, and the higher the measures are, the more susceptible investors are to the DE. As pointed out by Weber & Welfens (2008), Weber & Camerer's (1998) measure may be affected by market trends that could result in a lack of selling opportunities for positions that are either at gain (downward trend) or at loss (upward trend). This can explain why Odean's (1998) measure is even more popular despite requiring more information to compute paper gains and losses.

2.2 Econometric approaches

Econometric approaches rely on the estimation of a regression model. This model can be a simple ordinary least squares (OLS) regression model, a logistic regression model or a hazard model. In any case, the principle is to model the decision to sell or to hold a portfolio position and to explain this decision with (at least) one dummy variable indicating whether this position is at gain or at loss, as

6

in the very simple model below:

$$Sale_{i,j,t} = \beta_0 + \beta_1 Gain_{i,j,t} \tag{5}$$

$Sale_{i,j,t}$ is equal to one if investor i sells stock j on day t and zero otherwise. $Gain_{i,j,t}$ equals one if investor i's position in stock j is a gain and zero otherwise. In this simple case, the OLS estimate of β_1 in Equation 5 will be fully equivalent to Odean's (1998) DE measure[2], while the intercept estimate corresponds to the probability of selling a losing position (i.e., PLR_i in Equation 3). This simple OLS technique has been used by Chang et al. (2016) to study how the DE is affected by whether the asset held is a stock or a fund and by Frydman & Wang (2020) to analyze the effects of salience shocks on the DE.

Whatever the regression model used, the main advantage of these econometric approaches is that they allow to add variables and controls on the right-hand side of the model equation[3]. This is especially relevant when cross-sectional differences in the DE are analyzed and related to individual characteristics (e.g., investor's characteristics, trading behavior or portfolio size). For example, if the question to be answered is about the impact of financial literacy on the DE, the model equation will include a proxy for financial literacy (e.g., a score on a test or a dummy variable). However, if an investor's financial literacy tends to increase her asset holding period in general, the propensity to sell a position will be lower whatever that position is at gain or at loss. To further illustrate this case, a negative β_2 estimate in Equation 6 does not mean that financial literacy mitigates the DE and we need to include an interaction term in the model as well:

$$Sale_{i,j,t} = \beta_0 + \beta_1 Gain_{i,j,t} + \beta_2 Literacy_i + \beta_3 Gain_{i,j,t} * Literacy_i \tag{6}$$

The researcher should rather focus on the β_3 estimate in Equation 6, which will reflect the effect of financial literacy on the propensity to sell a winning position. If financial literacy reduces the DE, β_3 should be negative. This will be further discussed in Section 4.

---

[2]Of course, we should apply the same frequency for counting paper gains and losses to get identical results through both approaches.

[3]See Appendix D in Feng & Seasholes (2005) for a nice illustration of the relevance of applying an econometric approach.

Grinblatt & Keloharju (2001) are the first to depart from ratio-based methods and to use a regression model. To study how the DE is affected by past returns, they use a logit (rather than an OLS) regression to model the decision to sell each stock in a portfolio. This model relies on one observation for each position held by an investor on each day when this investor sells any stock. In other words, they use the same frequency for counting paper gains and losses as Odean (1998). The logit model takes the following form:

$$Sale_{i,j,t} = F(X_{i,j}\beta + Z_{i,j,t}\gamma) + \epsilon_{i,j,t} \tag{7}$$

In Equation 7, the right-hand side includes fixed covariates ($X_{i,j}$), such as investor i's financial literacy or gender, and time-varying covariates ($Z_{i,j,t}$), such as a trading gain indicator ($Gain_{i,j,t}$). F represents the cumulative density function of the standard logistic distribution.

In addition to the parameter associated with $Gain_{i,j,t}$, Grinblatt & Keloharju (2001) include variables such as the return associated with the most recent period or a dummy related to the magnitude of the recent return. This method allows analysis of the effects of these factors. Other explanatory variables can also be used to control for characteristics such as investor age or gender or portfolio size. This logit regression approach has been applied in a dozen empirical papers studying how the DE is affected by some factor (e.g., past returns for Kaustia (2010), investor intelligence in Grinblatt et al. (2012) and investor attention in Dierick et al. (2019)).

An alternative and very attractive approach to studying the DE was first proposed by Feng & Seasholes (2005). They use survival analysis, a popular tool used in health departments to analyze the occurrence of death based on several factors such as the administration of a given treatment. It consists in the estimation of a hazard rate, $h(t)$, which is a measure of conditional probability. Applied to the DE estimation, the hazard rate is the probability that the stock position will be sold during the day, given that the stock position was still held in the investor's portfolio until then. In other words, the hazard model describes how quickly investors will sell their positions: on each Day t after a stock position is bought, the hazard of that position being sold (conditional on the position surviving until the end of Day $t - 1$) is estimated.

These hazard models show some similarities with the logistic regression models presented above

since a sell indicator variable is regressed on the baseline hazard function ($h_0(t)$) and the independent variables may be constant (e.g., investor gender or financial literacy) or time-varying (e.g., an indicator informing on whether a position is at gain or not). However, hazard models have a great advantage over logit regression models since they explicitly model the time to an event, be it the death of a person, the failure of a machine, the sale of a portfolio position (e.g., Feng & Seasholes (2005)) or the repurchase of an asset after a sale (Strahilevitz et al. (2011)). The regression coefficients of Equation 8 are estimated using a maximum likelihood method.

$$h(t, x(t)) \quad = \quad h_0(t) \exp\left(\beta_1 x_1 + ... + \beta_p x_p + \gamma_1 z_{1,t} + ... + \gamma_q z_{q,t}\right) \tag{8}$$

The baseline hazard function may take different forms. While Feng & Seasholes (2005) used a Weibull proportional hazards model, Barber & Odean (2013) used a Cox proportional hazards model. The latter is more general in that the specification is semiparametric. This means that $h_0(t)$ is unknown. With the Weibull proportional hazards models, $h_0(t)$ has a functional form that allows the prediction of the hazard that an investor will sell her position on a given day given some known covariates. The parameters p and λ of the baseline hazard function can be estimated:

$$h_0(t) \quad = \quad p\lambda t^{p-1} \tag{9}$$

While the hazard function in Weibull model is monotonic over time, the Cox proportional hazards model allows the hazard function to increase <u>and</u> decrease over time. It can describe a "U" shape or a wide variety of shapes without specification of the underlying hazard function. In both the Weibull and Cox proportional hazards models, the hazard ratio (HR_i) of any coefficient β_i is constant over time and can be computed from the parameter estimate as follows:

$$HR(\beta_i) \quad = \quad \exp\left(\beta_i\right) \tag{10}$$

This hazard ratio provides information about the change in the hazard rate when the related independent variable increases by one unit. For example, if x_1 in Equation 8 is a dummy variable that is equal to one when the position is at gain and zero otherwise, $\exp\left(\beta_1\right) = 3$ means that the hazard

rate of selling when the position is at gain is three times the hazard rate otherwise. The hazard is increased by 200%, all other things being equal.

The issue related to the frequency at which paper gains and losses are determined is no longer relevant with hazard models. Hazard models are designed to model event occurrence data and the estimated hazard function explicitly accounts for the typical duration a position is held in an investor's portfolio through the baseline hazard function $(h_0(t))$. In most of the papers, the hazard models are estimated using a daily frequency. The baseline hazard function $h_0(t)$ represents the investor's behavior when all covariates are set to zero. In a way, it captures the typical holding period of an asset in the investor's portfolio. These characteristics make them quite suitable for analyzing the disposition effect, and it is thus natural that they are used in several papers published in top journals. For example, hazard models are used in Seru et al. (2010) to examine the effect of experience on the DE and in Heimer (2016) to test the relation with social interaction, as well as in some other papers [e.g., Coval & Shumway (2005), Frydman & Wang (2020)].

3 Disposition effect at the individual level

This section examines how the DE of individual retail investors is measured through the different approaches presented in the previous section. Trading data from a brokerage house are used to compute or estimate these measures empirically. This helps clarify whether and how these approaches differ in their results.

3.1 Data and Sample

The dataset comes from a large Belgian brokerage firm and consists of the trading accounts of 77,578 individual investors (entire dataset). The dataset also includes individual characteristics that are either sociodemographic or survey-based. The survey-based data are collected within the context of the MiFID regulation, which has been in place in the European Union member states since November

2007.[4] Another dataset including daily lowest and highest prices coming from Eurofidai and Bloomberg is also used to compute the paper gains or losses on each stock position held by an investor.

Table 2: Descriptive statistics for the sample of investors

	Entire dataset	Sample
# Investors	77,578	17,364
% of men	84.18	91.13
Age	44.55	45.64
# Trades	74.61	125.21
Trading Value	645.31	1,003.22
# Stocks	4.52	7.11
Portfolio Value	38.18	56.93
Holding Period (days)	n.a.	341.93

Notes: This table reports statistics and cross-sectional averages computed for both the entire dataset and the sample. *# Investors* is the number of investors in the entire dataset or the sample. *Age* is measured for each investor in January 2008. *# Trades* is the average number of trades during the whole sample period (January 2003–March 2012) and *Trading Value* is the average monetary value of these trades (thousands of euros). *Portfolio Value* is the average value of an investor's end-of-month portfolio (thousands of euros), while *# Stocks* is the average number of stocks in these portfolios. *Holding Period* refers to the stock holding period, expressed in number of days.

Since Section 4 will replicate a study examining the effect of financial literacy on the DE, the sample is restricted to the 25,277 investors for whom information on financial literacy is available since this is needed for the analysis in Section 4. Following Dhar & Zhu (2006), I kept the investors (17,651) who traded at minimum once a year on average (that is, those with at least 10 trades on the whole sample period). Since the disposition effect is related to the propensity to sell a position, only investors who sold at least one position during the whole sample period are considered. Applying these filters results in a sample of 17,364 investors.

Table 2 gives some basic statistics about both the entire dataset and the sample used in this paper. These statistics show that the sample is not very different from the full dataset. The average investor

---

[4]MiFID stands for the Markets in Financial Instruments Directive. MiFID I (2004/39/EC) was the first version of this directive, while its revision was implemented in January 2018 (known as MiFID II (2014/65/UE)). For more details on the MiFID regulation, please visit the European Commission website (`https://ec.europa.eu/info/law/markets-financial-instruments-mifid-ii-directive-2014-65-eu_en`).

is approximately 45 years old, and the proportion of men is high in both the full dataset and the sample. Of course, the activity of the investors in the sample is higher due to the elimination of very inactive investors. The statistics related to trading activity and portfolio size in the sample are quite in line with what other papers show. For example, Dhar & Zhu (2006) observe an average of 58 trades over their 6-year sample period (slightly less than 10 per year), while the investors in our sample trade approximately 12 times a year. In terms of portfolio size, the investors hold 7 stocks on average in their portfolios. For comparison purposes, in Dhar & Zhu (2006) and in Korniotis & Kumar (2013), they hold portfolios of 5 stocks on average.

3.2   Individual DE estimates

I first apply the ratio-based approaches at the individual level and compute Weber & Camerer's (1998) and Odean's (1998) DE measures for each retail investor in our sample. The first relies on Equation 1, while Equations 2 to 4 are used to apply Odean's approach. As already mentioned, the latter method has been applied with two different frequencies for determining paper gains and losses, i.e., either on each calendar (trading) day, as in Barber et al. (2007), or on selling days only, as in Odean (1998). Here, the measure is computed following both ways.

For the econometric approaches, a regression model is estimated with Equation 5. Unlike Chang et al. (2016), who use a linear regression that is strictly equivalent to Odean's DE, I use a logit model, which is more appropriate when the dependent variable is a dummy variable. The Cox and Weibull proportional hazards models are also applied. For the individual estimation of the DE, Equation 8 includes either a trading gain indicator (TGI) or a trading loss indicator (TLI) as the only covariate and reduces to one of the following equations:

$$h(t, TGI(t)) = h_0(t) \exp\left(\beta_1 . TGI_t\right) \tag{11}$$

$$h(t, TLI(t)) = h_0(t) \exp\left(\beta_1 . TLI_t\right) \tag{12}$$

The estimation relies on data observed on every day of the sample period (daily frequency). Estimating both Equations 11 and 12 could look redundant but it is not since there are many cases where the

average purchase price is between the lowest and the highest prices on a given day, which makes it impossible to determine whether the position is at gain or at loss on that day.

Table 3: Individual measures – Statistics

	N	$Mean$	$Q1$	$Median$	$Q3$
Ratio-based approaches					
WC	17364	35.913	9.091	36.000	66.667
$Odean_D$	14876	2.312	0.137	0.566	1.924
PGR_D	17043	3.523	0.376	0.984	2.867
PLR_D	15160	1.474	0.094	0.284	0.937
$Odean_S$	13519	14.893	2.086	12.500	27.083
PGR_S	15754	38.314	19.178	33.333	54.545
PLR_S	14379	23.630	7.353	16.667	33.333
Econometric approaches					
Logit regression					
Daily frequency (D)	17353	0.026	0.002	0.007	0.020
Sale-based frequency (S)	17076	0.430	0.059	0.186	0.364
Cox PH					
TGI	15910	10.714	1.846	4.507	16.478
TLI	16131	0.467	0.089	0.343	0.866
Weibull PH					
TGI	15502	8.600	2.084	5.018	14.881
TLI	15944	0.450	0.086	0.311	0.837

Notes: This table reports cross-sectional statistics on DE(-related) measures for the whole sample of investors. N is the number of investors for whom the measure can be computed. WC refers to Weber & Camerer's (1998) measure and $Odean$ to Odean's (1998), with a subscript indicating the frequency used to count the paper gains and losses (D for daily and S for sale-based frequency). PGR and PLR refer to the proportion of gains realized and the proportion of losses realized, respectively. The econometric approaches are logistic regressions and hazard models using either Cox or Weibull proportional hazard specifications. For the hazard models, the reported results are the average hazard ratios for both the trading gain indicator (TGI) and the trading loss indicator (TLI) in Equations 11 and 12.

Table 3 shows statistics on the individual DE measures for the 17,364 sample investors. First, it reports the number of investors for whom the alternative measures can be computed (N). For all but Weber & Camerer's (1998) measure, N is lower than the sample size. For example, Odean's measure is a difference between proportions which are computable only if the denominator is positive, i.e. if

13

the investor had a least one position at gain and one position at loss during the whole sample period. Overall, extreme situations lead to the impossibility of computing a given measure (ratio or parameter estimate) or to an extreme result that does not make sense. In the case of hazard models, the averages reported are average hazard ratios. These hazard ratios are given by the exponential of the coefficient estimates. Consequently, an estimate that is already excessively high leads to a very excessive hazard ratio. This leads us to apply a 80% winsorization to the individual hazard ratios before computing the averages reported in Table 3. Winsorizing the top 10% and bottom 10% of data points could look exaggerated but, overall, this is in line with the severe bunching of Odean's (1998) measure observed by Feng & Seasholes (2005) (up to 30% of all the accounts in their sample). A comparison of means and medians reported in Table 3 shows that the data are skewed to the right, which is likely to reflect some extreme behaviors (e.g., selling stocks nearly systematically for gains and never at a loss) observed for some investors in the sample.

When applied at the individual level, hazard models may provide high estimates when the investor systematically sells (or holds) a position in a specific context. Hazard ratios are computed as the exponential of these large numbers, which results in meaningless average hazard ratios. To minimize the impact of these outliers obtained with both Cox and Weibull proportional hazards models, an 80% winsorization has been applied to the vector of TGI and TLI estimates to compute the four means.

Since these alternative approaches have the same general goal, i.e., assessing investors' stronger propensity to sell winning stocks than to sell losing stocks, these measures should be highly correlated. While the ratio-based approaches can be directly compared, the econometric approaches may involve either a trading gain indicator or a trading loss indicator. Consequently, the results of these approaches might not be easily comparable. The hazard models do not provide a direct estimation of the DE, but using a trading gain (loss) indicator, the hazard ratios computed from these models show how being at gain (loss) on a given position impacts investors' propensity to sell this position. The magnitudes of the alternative measures may be very different. This is why Table 4 reports not only Pearson correlation coefficients (based on values[5]) but also Spearman correlations (based on ranks) across these alternative DE(-related) measures. The Spearman correlations are perhaps even more relevant than the Pearson

_____

[5]Because of many extreme values that are likely to bias the results, the Pearson correlation coefficients are computed after a 80% winsorization.

correlations since these measures should at least be consistent in ranking individual investors based on their increased (decreased) propensity to sell a position when it is at gain (loss).

Table 4 shows that the Spearman correlation coefficients are indeed higher in absolute terms, but despite being designed to reflect the same bias, the measures are not highly correlated. While they are the most widely used in the literature, the ratio-based measures are only moderately correlated. Even Odean's (1998) measures computed with both frequencies (daily and sale-based) are not highly correlated (62%, based on ranks). The same remark applies to the logistic approach implemented with both frequencies. More strikingly, the correlations between the variants of Odean's measure and logit estimates are high only when the same choice is made about the frequency at which paper gains and losses are counted. These results are very sensitive to this frequency.

In contrast, the estimates obtained from both the Cox and Weibull hazard models show very strong correlations[6]. The estimation of these hazard models uses data for every day a position is held or sold, but the model results are built on how long a given investor typically holds a position. It can also be observed that the results obtained from the hazard models are more strongly correlated with Odean's measure and the logit estimates when the paper gains and losses are counted on selling days only. A possible explanation for this is that retail investors do not monitor their portfolios frequently and that it is not relevant to consider the holding of any position on a given day as resulting from an actual holding decision. Investors may simply not follow their portfolios for several days, several weeks or even longer. However, investor monitoring frequency is difficult to capture since connection data are not often available in the datasets. Even when these specific data are available (see Dierick et al. (2019)), the connection rate cannot necessarily be interpreted as the monitoring frequency since investors may monitor their positions on other websites and decide to connect to the trading platform only when they want to trade. This issue will specifically be addressed in Section 5 using a simulation.

Even the simple Weber & Camerer's (1998) measure that relies only on realized gains and losses shows a higher correlation with the hazard model estimates than with Odean's measures or the logit estimates, especially when paper gains and losses are determined on a daily frequency. Again, this raises the question about the sensitivity of the latter approaches to investor trading frequency; if this

---

[6]As expected, estimates for the trading loss indicator (TLI) are negatively correlated with the other measures, but in absolute terms, these correlations are quite high.

Table 4: Pearson and Spearman correlations across DE measures

	W&C	$Odean_{All}$	$Odean_S$	$Logit_{All}$	$Logit_S$	Cox		Weibull	
						TGI	TLI	TGI	TLI
Panel A: Pearson correlation coefficients (based on values)									
W&C	1.000	0.129	0.509	0.103	0.395	0.629	-0.692	0.662	-0.683
$Odean_A$	0.129	1.000	0.383	0.721	0.254	0.217	-0.177	0.289	-0.211
$Odean_S$	0.509	0.383	1.000	0.293	0.894	0.508	-0.617	0.597	-0.628
$Logit_A$	0.103	0.721	0.293	1.000	0.163	0.103	-0.065	0.113	-0.059
$Logit_S$	0.395	0.254	0.894	0.163	1.000	0.457	-0.335	0.373	-0.298
Cox TGI	0.629	0.217	0.508	0.103	0.457	1.000	-0.729	0.867	-0.661
Cox TLI	-0.692	-0.177	-0.617	-0.065	-0.335	-0.729	1.000	-0.770	0.948
Weibull TGI	0.662	0.289	0.597	0.113	0.373	0.867	-0.770	1.000	-0.783
Weibull TLI	-0.683	-0.211	-0.628	-0.059	-0.298	-0.661	0.948	-0.783	1.000
Panel B: Spearman correlation coefficients (based on ranks)									
W&C	1.000	0.342	0.524	0.412	0.527	0.718	-0.739	0.731	-0.754
$Odean_A$	0.342	1.000	0.616	0.926	0.615	0.546	-0.441	0.619	-0.497
$Odean_S$	0.524	0.616	1.000	0.523	0.919	0.690	-0.660	0.710	-0.674
$Logit_A$	0.412	0.926	0.523	1.000	0.569	0.473	-0.372	0.549	-0.426
$Logit_S$	0.527	0.615	0.919	0.569	1.000	0.661	-0.600	0.646	-0.583
Cox TGI	0.718	0.546	0.690	0.473	0.661	1.000	-0.963	0.940	-0.915
Cox TLI	-0.739	-0.441	-0.660	-0.372	-0.600	-0.963	1.000	-0.914	0.951
Weibull TGI	0.731	0.619	0.710	0.549	0.646	0.940	-0.914	1.000	-0.958
Weibull TLI	-0.754	-0.497	-0.674	-0.426	-0.583	-0.915	0.951	-0.958	1.000

Notes: This table reports Pearson and Spearman correlation coefficients between alternative DE(-related) measures. For Odean's measures and the Logit models, the subscript refers to the frequency at which paper profits are computed and counted: daily (D) or sale-based (S). The last four measures refer to hazard ratios estimated from a Cox or Weibull proportional hazards model. Correlation coefficients across measures are computed both on values (Panel A) after winsorization and on ranks (Panel B) without winsorization. All the correlations reported in this table are highly significant (all p-values lower than 0.0001 in both panels).

sensitivity is high, it may call into question the conclusions of papers that assess the relationship between trading frequency (or experience, based on the number of trades made by an investor) and the DE. This issue is illustrated in the next section.

4   Investor sophistication and the disposition effect

As an illustration of the importance of the methodological choices, I will partially replicate the study of Dhar & Zhu (2006). The goal of their paper is to identify differences in the disposition effect across individuals and relate them to differences in investor sophistication. They measure the DE at the individual level and analyze how it is affected by investor financial literacy and trading frequency. Since they have no direct information about financial literacy, they use demographic variables that are correlated with financial literacy, i.e. annual income and occupational status. They conclude that both trading frequency (measured by the number of trades) and financial literacy mitigate the DE.

As already mentioned, my dataset includes individual sociodemographic and survey-based information. Instead of using proxies for financial literacy as in Dhar & Zhu (2006), I use a dummy variable based on a self-assessed financial literacy measure. This dummy equals one when the investor assesses himself to be financially literate (levels 4 or 5 on a scale from 1 to 5, indicating that an investor has good knowledge of financial markets and financial experience and has thoroughly mastered all aspects of the financial markets) and zero otherwise.

The main DE measure used by Dhar & Zhu (2006) is Odean's (1998) measure, with paper gains and losses counted on selling days only. The average DE measure obtained for their sample is 21%, while an average DE of 15% is computed for the investors in my sample using the same method (see Table 3). After observing that investors who are prone to the DE (i.e., who have a positive DE measure) have a lower level of financial literacy and a lower trading frequency, the authors build groups based on these dimensions. They observe that their proxies for financial literacy are negatively correlated with the DE and conclude that financial literacy mitigates this behavioral bias.

I replicate the approach of Dhar & Zhu (2006), except that I use the alternative DE measures presented in Section 2 instead of only one. Table 5 reports the average DE for groups of investors who

Table 5: Means across groups of investors

	High	*Low*	*H − L*
Panel A: Trading Frequency - Number of trades			
Group Size	8736	8628	
Weber-Camerer	31.459	40.423	−8.964***
Odean - All days	2.386	2.211	0.176
Odean - Selling days	13.177	17.753	−4.576***
Logit - All days	0.020	0.032	−0.012***
Logit - Selling days	0.195	0.678	−0.483***
Cox - TGI	8.213	13.728	−5.515***
Cox - TLI	0.515	0.412	0.103***
Weibull - TGI	7.419	10.048	−2.629***
Weibull - TLI	0.500	0.393	0.107***
Panel B: Financial Literacy			
Group Size	9580	7784	
Weber-Camerer	34.157	38.075	−3.918***
Odean - All days	2.044	2.661	−0.617***
Odean - Selling days	14.118	15.936	−1.817***
Logit - All days	0.024	0.029	−0.005**
Logit - Selling days	0.380	0.493	−0.114***
Cox - TGI	9.990	11.634	−1.645***
Cox - TLI	0.483	0.448	0.035***
Weibull - TGI	8.122	9.208	−1.087***
Weibull - TLI	0.466	0.430	0.036***

Notes: This table reports average DE(-related) measures and mean comparisons between sub-samples of investors. Panel A reports results for subsamples based on trading frequency (*High* or *Low*, defined as above or below the median number of trades) while Panel B reports the results for subsamples based on self-assessed financial literacy. *Group Size* is the size of the subsample. The different DE(-related) measures are those of Weber & Camerer (1998), those of Odean (1998) and the logit model estimates (both with daily and sale-based frequencies for the counting of paper profits) as well as the hazard ratios of both Cox and Weibull proportional hazards models for either a trading gain indicator (TGI) or a trading loss indicator (TLI). Significance levels for the mean comparisons are *p<0.1, **p<0.05, and ***p<0.01.

trade more or less (High versus Low) than the median investor and for groups of investors who assess themselves as benefiting from high financial literacy or not. In both panels, the *Low* group is often associated with a higher DE measure. Of course, it is the opposite for the estimates associated with the trading loss indicator (TLI) in both hazard model specifications: the closer to one the number is, the more likely she is to sell at loss.

Most of the differences in means have the expected sign and are significant. Panel B shows differences that are all in line with what Dhar & Zhu (2006) show: investors with higher financial literacy are less prone to the DE. All the results reported in Panel A but one are also in line with Dhar & Zhu's (2006) conclusions: trading frequency mitigates the DE. The only exception relates to Odean's measure based on a daily frequency for determining paper gains and losses (following Barber et al. (2007)). In this specific case, the difference has the opposite sign and is not significant. In other words, the frequency used to count paper gains and losses strongly affects Odean's measure and, most importantly, the conclusions of the analysis. Although most papers do not give arguments for the choice of a given frequency, this is not incidental at all.

Ratio-based approaches applied at the individual or group investor level have been very popular because they provide one unique number that is easy to interpret. A positive number means that the investor is prone to the DE. Using a very simple OLS regression model (see Equation 5 as in Chang et al. (2016)) also allows to determine whether this estimate is significantly different from 0. However, the econometric approaches also allow the researcher to explicitly test the relationship between the DE and any particular factor by including some covariates in the model equation.

To examine cross-sectional differences in Odean's (1998) measure of the DE, Dhar & Zhu (2006) regress the individual results on some covariates, such as trading activity. As clearly shown by Feng & Seasholes (2005), the main issue with this approach is that the DE measure itself may be mechanically linked to the behavioral variables that are considered as potential factors affecting the DE. The more often an investor trades, the higher the frequency of selling days is. Dhar & Zhu (2006) are aware of this sensitivity to trading frequency or portfolio size and replicate the analysis with Weber & Camerer's (1998) measure as a robustness check. Even if this measure is probably less sensitive to trading activity compared with Odean's measure, it is prone to bunching problems. Many investors only sell stocks

for gains or only sell stocks for losses. This is particularly true for investors who have underdiversified portfolios and are thus more likely to have only winners or only losers in their portfolios. This may result in a ratio that is equal to one, or negative one[7]. Panel A of Table 5 shows that Odean's measures and logit estimates are much lower when based on a daily frequency in determining the paper gains and losses. The reason for this is simple to understand: many more paper gains and/or losses are counted in this case while the number of realized gains and losses remains unchanged. Using a daily frequency for the counting of paper gains and losses reduces the magnitude of the ratio/estimate. When using a daily frequency, the implicit assumption is that a decision to keep (or to sell) the stock position is made every day. If this hypothesis is perhaps not too troublesome when we analyze this bias among professional investors, it becomes very so when we move to very inactive retail investors. Intuitively, counting the paper gains and losses on days when the investor does not monitor her portfolio provides an upward-biased estimate of the tendency to hold a stock position while limiting this counting to selling days may ignore decisions (to keep a position) that have been made at other times. It is then reasonable to consider that Odean's (1998) measure is mechanically linked to investor trading activity. It is then not surprising that the frequency at which paper gains and losses are counted affects the results. Appendix A focuses on this issue and shows how the conclusions may fully differ depending on the DE measure used.

Given the issues mentioned above and contrary to Dhar & Zhu (2006), I favor econometric approaches for investigating cross-sectional differences in financial literacy and trading activity. For both the logistic regressions and hazard models, different model specifications are used: first a simple model with the trading indicator only and then several models including a financial literacy and/or a trading activity variable. I also add some interaction terms because they are most relevant to answering the key questions. Indeed, investors with a higher level of financial literacy may hold their positions longer, regardless of whether these positions are at gain or at loss, but these investors could be more inclined to keep a winning position even longer.

The logistic regression models are estimated using seven different specifications. Besides the many

---

[7]This bunching phenomenon is discussed in Feng & Seasholes (2005). It also affects Odean's (1998) measure, which sometimes cannot even be computed because either gain or loss opportunities are not available for a given investor.

advantages that they have on ratio-based approaches, logistic regression models (as well as OLS regression models) share a weak point with Odean's (1998) measure. Looking at Equations 5 and 7, it is easy to understand that the covariate $Gain_{i,j,t}$, which indicates whether the position j of investor i is at a gain (1) or not (0) on Day t, needs to be determined at a given frequency. Again the choice of this frequency (every day in the sample period or on selling days only) represents an implicit assumption that could be somewhat arbitrary. To better understand how the conclusions may be affected by this choice, I estimate the logistic regression models using both frequencies.

Table 6 reports the estimates and the average marginal effects of the alternative models. Whatever the model specification, the estimate associated with the trading gain indicator is always positive and significant, which reflects that retail investors are prone to the disposition effect in general. Model 1 is the simplest model, with a trading gain indicator only. The positive and significant estimate associated to the trading gain indicator ($Gain$) is in line with the existence of the DE for the investors of our sample. To check the relationship between investor sophistication and the DE, it could look natural to include proxies of investor sophistication, such as financial literacy and trading activity, in the model (see Model 2) and to build conclusions based on the estimates of their parameters. However, this is clearly not a good idea. In Model 2, it can be observed that trading activity has a positive impact on the probability to sell a portfolio position. Actually, it is not surprising that the more an investor trades, the higher her probability to sell a position. To analyze the impact of trading activity on the DE, Models 4 and 5 are more informative since the parameter estimate of the interaction term ($Gain * Activity$) provides information about how trading activity affects the probability to sell a position at gain (rather than any position). In Model 4, the estimate associated to the interaction term is negative and significant, after controlling for the general impact of trading activity on the selling probability. This means that a trading gain indicator equal to one is associated with a decrease in the propensity to sell a position sooner for investors with a higher trading activity. Similar conclusions about financial literacy can be drawn from Model 3. For the complete model (Model 5), Panels A and B of Table 6 show that the estimates associated with both interaction terms are negative and significant. This can be interpreted as a negative effect of both financial literacy and trading activity on selling a winning position sooner. While these effects are statistically significant, the average marginal effects are quite small compared with the average marginal effect associated to the trading gain indicator.

Table 6: Results of the Logit Regression Analysis

Panel A: Daily frequency (all days taken into account)

	Model 1	Model 2	Model 3	Model 4	Model 5
Gain	1.189***	1.148***	1.247***	1.850***	1.860***
	0.006	0.006	0.006	0.009	0.009
Literacy		−0.233***	0.083***		−0.219***
		−0.001	0.000	.	−0.001
Activity		0.468***		0.530***	0.545***
		0.002		0.003	0.003
*Gain * Literacy*			−0.091***		−0.022***
			0.000		0.000
*Gain * Activity*				−0.124***	−0.123***
				−0.001	−0.001
LR test	277498.125***	611451.316***	277976.899***	608895.708***	618175.256***

Number of observations (Investor x Stock Position x Day): 159,357,416

Panel B: Frequency of sales (only days with sales)

	Model 1	Model 2	Model 3	Model 4	Model 5
Gain	0.972***	0.972***	1.023***	1.471***	1.478***
	0.091	0.089	0.095	0.134	0.134
Literacy		−0.148***	−0.336***		−0.139***
		−0.013	−0.031	.	−0.013
Activity		−0.341***		−0.299***	−0.289***
		−0.031		−0.027	−0.026
Gain*Literacy			−0.072***		−0.016***
			−0.007		−0.001
Gain*Activity				−0.087***	−0.086***
				−0.008	−0.008
LR test	165958.693***	343536.421***	188306.465***	342578.315***	345810.790***

Number of observations (Investor x Stock Position x Day): 7,534,870

Notes: This table reports the estimates for five different logit model specifications obtained by using a daily frequency to count the paper gains and losses (Panel A) or using the frequency of sales (Panel B). For each independent variable included in a given model, the first row reports the parameter estimate and the associated hazard ratio appears on the next row. The dependent variable is a dummy variable which is equal to 1 if the position was sold on that day and 0 otherwise. *Gain* is a dummy variable that equals 1 when the position is at gain and 0 otherwise. *Literacy* is a dummy variable associated with investor financial literacy. This dummy equals one when the investor assesses himself as financially literate (levels 4 or 5 on a scale from 1 to 5, indicating that an investor has good knowledge of financial markets and financial experience and has thoroughly mastered all aspects of financial markets) or zero otherwise. *Activity* refers to the natural logarithm of the number of trades recorded for the investor during the whole sample period. The last two covariates are interaction terms. *LR test* refers to the log-likelihood ratio test (comparison with the null model). Significance levels for the mean comparisons are *p<0.1, **p<0.05, and ***p<0.01.

This means that financial literacy and trading activity only slightly mitigate the DE.

However, Panels A and B of Table 6 show that the results regarding trading activity and financial literacy are sensitive to the frequency at which paper gains and losses are determined. Specifically, the estimate associated with trading activity is always highly significant, but it is positive when a daily frequency is used to determine the paper gains and losses, while it is negative when only selling days are taken into account. Again, as for the use of Odean's measure, it is impossible to give a general advice. Except if detailed information about when investors actually make a decision to sell or to hold a stock position is available, the choice of a frequency could be arbitrary. On the one hand, when using a daily frequency, we implicitly assume that a decision to hold the position is made by the retail investor on each day a position is kept in the portfolio, whether she makes it on purpose or she is simply too busy or unable to monitor her portfolio or to connect to the trading platform. On the other hand, using observations collected on selling days only relies on another implicit assumption: retail investors never decide to keep their positions on a given day unless they decide to sell at least one position, which seems to be a very strong hypothesis.

Ideally, we should rely on days when a decision (to sell or to keep) is explicitly made, but determining when this decision occurs is quite difficult. Relying on a dataset that provides information about when retail investors connect to the trading platform (as in Dierick et al. (2019)) could be an advantage. However, investors have access to alternative websites providing complete information about any asset and they could connect to the trading platform only once they have decided to trade. This means that, while useful, login data may only be considered as a proxy for the timing of actual investors' decisions to keep or sell their positions.

Hazard models may be considered as a solution to the issue related to the frequency at which investors actually decide to sell or to keep an asset position. While they share some characteristics with the logistic regressions, hazard models have the advantage to include a baseline hazard function describing how the hazard rate for the occurrence of a sale per time unit (e.g., per day) changes over time at baseline levels of possible covariates. With this baseline hazard function, the hazard models are able to account for how long an investor or a group of investors typically holds a position. In other words, the hazard models are able to take the typical frequency at which investors make their decisions

23

into account through the baseline hazard function. We don't need to make any assumption about this and the available daily observations can be used to estimate the model parameters.

In addition to this baseline hazard function, the effect parameters describe how the hazard varies in response to changes in the explanatory variables. In the context of a DE analysis, a trading gain indicator which is equal to 1 when the position is at gain (and 0 otherwise) is included on the right-hand side of the equation and other variables that are likely to affect the propensity to sell a position sooner may also be added. This makes it easy to interpret how a change in an independent variable affects the hazards of selling the position. This principle makes hazard models appropriate to analyze the DE

Another enormous advantage is that hazard models allow for varying probabilities for selling a position over time. While Feng & Seasholes (2005) use a Weibull proportional hazards model, Barber & Odean (2013) use a Cox proportional hazards model to analyze the DE. Even if both models have been used to estimate the individual DE, I use Cox models in this section because the estimation is less resource-intensive and because the Weibull model[8] is more restrictive than the Cox model. The latter is semiparametric and allows the estimation of hazard ratios without any assumption on the form of the baseline hazard function. Since I am not interested in forecasting the probability of occurrence of a sale, in which case the baseline hazard function would need to be estimated, this constitutes a major advantage of the Cox proportional hazards model.

Table 7 reports the results for different specifications of Cox proportional hazards models, some of them including interaction terms. All these models include random effects to account for investor characteristics that influence the hazard of the occurrence of the sale. In other words, these investor-specific random effects[9] modify the baseline hazard function for distinct investors. Model 1 is based either on Equation 11 for Panel A or Equation 12 for Panel B, while other models also include other covariates, as in Equation 8 (e.g., financial literacy or trading activity).

First, the estimates regarding the trading gain/loss indicator always have the expected sign and

---

[8]Weibull proportional hazards models with the same specifications as in Table 7 are also estimated. The results are not reported here, but they lead to the same conclusions. These results are available upon request. Table 4 also shows that the individual estimates obtained from the Cox and Weibull proportional hazards models are highly correlated.

[9]Instead of using random effects, a simple clustering of standard errors is also possible. Since this clustering is more popular in the finance literature, I also report the results computed with this approach (without random effects but with robust standard errors) in Appendix B.

Table 7: Results of the Cox proportional hazards models

	Model 1	Model 2	Model 3	Model 4	Model 5
Panel A: Models including a trading gain indicator					
Gain	1.24***	1.24***	1.30***	1.94***	1.96***
	3.45	3.44	3.66	6.95	7.09
Literacy		−0.15***	0.02		−0.11***
		0.86	1.02		0.90
Activity		0.32***		0.40***	0.41***
		1.38		1.49	1.50
Gain ∗ Literacy			−0.10***		−0.05***
			0.91		0.95
Gain ∗ Activity				−0.13***	−0.12***
				0.88	0.88
LR test	487087.20***	489336.88***	487351.82***	492208.36***	492370.95***
Number of observations (Investor x Stock Position x Day): 159,357,416					
Panel B: Models including a trading loss indicator					
Loss	−0.70***	−0.70***	−0.74***	−1.72***	−1.72***
	0.50	0.50	0.48	0.18	0.18
Literacy		−0.14***	−0.05***		−0.14***
		0.87	0.95	.	0.87
Activity		0.33***		0.25***	0.26***
		1.39		1.29	1.30
Loss ∗ Literacy			0.06***		0.00
			1.07		1.00
Loss ∗ Activity				0.19***	0.19***
				1.21	1.20
LR test	357821.12***	360261.65***	357935.60***	366523.36***	366602.20***
Number of observations (Investor x Stock Position x Day): 159,357,416					

Notes: This table reports the results for ten different specifications of Cox proportional hazards models including random effects. The five models in Panel A include a trading gain indicator (*Gain*) while the other five models in Panel B include a trading loss indicator (*Loss*). For each independent variable included in a given model, the first row reports the parameter estimate and the associated hazard ratio appears on the next row. *Gain* (*Loss*) is a dummy variable that equals 1 when the position is at gain (loss) and zero otherwise. *Literacy* is a dummy variable that equals one when the investor assesses himself as financially literate (levels 4 or 5 on a scale from 1 to 5, indicating that an investor has good knowledge of financial markets and financial experience and has thoroughly mastered all aspects of financial markets) or zero otherwise. *Activity* refers to the natural logarithm of the number of trades recorded for the investor during the whole sample period. *Gain/loss ∗ Literacy* and *Gain/loss ∗ Activity* are interaction terms between the trading indicator (*Gain* or *Loss*) and *Literacy* and *Activity*, respectively. *LR test* refers to the log-likelihood ratio test (comparison with the null model). Significance levels are *p$<$0.1, **p$<$0.05, and ***p$<$0.01.

are highly significant. This is in line with the DE affecting the investors in the sample. Second, Model 2 simply includes the variables associated with financial literacy and activity, but for the same reasons discussed in regard to the logistic regression approach, this does not allow us to interpret any relationship between these variables and the DE. Financial literacy decreases the propensity to sell a position, whatever it is a winner or a loser, while high trading activity increases the propensity to sell any position. These partial conclusions could mean that investors with a high level of financial literacy tend to hold their positions longer, while those who are more active sell their positions more quickly. However this does not tell anything about the impact of financial literacy and trading activity on the DE.

In Model 3, the coefficient of the interaction term ($Gain * Literacy$) is more interesting because the general effect of financial literacy on the propensity to sell a position can be disentangled from its specific effect when the position is at gain. The estimate associated with financial literacy becomes statistically nonsignificant while the interaction term is negative and statistically significant. This means that financial literacy has no effect on the propensity to sell a position in general but that it reduces the propensity to sell a winning position. This can be considered a positive effect reducing the DE. From Model 3 including a trading loss indicator (Panel B), it can be observed that the positive effect of the interaction term ($Financial\ Literacy * Loss$) more than compensates for the negative effect of financial literacy on the propensity to sell a position. The hazard ratio for the variable $Literacy$ (0.95) means that the hazard rate is reduced by 5% when the investor shows a high financial literacy but the hazard ratio for the interaction term (1.07) means that the hazard rate is increased by 7% when investor's financial literacy is high and when the position is at a loss. For Model 4, the results reported in both panels may be interpreted the same way. There is a general effect of trading activity that increases the propensity to sell overall, but this higher trading activity results in holding winning positions longer and selling losing positions sooner.

Finally, Model 5 includes both variables ($Literacy$ and $Activity$) as well as the interaction terms. The results show that financial literacy and trading activity have a beneficial effect in terms of reducing the DE. If we look at the results reported in Panels A and B separately, financial literacy seems to diminish the odds of selling a position when it is at gain, while there is no effect when the position is at loss. In other words, the effect of financial literacy on the propensity to sell a winning position

26

is confirmed, but financial literacy does not seem to reduce the holding period of losing positions. Trading activity has a beneficial effect in both cases.

Overall, I show that both financial literacy and trading activity mitigate the DE and that for financial literacy, the effect is observed only for positions at gain. These results are overall in line with the conclusions obtained by Dhar & Zhu (2006). However, there are reasons to think the results presented in this paper are more robust. First, I empirically show that the most popular measure (Odean (1998)) could be misleading because the results based on it are affected by trading activity. If a sale-based frequency is used to determine and count paper gains and losses (as in Dhar & Zhu (2006)), a comparison of the mean DE between active and nonactive investors leads to the conclusion that investor activity reduces the DE. If a daily frequency is used, this conclusion no longer holds. In addition, I show empirically that analyzing cross-sectional differences in the DE could also be risky since the individual measure itself is affected by trading activity (see Appendix A). This issue is probably less sensitive when we look at the relationship with financial literacy because the DE measure is not (or at least less) affected by this element.

Second, I implement a multivariate analysis using different methodological approaches. The first one is a logistic regression, which allows the addition of variables and controls on the right-hand side of the equation. In analyzing the relationship between a given variable and the DE, it is natural to include this variable in the model, but it is also necessary to include interaction term(s), as in Dierick et al. (2019) or Frydman & Wang (2020). However, with logistic regressions, a choice again has to be made between using a daily frequency or a sale-based frequency. Intuitively, one would think that this choice should be based on how frequently an investor monitors her portfolio and actually makes her decisions, but this criterion remains somewhat arbitrary, especially when the data do not provide any information about when the investors connect to the trading platform.

Hazard models may seem to be the right tool to solve this issue but, based on empirical data, it is quite difficult to prove since the true DE levels of the retail investors are unknown. To answer the question of whether hazard models perform better and are able to take the baseline frequency of investors' decisions into account to assess their DE levels, I use simulation.

5 Which measure should we favor?

This section provides some simulation-based results illustrating the advantages and drawbacks of different approaches to measuring the DE. The first simulation is focused on ratio-based approaches only. It shows that these DE measures may provide very different results. The second simulation illustrates the importance of accounting correctly for the frequency of portfolio monitoring.

According to Weber & Welfens (2008), Weber & Camerer's (1998) measure is sensitive to market trends because they could cause a lack of selling opportunities for positions that are either at gain (downward trend) or at loss (upward trend). To better understand this sensitivity, let's consider a simple example[10] with two investors, A and B, who are not prone to the DE. Since these investors are not prone to the DE, their propensity to sell a given position does not depend on whether this position is at gain or at loss. Based on this assumption, the proportion of gains realized (PGR) and the proportion of losses realized (PLR) should be identical on average. In this example, the number of realized gains (NRG) and the number of paper gains (NPG) represent 30% and 70% of the number of selling opportunities for positions at gain, respectively. The same proportions are used for the number of realized losses (NRL) and the number of paper losses (NPL).

Investor	NRG	NPG	NRL	NPL	W&C	PGR	PLR	PGR−PLR
A	18	42	12	28	30%	30%	30%	0%
B	30	70	6	14	67%	30%	30%	0%

Based on these numbers, Odean's (1998) measure (i.e., the difference (PGR-PLR)) is equal to 0% for both investors. However, it can be observed that Weber & Camerer's (1998) measure ($W\&C$) indicates a positive DE and that the measure is positively correlated with the proportion of selling opportunities that are at gain, which are more likely to occur in a bullish market.

However, as shown in previous sections, Odean's (1998) measure and the logit regression approach have also their limits and could lead to different results depending on the frequency at which the paper gains and losses are determined and counted. Ideally, this frequency should accurately reflect

---

[10]This example is similar to Example 1 reported in Appendix D of Feng & Seasholes (2005). However the probability of selling a position is set at 30% (instead of 25%) to stick to the parameter used in my simulation and the number of selling opportunities for positions at gain is higher than the corresponding number for positions at loss (as this is likely to occur when the market trend is bullish).

the frequency at which investors actually make their decisions. In other words, the paper gains and losses should be determined and counted on days when the retail investor makes a decision, regardless of whether she decides to sell or to hold a stock position. Unfortunately, this precise information seems impossible to observe in the real world and, at best, only connection data could help estimate these decision points, which remains insufficient today given that investors may connect on the internet and make decisions to hold their positions without connecting to the trading platform. To avoid rough approximations and arbitrary choices regarding the frequency at which investors actually make their decisions, hazard models are probably a good solution since they are able to account for the baseline frequency of investor decisions. This will be illustrated through the second simulation.

Simulation allows a control of parameters that are sometimes difficult to observe from empirical data. First, it is possible to control for an investor's susceptibility to the DE by determining the likelihoods of selling winning and losing positions. Second, trends in stock prices can be controlled for to assess the sensitivity of the measures to these trends. Finally, it is also possible to precisely know when an investor makes actual decisions to keep or to hold stock positions since the simulation may randomly determine whether an investor is allowed to make a decision on a given day. Unlike the DE level and the monitoring frequency, market trends are easily observable. However, the simulation makes it possible to match a bullish or bearish period of the market to the trading period of an investor.

The first simulation assumes that investors make their decisions every day and I focus on whether and how the ratio-based DE measures are affected by a market trend or by the frequency at which paper gains and losses are counted (daily or only on selling days). In the second simulation, I further analyze how the daily probability of making a (holding or selling) decision on any position of the portfolio may relate to the accuracy of the alternative approaches. For the sake of simplicity, this daily probability is sometimes called "connection rate". Of course, in the real world, a retail investor can monitor her portfolio on the internet and connect to the trading platform only when she decides to actually trade. In this section, the connection rate covers not only the probability for an investor to connect to the trading platform but also the cases when she decides not to connect (and not to trade) after monitoring her portfolio on any website. To be clear, the second simulation differs from the first one in that the connection rate is used to simulate whether the investor makes a decision or not on an given day, regardless of whether this decision means holding or trading a position.

5.1 First simulation

The first simulation is based on 250 iterations including 1500 market trading days. It involves three investor-types with different DE levels and five stocks with normally distributed returns with fixed volatility (0.10%) and means of -0.2%, -0.1%, 0%, 0.1% and 0.2%. On each day, the returns of these stocks are randomly generated. On Day 0 ($t = 0$), the prices of these stocks all equal \$100, and subsequent prices are computed from the simulated returns. The initial portfolio of any investor contains one share of every stock, assumed to have been bought at \$100 on Day 0. Investors have no budget constraint; no short-selling is allowed.

The three investor-types only differ in their DE levels: 0%, 10% and 20%. For each investor, the baseline probability of buying one additional share and of selling an existing position is arbitrarily set to 30%[11] during any trading day. According to Weber & Camerer (1998), the DE is "the tendency to sell assets that have gained value ('winners') and keep assets that have lost value ('losers')". Therefore, an investor prone to the DE should show a higher propensity (or probability) to sell a portfolio position at gain compared with one that is at loss. In the simulation process, I translated this definition as a difference between the probability to sell a position at gain and the probability to sell a position at loss. For a 20%-DE investor, the probability of selling a position is adjusted according to whether the stock is at gain (30% + (20%/2) = 40%) or at loss (30% - (20%/2) = 20%). The difference between these probabilities is equal to 20%.

Even if is in line with the definition found in Weber & Camerer (1998), the DE is generated as a higher probability to sell winners compared to losers, which is exactly how Odean computes his measure (i.e., a difference between PGR and PLR). It could be argued that using these simulated data for performance comparisons would lead to a bias in favor of Odean's measure. Let's keep this in mind when examining the results. However, at least for a zero-DE investor, there is no bias in favor of one measure over others because the baseline probability of selling a position is identical, whether at gain or at loss (compared with the average purchase price of the position).

Using these simulated data, both Weber & Camerer's (1998) and Odean's (1998) measures of the DE are computed. The first measure is computed following Equation 1 while the second is based on Equations 2 to 4. Table 8 reports the Pearson (values) and Spearman (ranks) correlation coefficients

---

[11]Setting this probability to 20% does not affect the conclusions.

between these ratio-based DE measures computed from the simulation data. Although both measures are designed to capture the same bias, they are weakly correlated, except for the two variants of Odean's (1998) measure, which exhibit a (nearly) perfect correlation.

Table 8: Correlations across DE measures based on simulation data

Measure	W&C	$Odean_S$	$Odean_D$
Panel A: Pearson correlations			
W&C	1.0000	0.0937	0.0979
$Odean_S$	0.0937	1.0000	0.9584
$Odean_D$	0.0979	0.9584	1.0000
Panel B: Spearman correlations			
W&C	1.0000	0.1693	0.1839
$Odean_S$	0.1693	1.0000	0.9615
$Odean_D$	0.1839	0.9615	1.0000

Notes: This table reports the correlation coefficients between Weber & Camerer's (1998) and Odean's (1998) DE measures. For the latter, the subscript refers to the frequency at which paper profits are computed and counted: either sale-based frequency (S) or daily (D). The correlations across measures are computed based both on values (Panel A) and on ranks (Panel B). All the correlations reported in this table are highly significant (p-values lower than 0.0001).

Table 9 shows these one-sample comparisons of the different measures using the true DE level (0%, 10% or 20% used in the simulation) in the null hypothesis[12]. From Panel A, the sensitivity of Weber & Camerer's (1998) measure to a market trend is really striking. Whatever the true DE level, all the average values for Weber & Camerer's (1998) measure are significantly different from the true value, except when there is no trend in the returns. Even when the sales are generated in the same way for winners and losers ($DE = 0\%$) and are simply based on a fixed probability to sell, the measure is clearly affected by a market trend and this poor performance cannot be explained by the method used to generate the true DE in the simulation.

Although the results in Panel B are more in line with the true DE value, significant differences are observed depending on whether a daily (D) or a sale-based (S) frequency is chosen. Using a sale-based frequency results in a DE measurement estimation bias that increases with the true DE

---

[12]Let's keep in mind that the DE is generated through a probability gap between selling stocks at gain and at loss, which is exactly the formula for Odean's measure.

Table 9: Performance of ratio-based measures of the disposition effect

Frequency	Trend	DE = 0%	DE = 10%	DE = 20%
Panel A: Weber & Camerer's (1998) measure				
	-0.2%	-84.423^{***}	-81.931^{***}	-79.692^{***}
	-0.1%	-57.365^{***}	-50.070^{***}	-43.099^{***}
	0.0%	-0.467	9.618	20.516
	0.1%	56.189^{***}	62.890^{***}	69.597^{***}
	0.2%	84.956^{***}	87.424^{***}	90.003^{***}
Panel B: Odean's (1998) measure				
D	-0.2%	0.089	9.754	19.106^{*}
D	-0.1%	-0.256	9.812	20.240
D	0.0%	0.193	10.137	19.844
D	0.1%	-0.215	9.666	19.943
D	0.2%	0.119	9.978	20.400
S	-0.2%	0.321	11.166^{**}	23.098^{***}
S	-0.1%	-0.452	11.280^{***}	23.963^{***}
S	0.0%	0.124	12.141^{***}	23.796^{***}
S	0.1%	-0.695^{**}	11.207^{***}	23.708^{***}
S	0.2%	0.424	11.868^{***}	24.074^{***}

Notes: Frequency is the frequency at which paper profits are computed: daily (D) or sale-based (S). *Trend* is the mean return used for the simulation of stock returns. The other five columns ($DE = x$) refer to the specific DE values used in the simulation. The DE measures are expressed in percentages, and the null hypothesis used for the test is based on the true DE value. Significance levels are $^{*}p<0.1$, $^{**}p<0.05$, and $^{***}p<0.01$.

value. Regardless of the trend and the true DE level, the estimates provided when a daily frequency is chosen are not statistically different from the true DE value. At first glance, this seems in line with Barber et al.'s (2007) assertion that without resource constraints, a daily frequency should be chosen. However, in the first simulation, decisions to sell a position (or not) are made every day by the machine, while in the real world, most retail investors do not monitor their portfolios daily. It seems strange to consider an investor's decision to keep a winner/loser position on days when the investor is not even aware of the evolution of his positions. To account for this, a parameter related to the connection rate is added in the second simulation.

5.2 Second simulation

The goal of this simulation is to assess how monitoring intensity or the frequency at which the investor makes a decision affects the DE estimation. The exercise relies on 250 iterations divided into five groups of equal size across which the daily connection probability varies. In the first group, a daily connection is assumed for the three investor-types, but for every other 30-iteration group, connection probabilities of 50%, 20%, 10% and 5% for every investor-type are considered. Whether an investor connects to the trading platform is randomly determined for each day. An investor is able to trade on a given day only if he connects to the platform.

In addition to computing Odean's DE with daily (D) and sale-based (S) frequencies used so far to determine and count paper gains and losses, I also compute this measure based on the connection frequency (C) since the simulated data allow a clear identification of the days where a connection was randomly determined. Each panel of Table 10 refers to a given connection rate. In Panel A, the results associated with connecting every day are similar to those in Table 9, and the DE measures using the D and C frequencies are thus identical and perform well. The next panels show that the daily frequency recommended by Barber et al. (2007) is sometimes the worst solution. When the connection probability is low, the "daily-frequency" ratio is often significantly lower than the true DE value (approximately the true DE value multiplied by the connection probability). When counting the paper gains and losses, an arbitrary choice regarding the frequency may bias the results more seriously than Barber et al. (2007) suggest. Actually, Barber et al. (2007) are right when they claim that comparing different DE estimates across large and small investors (with some correlation with trading frequency) is difficult, but they seem to be wrong when suggesting that considering a daily frequency when counting the paper profits solves the problem.

Consequently, counting paper profits on connection days may be preferable but this information is rarely available in the datasets provided by brokerage firms[13]. Moreover, many investors may monitor their positions on any financial website and connect to the platform only when they need to trade. Based on these arguments, a practical choice could be to count paper profits on days when some trading activity was observed for the account (e.g., trading, order submission) since it seems reasonable to assume a good correlation between monitoring frequency and trading activity.

---

[13]To my knowledge, the only exception is the dataset used by Dierick et al. (2019).

Table 10: Disposition effect 'à la Odean' and the connection rate

Frequency	$DE = 0\%$	$DE = 10\%$	$DE = 20\%$
Panel A: Connection probability = 100%			
D	0.143	10.192	20.012
S	0.162	12.428***	24.452***
C	0.143	10.192	20.012
Panel B: Connection probability = 50%			
D	−0.121	5.202***	10.044***
S	−0.407	12.32***	24.435***
C	−0.242	10.086	19.852
Panel C: Connection probability = 20%			
D	0.116	2.040***	4.037***
S	0.648	12.382***	24.634***
C	0.649	10.099	20.013
Panel D: Connection probability = 10%			
D	−0.003	1.092***	1.993***
S	−0.773	13.795***	24.884***
C	−0.286	11.044	19.713
Panel E: Connection probability = 5%			
D	−0.060	0.559***	1.068***
S	−1.253	12.153*	26.976***
C	−1.352	10.837	21.815*

Notes: Frequency is the frequency at which paper gains and losses are determined and counted: daily (D), sale-based (S), and connection-based (C). The other columns ($DE = x$) refer to the specific DE values used in the simulation. All Odean's (1998) DE measures are expressed in percentages and are compared with the true DE value appearing in the first row of the table. Significance levels are *p<0.1, **p<0.05, and ***p<0.01.

Luckily, hazard models provide a much better solution to this monitoring frequency problem. They are designed to model event occurrence data. In a way, the monitoring frequency, connection rate and typical holding period of stock positions are individual habits that can be captured through the baseline hazard function ($h_0(t)$). This is the hazard function obtained when all covariates are set to zero. The estimates obtained for the hazard function provide information about how the presence of nonnull covariates (e.g., a trading gain indicator equal to one) may impact the hazard of selling the position.

Using data from the second simulation, I estimate two simple Cox proportional hazards models

Table 11: Hazard model estimates and the connection rate

Connection Rate	DE = 0%		DE = 10%		DE = 20%	
	TGI	TLI	TGI	TLI	TGI	TLI
Value	1.000	1.000	1.400	0.714	2.000	0.500
5%	1.006	1.066	1.563*	0.694*	2.228**	0.478**
10%	1.020	1.012	1.482	0.699	2.134*	0.483*
20%	1.027	0.987	1.453	0.701	2.058	0.494
50%	0.995	1.012	1.464***	0.689***	2.117***	0.475***
100%	1.006	0.995	1.510***	0.665***	2.277***	0.442***

Notes: The TGI (TLI) columns report the hazard ratios associated with the trading gain (loss) indicator in a Cox proportional hazards model. $Value$ provides the parameter value based on the true DE level used in the simulation. The columns ($DE = x$) refer to the specific DE values used in the simulation. The hypothesis that the average estimate is equal to its $Value$ is tested. Significance levels are *p<0.1, **p<0.05, and ***p<0.01.

based on Equations 11 and 12. Each of them includes only one covariate, a trading gain/loss indicator. Table 11 reports the hazard ratios obtained with these simple Cox proportional hazards model specifications. First, it is striking how the estimates are stable across the five connection frequencies, for any of true DE values considered. An ANOVA shows that the equality of the model parameters (whether TGI or TLI) across the five connection rate categories cannot be rejected. This means that this approach gives results that are not sensitive to the connection rate. Second, when the DE value used in the simulation is zero, the hazard ratios associated with TGI or TLI are close to one. Indeed, the estimates do not differ statistically from zero (even at a significance level of 10%). Third, Table 11 reports hazard ratios for TGI (TLI) that are increasing (decreasing) with the DE value used in the simulation, in line with expectations. I test whether all these hazard ratios are statistically different from the expectation based on the true DE used in the simulation. If we look at the column associated to TGI for an investor with a 10% DE and a connection rate of 10%, the estimate (1.482) is not statistically different from 1.4 (i.e., 35%/25%), the hazard ratio that should be obtain based on the hypothesis of a 10% DE used in the simulation. However Table 11 shows that some estimates statistically differ from the expected values (1.4, 0.714, 2 and 0.5 for the last 4 columns, respectively) but these estimates are quite close. Let's remember that the DE is generated as a higher probability

to sell winners compared to losers, which is exactly how Odean computes his measure. Despite this, hazard models outperform Odean's measure. Let's take the example of an investor who is monitoring her portfolio every two weeks (connection rate = 10%) and whose DE is 20%. Based on the results reported in Table 11, the investor is likely to have an estimated probability to sell at gain (at loss) slightly higher that 40% (slightly lower than 20%). For example, these probabilities could be equal to 40.8% and 19.12% (= 40.8%/2, 134 in order to respect the hazard ratio). This is quite close to the true numbers and would mean an estimated DE of 40.8% − 19.12% = 21.88%. Compared with Odean's results reported in Table 10, it is a better estimation of the true DE, except when the actual timing of investor's decisions is known (which never happens in the real world).

All these results confirm that hazard models are appropriate to the disposition effect measurement, especially when individual investors exhibit very different behaviors in terms of portfolio monitoring and/or trading frequency.

6   Conclusion

The disposition effect has been largely confirmed in hundreds of published papers. Surprisingly, while the methods used to measure this bias differ across these papers, little attention has been devoted to the measurement issue, and most of the time, no clear argumentation has been proposed for the use of a given approach. This paper aims to show how important the choice of measure and the implementation of this choice are.

The first contribution to the literature relies on an exhaustive analysis of different empirical approaches to the DE measurement. These approaches are based either on ratios or on econometric models but, while they are supposed to measure the same bias, the different measures are often weakly correlated. I show that Weber & Camerer's (1998) measure is sensitive to market trends and that Odean's (1998) measure and a logit regression approach could lead to different results depending on the frequency at which the paper gains and losses are determined and counted. While this frequency should accurately reflect the frequency at which investors make their decisions, the latter can never be observed. Login data are rarely available and, when they are, they do not allow to estimate the accurate timing of individual investors' decisions. Indeed, investors may connect on the internet and

decide to hold their positions without loging to the trading platform.

Second, this paper provides some methodological insights regarding how cross-sectional differences in the disposition effect should be analyzed, especially when the DE measure itself maybe mechanically linked to a variable that is considered as a potential factor affecting the DE. The pitfalls are illustrated through a partial replication of Dhar & Zhu's (2006) study. My results are rather in line with theirs but are more robust. Using an approach based on hazard models, I show that financial literacy and trading activity slightly mitigate the DE.

Last but not least, I show that hazard models are the most suitable for measuring and analyzing the disposition effect. They have an enormous advantage over other approaches since they rely on a baseline hazard function that accounts for how long an investor (or a group of investors) typically holds a position. Using a simulation, I show that thanks to this baseline hazard function, hazard models can be estimated identically for any individual, be it a day trader or a typical retail investor who monitors his portfolio very infrequently. Unlike other measures, hazard models do not require any assumptions about whether an investor makes a decision every day, once a week or only on selling/trading days, and the hazard ratios of the trading gain/loss indicator estimated using hazard models are not sensitive to the frequency at which investors make their decisions. For each investor (or group of investors), the hazard function provides information about how a change in a given explanatory variable affects the conditional probability of selling a position compared with the typical probability represented by the baseline hazard function.

References

Barber, B. M., Lee, Y.-T., Liu, Y.-J. & Odean, T. (2007), 'Is the aggregate investor reluctant to realise losses? Evidence from Taiwan', *European Financial Management* **13**(4), 423–447.

Barber, B. M. & Odean, T. (1999), 'The courage of misguided convictions', *Financial Analysts Journal* **55**(6), 41–55.

Barber, B. M. & Odean, T. (2013), The behavior of individual investors, *in* 'Handbook of the Economics of Finance', Vol. 2, Elsevier, pp. 1533–1570.

Birru, J. (2015), 'Confusion of confusions: A Test of the Disposition Effect and Momentum', *Review of Financial Studies* **28**(7), 1849–1873.

Chang, T. Y., Solomon, D. H. & Westerfield, M. M. (2016), 'Looking for someone to blame: Delegation, cognitive dissonance, and the disposition effect', *Journal of Finance* **71**(1), 267–302.

Chui, P. M. W. (2001), 'An Experimental Study of the Disposition Effect: Evidence From Macau', *Journal of Psychology and Financial Markets* **2**(4), 216–222.

Coval, J. D. & Shumway, T. (2005), 'Do behavioral biases affect prices?', *Journal of Finance* **60**(1), 1–34.

Da Costa, N., Goulart, M., Cupertino, C., Macedo, J. & Da Silva, S. (2013), 'The disposition effect and investor experience', *Journal of Banking & Finance* **37**(5), 1669–1675.

Dhar, R. & Zhu, N. (2006), 'Up close and personal: Investor sophistication and the disposition effect', *Management Science* **52**(5), 726–740.

Dierick, N., Heyman, D., Inghelbrecht, K. & Stieperaere, H. (2019), 'Financial attention and the disposition effect', *Journal of Economic Behavior and Organization* **163**, 190–217.

Feng, L. & Seasholes, M. S. (2005), 'Do investor sophistication and trading experience eliminate behavioral biases in financial markets?', *Review of Finance* **9**(3), 305–351.

Fischbacher, U., Hoffmann, G. & Schudy, S. (2017), 'The Causal Effect of Stop-Loss and Take-Gain Orders on the Disposition Effect', *The Review of Financial Studies* **30**(6), 2110–2129. Publisher: Oxford Academic.

Frydman, C. & Rangel, A. (2014), 'Debiasing the disposition effect by reducing the saliency of information about a stock's purchase price', *Journal of Economic Behavior & Organization* **107**, 541–552.

Frydman, C. & Wang, B. (2020), 'The Impact of Salience on Investor Behavior: Evidence from a Natural Experiment', *Journal of Finance* **75**(1), 229–276.

Goetzmann, W. N. & Massa, M. (2008), 'Disposition Matters: Volume, Volatility, and Price Impact of a Behavioral Bias', *The Journal of Portfolio Management* **34**(2), 103–125.

Grinblatt, M. & Keloharju, M. (2001), 'What makes investors trade?', *Journal of Finance* **56**(2), 589–616.

Grinblatt, M., Keloharju, M. & Linnainmaa, J. T. (2012), 'Iq, trading behavior and performance', *Journal of Financial Economics* **104**(2), 339–362.

Heimer, R. Z. (2016), 'Peer pressure: Social interaction and the disposition effect', *Review of Financial Studies* **29**(11), 3177–3209.

Janssen, D. J., Li, J., Qiu, J. & Weitzel, U. (2020), 'The Disposition Effect and Underreaction to Private Information', *Journal of Economic Dynamics and Control* **113**, 1–28.

Kaustia, M. (2010), 'Prospect Theory and the Disposition Effect', *Journal of Financial and Quantative Analysis* **45**(3), 791–812.

Korniotis, G. M. & Kumar, A. (2013), 'Do Portfolio Distortions Reflect Superior Information or Psychological Biases?', *Journal of Financial and Quantative Analysis* **48**(1), 1–45.

Lehenkari, M. (2012), 'In Search of the Underlying Mechanism of the Disposition Effect', *Journal of Behavioral Decision Making* **25**(2), 196–209.

Odean, T. (1998), 'Are investors reluctant to realize their losses?', *Journal of Finance* **53**, 1775–1798.

Ploner, M. (2017), 'Hold on to it? An Experimental Analysis of the Disposition Effect', *Judgement and Decision Making* **12**(2), 118–127.

Rau, H. A. (2015), 'The disposition effect in team investment decisions: Experimental evidence', *Journal of Banking & Finance* **61**, 272–282.

Richards, D. W., Fenton-O'Creevy, M., Rutterord, J. & Kodwani, D. G. (2018), 'Is the disposition effect related to investors' reliance on System 1 and System 2 processes or their strategy of emotion regulation?', *Journal of Economic Psychology* **66**, 79–92. Publisher: North-Holland.

Schlarbaum, G. G., Lewellen, W. G. & Lease, R. C. (1978), 'Realized Returns on Common Stock Investments: The Experience of Individual Investors', *Journal of Business* **51**(2), 299–325.

Seru, A., Shumway, T. & Stoffman, N. (2010), 'Learning by Trading', *Review of Financial Studies* **23**(2), 705–739.

Shapira, Z. & Venezia, I. (2001), 'Patterns of Behavior of Professionally Managed and Independent Investors', *Journal of Banking and Finance* **25**(8), 1573–1587.

Shefrin, H. & Statman, M. (1985), 'The disposition effect to sell winners too early and ride losers too long: Theory and evidence', *Journal of Finance* **40**, 777–790.

Strahilevitz, M. A., Odean, T. & Barber, B. M. (2011), 'Once Burned, Twice Shy: How Naive Learning, Counterfactuals, and Regret Affect the Repurchase of Stocks Previously Sold', *Journal of Marketing Research* **48**, S102–S120.

Trejos, C., van Deemen, A., Rodríguez, Y. E. & Gómez, J. M. (2019), 'Overconfidence and disposition effect in the stock market: A micro world based setting', *Journal of Behavioral and Experimental Finance* **21**, 61–69.

Vaarmets, T., Liivamägi, K. & Talpsepp, T. (2019), 'How Does Learning and Education Help to Overcome the Disposition Effect?*', *Review of Finance* **23**(4), 801–830.

Weber, M. & Camerer, C. F. (1998), 'The disposition effect in securities trading: An experimental analysis', *Journal of Economic Behavior and Organization* **33**, 167–184.

Weber, M. & Welfens, F. (2008), 'Splitting the disposition effect: Asymmetric reactions towards 'selling winners' and 'holding losers'', *SSRN Electronic Journal* .

Appendix

A Cross-sectional differences in the disposition effect

Dhar & Zhu (2006) regress individual Odean's (1998) measures on some covariates, such as trading activity and control variables, to examine cross-sectional differences. As clearly shown by Feng & Seasholes (2005), the main issue with this approach is that the DE measure itself may be mechanically linked to the behavioral variables that are considered as potential factors affecting the DE. In their example presented in Appendix D, Feng & Seasholes (2005) focus on how portfolio size may affect several ratio-based measures.

This appendix provides another illustration of this issue and shows how the conclusions may differ entirely depending on the DE measure used. OLS regression models are used and only differ in terms of the dependent variable:

$$DE_i = \beta_0 + \beta_1 X_1 + ... + \beta_p X_p \tag{13}$$

The dependent variable is one of the alternative measures and I use covariates referring to self-assessed financial literacy[14] and to activity (measured by the logarithm of the total number of trades by the investor) as well as some control variables (age, gender, and portfolio size).

The results in Table 12 show that this kind of analysis is risky when the measure used as the dependent variable is sensitive to a change in some covariates. This observation is striking when we look at the estimates associated with *Activity* in the alternative models. Conclusions based on this approach may be contradictory, and this is particularly the case with Odean's (1998) measure with different frequencies at which paper gains and losses are counted. From this illustration, it is important to note that the analysis of cross-sectional differences in the DE should rely on a multivariate approach rather than any approach based on individual DE measures.

---

[14]Dhar & Zhu (2006) use proxies for financial literacy (i.e., income and type of occupation).

Table 12: Analysis of cross-sectional differences in the DE

	$W\&C$	$Odean_{All}$	$Odean_S$	$Logit_{All}$	$Logit_S$	$Cox\ TGI$	$Cox\ TLI$
Intercept	35.110***	−0.440	27.051***	−0.009***	0.263***	14.452***	0.281***
Literacy	−0.400	−0.445***	−0.775**	−0.002***	−0.010***	−0.520***	0.008
Activity	−1.349***	1.135***	−1.142***	0.010***	0.000	−1.371***	0.063***
Age	−0.022	−0.025***	−0.077***	0.000***	−0.001***	0.014**	−0.001**
Gender	−2.269**	−0.372**	−1.361**	−0.001*	−0.008*	−0.971***	0.034***
Portfolio size	0.461***	−0.126***	−0.266***	−0.001***	−0.004***	0.076***	−0.007***
R^2	0.013	0.064	0.027	0.137	0.078	0.020	0.034

Notes: This table reports the results (OLS model estimates) of the analysis of cross-sectional differences in the DE. The dependent variable is one of the alternative DE measures. *Literacy* is a dummy variable associated with investor financial literacy. This dummy equals one when the investor assesses himself as financially literate (levels 4 or 5 on a scale from 1 to 5, indicating that an investor has good knowledge of financial markets and financial experience and has thoroughly mastered all aspects of financial markets) or zero otherwise. *Activity* refers to the natural logarithm of the number of trades recorded for the investor during the whole sample period. *Age* is the investor's age in January 2008. *Gender* is a dummy that equals 1 for a man and zero otherwise. *Portfolio size* is the average number of stocks held in the investor's portfolio. Significance levels are *p<0.1, **p<0.05, and ***p<0.01.

B Other results using hazard models

In Section 4, Table 7 reports the results obtained with different Cox proportional hazard model specifications. These models are enhanced through the incorporation of random effect terms to account for within-investor characteristics that affect the hazard of the occurrence of the sale decision. Table 13 reports the results obtained with the same Cox proportional hazards models without random effects. Instead, robust standard errors are computed. Most of the results are in line with what Table 7 shows, except the estimate for the *TGI * Literacy* interaction term in Model *G*7.

Table 13: Estimates from Cox proportional hazards models with robust standard errors

	Model 1	Model 2	Model 3	Model 4	Model 5
Panel A: Models including a trading gain indicator					
Gain	1.10***	1.09***	1.14***	1.55***	1.56***
	2.99	2.98	3.12	4.69	4.75
Literacy		−0.17***	0.01		−0.14***
		0.85	1.01		0.87
Activity		0.30***		0.34***	0.35***
		1.34		1.40	1.42
Gain ∗ Literacy			−0.07***		−0.03
			0.93		0.97
Gain ∗ Activity				−0.08***	−0.08***
				0.92	0.92
LR test	153451.84***	229921.96***	153716.71***	228144.32***	231459.99***
Number of observations (Investor x Stock Position x Day): 159,357,416					
Panel B: Models including a trading loss indicator					
Loss	−0.59***	−0.58***	−0.62***	−1.33***	−1.33***
	0.56	0.56	0.54	0.26	0.26
Literacy		−0.16***	−0.04**		−0.15***
		0.85	0.96		0.86
Activity		0.30***		0.24***	0.25***
		1.34		1.27	1.28
Loss ∗ Literacy			0.06**		−0.01
			1.06		0.99
Loss ∗ Activity				0.14***	0.14***
				1.15	1.15
LR test	42327.67***	118532.31***	42493.09***	119708.54***	122709.99***
Number of observations (Investor x Stock Position x Day): 159,357,416					

Notes: This table reports the results for ten different specifications of Cox proportional hazards models. No random effect is included here, but robust standard errors have been computed. The five models in Panel A include a trading gain indicator (*Gain*) while the models in Panel B include a trading loss indicator (*Loss*). For each independent variable included in a given model, the first row reports the parameter estimate and the associated hazard ratio appears on the next row. *Gain* (*Loss*) is a dummy variable that equals 1 when the position is at gain (loss) and zero otherwise. *Literacy* is a dummy variable that equals one when the investor assesses himself as financially literate (levels 4 or 5 on a scale from 1 to 5, indicating that an investor has good knowledge of financial markets and financial experience and has thoroughly mastered all aspects of financial markets) or zero otherwise. *Activity* refers to the natural logarithm of the number of trades recorded for the investor during the whole sample period. *Gain/Loss ∗ Literacy* and *Gain/Loss ∗ Activity* are interaction terms between the trading indicator (*Gain* or *Loss*) and *Literacy* and *Activity*, respectively. *LR test* refers to the log-likelihood ratio test (comparison with the null model). Significance levels are *p<0.1, **p<0.05, and ***p<0.01.

Acknowledgments