Improved Mutual Fund Investment Choice Architecture

Philip W. S. Newall

Technical University of Munich

Katie N. Parker

University College London

Author Details

Dr Philip W. S. Newall, Chair of Marketing, Technical University of Munich, Arcisstrasse 21, 80333 Munich, Germany pnew@tum.de

Dr Katie N. Parker, Department of Experimental Psychology, University College London, 26 Bedford Way London, UK WC1H 0AP katie.parker.11@ucl.ac.uk

Software, Open Access Data and Materials

For open access to the materials and data reported in this paper please visit osf.io/kpz5r.

Abstract

Two choice architecture interventions were explored to debias investors' irrational preference for mutual funds with high past returns rather than funds with low fees. A simple choice task was used involving a direct trade-off between maximizing past returns and minimizing fees. In the first intervention, warning investors that, "Some people invest based on past performance, but funds with low fees have the highest future results" was more effective than three other disclosure statements, including the US financial regulator's, "Past performance does not guarantee future results". The second intervention involved converting mutual fund annual percentage fees into a 10 year dollar cost equivalent. This intervention also improved investors' fee sensitivity, and remained effective even as past returns increased. Financially literate participants were surprisingly more likely to irrationally maximize past returns in their investment choices.

Keywords: personal finance; investing; disclosure; financial literacy; nudging

Instead of having to buy hundreds of individual stocks, investors can form diversified portfolios by just buying a few mutual funds (each containing many stocks). But potential mutual fund investors have to consider many fund attributes. Two of the most important attributes are the fund's past performance and the fund's annual percentage fee (Wilcox, 2003). The US Securities and Exchange Commission (SEC) requires mutual funds to tell potential investors that, "Past performance does not guarantee future results". Although the SEC does not insist on a specific wording, this is the phrase they themselves use (SEC, 2003). This is because the evidence is clear: Past performance does not persist (Carhart, 1997).

Mutual funds, and stocks more generally, are commonly described as obeying a "random walk", where each price movement is independent of what has occurred in the past (Malkiel, 2016). Nonetheless, many investors do incorrectly assume that high past performance is the best guide to high future returns (Greenwood & Shleifer, 2014; Ippolito, 1992). Minimizing fees is, however, an effective strategy for increasing future expected returns for a given level of risk (Malkiel, 2003; Malkiel, 2016), as fees reduce investment returns one-for-one on average (Sharpe, 1991), but many mutual fund investors fail to realize this (Barber, Odean, & Zheng, 2005; Choi, Laibson, & Madrian, 2010). This paper adds to the literature on investor debiasing interventions, by finding two interventions that can help mutual fund investors to minimize fees.

Back of the envelope calculations show that investors' losses are massive. In 2014, over $8 trillion was invested in US equity mutual funds with an average fee of 0.7% a year (Investment Company Institute, 2015), but low-fee index funds with fees of 0.1% a year or lower are available. A 0.1% reduction in average fees paid would save investors $8 billion in a single year. Many investors may not understand the importance of fees, and the irrelevance

of past performance, given the industry's tendency to prominently advertise funds with spectacular past performance (Barber et al., 2005; Jain & Wu, 2000; Koehler & Mercer, 2009).

Investors' preference for maximizing past performance rather than minimizing fees is difficult to debias. Choi et al. (2010) presented a sample of highly-educated and financially literate Harvard staff members in a mutual fund choice task allocating $10,000 between four mutual funds. The funds were selected so that the highest-fee fund had the highest past returns, but the lowest-fee fund maximized future returns. Choi et al.'s participants were highly financially incentivized, each with a maximum performance-related payment differential of $94. Nonetheless, the irrational past return maximizing strategy was extremely resistant to debiasing. The most effective intervention involved converting mutual fund fees into corresponding dollar cost equivalents (e.g., presenting a 1% fee on $10,000 as $100), but this reduced the average fee paid by only $24 from $456 to $432, compared to a minimum potential fee of $310.

Evidence shows that "Past performance does not guarantee future results", the standard form of current financial disclaimers, does not achieve its aims, and can be improved. Two studies showed that investors can be helped toward better investment choices by disclaimers directly mentioning fees (Fisch & Wilkinson-Ryan, 2014; Mercer, Palmiter, & Taha, 2010). However, consider their experimental stimuli:

Mercer et al. (2010) p.445:

"Do not expect the fund's quoted past performance to continue in the future. Studies show that mutual funds that have outperformed their peers in the past generally do not outperform in the future. Strong past performance is often a matter of chance."

Fisch and Wilkinson-Ryan (2014) p.633:

"In making your investment decision, you may want to consider the following information: The most important single factor in mutual fund performance is the fund's operating expenses (in other words, its fees)."

Although both disclaimers tell investors about the importance of fees, these disclaimers are also much longer than the real world status quo. Long disclaimers will have higher implementation costs than short disclaimers. Also, these disclaimers tell investors about the importance of fees, but fail to emphasize this point. It may well be the case that disclaimers with a greater attention to investor psychology could be even more effective. The first experiment in this paper tests the efficacy of three alternative short-form disclaimers:

Fees: "The most important single factor in mutual fund performance is the fund's fees"

Loss: "Investments may go up or down in value, but fund fees represent a permanent loss of wealth"

Social comparison: "Some people invest based on past performance, but funds with low fees have the highest future results"

The fees condition was a shortened version of Fisch and Wilkinson-Ryan's (2014) fee disclaimer. The loss disclaimer was designed to utilize mental accounting (R. Thaler, 1985) and loss aversion (Kahneman & Tversky, 1979). Fund fees are netted against fund performance, so that investors only observe a return net of all fees in their annual statements. It was hypothesized that describing the impact of fees separately from other drivers of fund performance, and framing fees as a loss from the status quo of no fees, may increase fee salience. The social comparison disclaimer was designed to help participants understand why past performance is frequently mentioned (since many investors incorrectly invest based on this metric), and yet nonetheless discount this information (and hence achieve superior

expected returns). The social comparison disclaimer was designed to tap into social comparison, an important driver of behavior (Buunk & Gibbons, 2007). Disclaimers were designed to be as concise as possible, with the three new conditions ranging between 13 and 17 words each. Care was taken to make specific language as similar as possible across all conditions.

The second experiment in this paper directly manipulated the salience of mutual fund fees by framing fees as the total dollar cost over 10 years. Mutual fund fees are charged as "small" annual percentages, usually 2% or less, which nonetheless add up to large losses of wealth over long time horizons. Prior work has explored reframing mutual fund fees as their corresponding dollar cost, but this nudge is only marginally-effective (Choi et al., 2010; Hastings & Tejeda-Ashton, 2008), and is actually counter-productive when the corresponding dollar cost is "small", i.e., $10-$15 (Newall & Love, 2015).

Many investors do not understand how to perform financial math with percentages (Newall, 2016). Many investors think a return sequence of +10%, - 10% = 0. Correctly multiplying these returns yields the answer of = -0.1%. It is highly plausible that investors do the same thing with percentage fees which have been reframed as dollar costs. However, this means the investors will underestimate how the amount of money lost to fees grows as the investment increases in value. Therefore, one solution is to frame fees as the total dollar cost incurred over long time periods, such as 10 years (similar to framing car gas economy as the cost of gas for driving 100,000 miles; Camilleri & Larrick, 2014). This has an advantage, in that a 1% fee will have an increasing 10 year dollar cost fee as mutual fund returns increase (since 1% is taken from a larger account value). For example, a 1% fee corresponds approximately to $149 on an initial balance of $1,000 if after-fee returns are 4%, but

increases to $208 if after-fee returns are 8%. This equilibrating mechanism is hypothesized to keep fees salient under the 10 year dollar cost framing, even when past returns are high.

The SEC does require mutual fund providers to disclose the long-run impact of their fees, but this information is relegated to often-long prospectuses of technical information (SEC, 2017), which many investors are likely to ignore. Many investors are likely to just view the fund's summary information on mutual fund comparison sites, which tend to just include the fund's annual percentage fee. Figure 1 shows an example. This is the first experiment, to our knowledge, to explore the extent to which directly reframing this information as a long-term currency cost can help investors to become more aware of mutual fund fees.

Figure 1. Example of mutual fund information from Yahoo Finance

Two experiments were conducted to test these hypotheses. Experiment 1 compares four mutual fund disclaimers in a simple choice task. Experiment 2 uses the same task to contrast the salience of equivalent percentage and long-term dollar cost fees, and how this changes as past returns increase.

Experiment 1

Method

Participants. Participants aged 18 and over and from the US were recruited (N = 1,003). Participants were recruited from Amazon Mechanical Turk, an online crowdsourcing site commonly used for psychological research and shown to provide more attentive participants than university participant pools (Hauser & Schwarz, 2016; Ramsey, Thompson, McKenzie, & Rosenbaum, 2016). Participants were paid a baseline fee of $0.10, with a further $0.10 of incentive based on their choices in the mutual fund selection task. Although this is a low absolute level of incentive-based pay, it did provide a high relative level of incentives, giving participants a chance to double their total earnings. Previous research shows that even with high financial incentives, most participants incorrectly maximize past returns instead of minimizing fees (Choi et al., 2010). Participants had a mean age of 36.2 years, 59.2% were female, and 52.5% of the sample had at least a college degree.

Although participants were drawn from a general population pool, they were not unfamiliar with financial concepts. Participants had a mean financial literacy score of 8.6 out of 13 (Fernandes, Lynch, & Netemeyer, 2014). Importantly, participants were more financially literate than in Fernandes et al.'s three studies, which had mean financial literacy scores ranging between 7.27 – 7.81.

Procedure. A between-participants design was used, with participants randomly assigned to one of four financial disclaimer conditions. Participants completed the main choice task before providing demographics and other individual difference information. The three treatment disclaimers were tested against a control disclaimer, "Past performance does not guarantee future results".

Materials. Participants were given a short introduction to mutual fund basics and terminology at the beginning of the task. The aim was to give sufficient background to someone with zero prior knowledge, but to keep the task short and interesting:

> Earn a bonus of up to $0.10 on this question. The size of your bonus is related to the average future performance of an investment like the one you choose.
>
> Investors can use mutual funds to invest in the stock market. Stock mutual funds combine the money of many investors, and use this money to buy a portfolio of stocks. Buying a single mutual fund is easier for investors than buying many different individual stocks. Mutual funds charge fees in return for this service. Mutual funds can be assessed over numerous criteria, but there are two key features. A mutual fund's annual percentage fee gets taken out of an investment in the fund every year. Meanwhile, a mutual fund's past performance is the total return that previous investors took home from investing in that fund, after all fees had been subtracted.
>
> Your task is to choose one of these four funds to invest in. Each fund follows a similar strategy, and was launched at the same time ten years ago.

The final sentence informs participants that differences in fund performance are not due to one fund investing in a riskier portfolio than the others – the only reason for expecting past performance differences to extend into the future (Malkiel, 2003; Malkiel, 2016). The prominent placing of this sentence should help inform participants that the funds do not differ

in terms of strategy, average risk level, or assets, and so preferences over these attributes are not relevant to the decision. Performance differences such as these would typically occur due to portfolio allocation differences between various active fund managers, and will not be predictive of future performance (Carhart, 1997). Therefore, participants should normatively ignore the past performance, and instead minimize fees, which is predictive of high future performance (Malkiel, 2003; Malkiel, 2016)

In line with previous mutual fund debiasing experiments, there was a positive association between past returns and fees, providing participants with a dilemma over whether to minimize fees or maximize past returns (Choi et al., 2010; Mauck & Salzsieder, 2017; Newall & Love, 2015). Table 1 shows the four mutual funds participants were asked to choose from, ranging from a fund with fees of 0.5% and past performance of 2%, to a fund with fees of 2% and past performance of 8%. These values were used to tempt participants toward maximizing past performance in the baseline condition, which the treatment disclaimers attempted to debias participants away from.

Table 1

Mutual Funds on Offer in Experiment 1

Fund cues	Fund A	Fund B	Fund C	Fund D
Fees	2.0% a year	1.5% a year	1.0% a year	0.5% a year
Past performance	8.0% a year	6.0% a year	4.0% a year	2.0% a year

Immediately below this table was the financial disclaimer, e.g., “Past performance does not guarantee future results”. Importantly, each of the four disclaimers provides effectively equivalent information to investors: That selection should not be based on past

performance, and therefore should be based on fund fees (the only other cue that varies). Any differences across the four disclaimers must therefore be due to investor psychology.

Measures. The dependent variable was fund selection from the four funds in Table 1. The funds offer a monotonic trade-off between maximizing past returns and minimizing fees. This forced choice paradigm was chosen over the alternative: A hypothetical portfolio allocation task, where money is allocated continuously between the funds on offer (Choi et al., 2010; Fisch & Wilkinson-Ryan, 2014). Hypothetical portfolio allocation tasks may underestimate real world effect sizes if experimental participants respond heuristically because they have less cognitive resources than real world investors. The naïve diversification heuristic would involve putting an equal allocation of cash in each fund (Benartzi & Thaler, 2001), and appears prevalent in similar investment choice tasks (Bateman, Dobrescu, Newell, Ortmann, & Thorp, 2016; Mauck & Salzsieder, 2017). A forced choice paradigm avoids these potential problems. Furthermore, it has been suggested that a failure to minimize mutual fund fees is driven by the naïve diversification heuristic (Mauck & Salzsieder, 2017). Ruling out naïve diversification allows this experiment to see whether additional factors underlie investors' inability to minimize mutual fund fees. Previous research shows that hypothetical investment choices are substantially influenced by the range of investments on offer (Vlaev, Chater, & Stewart, 2007), so having an even number of funds prevents participants from picking a focal central option.

Bonus size varied in $0.02 increments across Table 1, from $0.10 for the fee-minimizing fund to $0.04 for the past return-maximizing fund, reflecting the true negative relationship between fees and future performance.

After the main choice task participants answered a number of individual difference blocks in randomized order. Participants gave demographic information of age, education, and gender. Participants answered a 13-part financial literacy scale (Fernandes et al., 2014), and the short form scale for social comparison orientation (Gibbons & Buunk, 1999). Participants also provided a short measure of loss aversion (Harinck, Van Beest, Van Dijk, & Van Zeeland, 2012).

It was hypothesized that the success of the three new financial disclaimers may depend on relevant individual difference variables (all disclaimers with financial literacy, the loss disclaimer with loss aversion, and the social comparison disclaimer with social comparison orientation). It is important to examine whether specific disclaimers may be especially effective with sub-groups of the population.

Materials and data for the experiments reported in this paper can be accessed from osf.io/kpz5r.

Results

In the experiments that follow, the main dependent variable was the choice of one of four mutual funds, violating the standard assumption of ordinary least squares (where the dependent variable can take on any value). Because the mutual funds on offer have a monotonic relationship between maximizing past performance and minimizing fees, an ordinal logistic regression was used. Ordinal logistic regression utilizes this ordering of the dependent variable and leads to increased statistical power compared to non-ordered models, such as multinomial logistic regression (Liu, 2015).

Table 2 shows the percentage of responses per cell. As can be seen, the percentage of participants minimizing fees almost doubled from 12.3% in the control to 20.7% in the social comparison condition. This is a remarkably effective nudge for such a simple disclaimer, in a choice task that is designed to tempt participants towards the common error of maximizing past returns. The loss disclaimer did not lead to a significant improvement in fee-sensitivity compared to the control condition ($B = 0.01$, $z = 0.11$, $p = .916$, 95% CI [-0.24, 0.26]), but the fees condition did ($B = 0.33$, $z = 2.25$, $p = .024$, 95% CI [0.04, 0.61]) and the social comparison condition also did ($B = 0.59$, $z = 4.29$, $p < .001$, 95% CI [0.32, 0.87]). A post-hoc comparison revealed that the improvement in fee-sensitivity was greater in the social comparison condition than in the fees condition ($B = 0.45$, $z = 2.78$, $p = .006$, 95% CI [0.13, 0.77]). "Some people invest based on past performance, but funds with low fees have the highest future results" was a more effective disclaimer than either the real-world status quo, or the fees disclaimer based on previous experimental debiasing research (Fisch & Wilkinson-Ryan, 2014).

Table 2

Percentage of Responses Per Experimental Cell

Response	Control	Fees	Losses	Social comparison
Fund A (Maximize past performance)	43.1%	32.7%	38.0%	28.4%
Fund B	27.8%	28.2%	28.2%	27.9%
Fund C	16.9%	21.2%	19.0%	23.0%
Fund D (Minimize fees)	12.3%	17.9%	14.7%	20.7%

Note. Disclaimer texts for each condition were as follows. Control: "Past performance does not guarantee future results"; Fees: "The most important single factor in mutual fund

performance is the fund's fees"; Loss: "Investments may go up or down in value, but fund fees represent a permanent loss of wealth"; Social comparison: "Some people invest based on past performance, but funds with low fees have the highest future results".

The ordinal logistic regression depends on the "proportional odds assumption": That the predictors have identical effects across the three transitions between the four outcomes of the dependent variable (Williams, 2006). However, a Brant test showed that this assumption was not valid for either the social comparison (χ^2 (2) = 9.09, p = .011) or fees disclaimer conditions (χ^2 (2) = 25.36, p < .001). Therefore, a partial proportional odds model was fitted where the effect of these two disclaimers was estimated over the three transition probabilities across the dependent variable (Williams, 2006). All three transition probabilities were significant for the social comparison disclaimer (p-values < .01), showing that participants were always more fee-sensitive with this disclaimer compared to the control condition. For the fees disclaimer, the transition probability between Funds A and B was not significant (z = -0.14, p = .887), but the two other transition probabilities were (p-values < .001). This shows that the fees disclaimer was not effective in changing behavior when fees and past performance were extremely high (Funds A and B), but was otherwise effective.

Participants were on average financially literate, with a mean score of 8.6 out of 13. The original regression model was re-run, adding a regression term for standardized financial literacy (see Model 2 in Table 3). Interestingly, there was a statistically significant negative relationship between financial literacy and fee-sensitivity (B = -0.18, z = -3.04, p = .002, 95% CI [-0.29, -0.06]). More financially-literate participants were more likely to choose the mutual fund with high past returns and high fees. Analysis of marginal effects showed that a one standard deviation increase in financial literacy was associated with a four percentage

point increase in probability of choosing the highest-fee fund. This relationship remained statistically significant after controlling for demographic variables of age, education, and gender (Model 3 in Table 3).

Table 3

Ordinal Logistic Regression Predicting Fee Sensitivity in Mutual Fund Choice
Dependent variable: Mutual fund choice

	M1	M2	M3
Social comparison	0.60***	0.62***	0.62***
Fees framing	0.33*	0.35*	0.34*
Loss disclaimer	0.01	0.02	0.01
Financial literacy		-0.18**	-0.13*
Age			-0.01
Education			0.41
			0.52
			0.33
Gender			0.20

Significance levels: *p < .05 **p < .01 ***p < .001.

Participants also had a mean social comparison index of 18.6 out of 30, and a mean loss aversion parameter of 6.2. These individual difference parameters were interacted with the three treatment conditions to test for ex-ante plausible individual differences. Separate regression models added the main effect of the individual difference parameter and an interaction term between the parameter and treatment condition. Specifically, the three treatment conditions were each separately interacted with financial literacy, social comparison with social comparison orientation, and the loss treatment with loss aversion. A natural logarithm transformation was first performed on the loss aversion parameter to reduce positive skewness. None of these models' interaction terms were statistically significant,

suggesting that the effectiveness of these financial disclaimers does not vary between individuals on the basis of these measures.[1] The only marginally significant interaction term was between the social comparison condition and social comparison orientation ($B = 0.23$, $z = 1.79$, $p = .073$, 95% CI [-0.02, 0.48]). This suggests a potential trend where this disclaimer may work even better in individuals with high social comparison orientation, and provides some evidence for the hypothesized mechanism via social comparison. However, this does not affect the results on a practical level, since this was already the most effective disclaimer across all participants.

Experiment 1b

It could be argued that the social comparison disclaimer only worked because of a demand effect, i.e., that this disclaimer made it clearer to participants which mutual fund would maximize their bonus payment. A non-incentivized replication was designed to test this explanation.

Method

Another 761 participants were recruited from the same pool (no participant took part in more than one experiment), and were randomly assigned to one of the control, fees, or social comparison disclaimer conditions. The loss condition was not used since this disclaimer was not more effective than the control condition in Experiment 1. Individual difference measures were not collected on this occasion. Participants were compensated with a baseline payment of $0.20, with no incentive-based payment, thus eliminating any potential demand effects from the incentive payment.

[1] Statistical significance of interaction models in non-linear probability models can be difficult to interpret, so it is recommended to also check the significance of interaction effects under ordinary least squares (Ai & Norton, 2003). Statistical significance of all interaction effects reported in this paper was unaffected under ordinary least squares, and so will not be reported.

Results

The social comparison disclaimer increased the percentage of fee-minimizing responses from 13.2% in the control condition to 21.0%, as seen in Table 4. The results of an ordinal linear regression showed that the social comparison disclaimer was again effective compared to the control disclaimer (B = 0.51, z = 3.23, p = .001, 95% CI [0.20, 0.82]) but the fees disclaimer was not on this occasion (B = 0.17, z = 1.04, p = .298, 95% CI [-0.15, 0.50]). The social comparison disclaimer was again more effective than the fees disclaimer (B = 0.32, z = 1.97, p = .049, 95% CI [0.01, 0.64]).

Table 4

Percentage of Responses Per Experimental Cell in Experiment 1b

Response	Control	Fees	Social comparison
Fund A (Maximize past performance)	45.4%	37.1%	32.1%
Fund B	22.2%	22.4%	22.3%
Fund C	19.3%	22.7%	24.7%
Fund D (Minimize fees)	13.2%	17.9%	21.0%

Results of a Brant test showed that the parallel lines assumption held for the social comparison disclaimer (χ^2 (2) = 1.79, p = .408) but not for the fees disclaimer (χ^2 (2) = 11.40, p = .003). Results of a partial proportional odds model showed that the fees disclaimer was effective for the lowest levels of fees and past performance between Funds C and D (z = 3.12, p = .002), but was not otherwise effective (p-values > .292). On a practical level these results confirm the incentivized experiment: The social comparison disclaimer worked best out of those tested.

Experiment 1c

It could further be argued that the social comparison disclaimer worked best out of those tested because it gave overly strong guidance to investors. A very high fee fund might, by chance, outperform the market and its rivals for a long time. It might be argued that the disclaimer should relate to the performance of funds on average, rather than reflecting an exact prediction for any given fund. Therefore a replication experiment was conducted, to see if the following weaker social comparison disclaimer remained the most effective nudge:

"Some people invest based on past performance, but on average funds with low fees have the highest future results." (Adding the words "on average".)

Method

Another 754 participants were recruited from the same pool, and were randomly assigned to one of the control, fees, or social comparison (alternative) disclaimer conditions. Participants were rewarded as in Experiment 1, with $0.10 of baseline payment, and up to $0.10 of performance-based payment depending on their choice from the four funds. Individual difference measures were again not collected on this occasion.

Results

The social comparison disclaimer increased the percentage of fee-minimizing responses from 7.6% in the control condition to 17.9%, as seen in Table 5. The results of an ordinal linear regression showed that the social comparison disclaimer was again effective compared to the control disclaimer ($B = 0.91$, $z = 5.63$, $p < .001$, 95% CI [0.60, 1.23]) as was the fees disclaimer this time ($B = 0.46$, $z = 2.84$, $p = .005$, 95% CI [0.14, 0.78]). But the social comparison disclaimer was again more effective than the fees disclaimer ($B = 0.42$, $z = 2.54$, $p = .011$, 95% CI [0.10, 0.74]).

Table 5

Percentage of Responses Per Experimental Cell in Experiment 1c

Response	Control	Fees	Social comparison (alternative)
Fund A (Maximize past performance)	46.9%	42.0%	27.1%
Fund B	31.3%	22.8%	25.3%
Fund C	14.2%	13.6%	29.7%
Fund D (Minimize fees)	7.6%	21.6%	17.9%

Results of a Brant test showed that the parallel lines assumption held for the social comparison disclaimer (χ^2 (2) = 3.28, p = .194) but not for the fees disclaimer (χ^2 (2) = 14.05, p < .001). Results of a partial proportional odds model showed that the fees disclaimer was not effective for the highest levels of fees and past performance between Funds A and B (z = 1.39, p = .164). As can be seen from Table 5, only 27.1% of participants in the social disclaimer condition maximized past returns (and fees), but 42.0% in the fees condition did (compared to 46.9% in the control condition). Despite the weaker wording, the social comparison disclaimer still the most effective out of those tested.

Experiment 2

Experiment 2's goal was to discover whether part of investors' reluctance to minimize fees is due to a misunderstanding of the long-run impact of small annual percentage fees around 0.5% to 2%. Fees were either framed as percentages, or as the equivalent 10 year currency cost that an initial investment of $1,000 would have incurred. Previous experiments converting a mutual fund fee into its corresponding *annual* currency cost showed limited ability to increase fee-sensitivity (Choi et al., 2010; Newall & Love, 2015). However, it was

hypothesized that the 10 year currency cost frame could increase the salience of fees, even when past returns were very high.

Method

Experiment two was a 2x2 within-subjects experiment, manipulating fee framing (percentages, long-term dollar cost) and past returns (low, high). A within-subjects design was used to maximize power, and because no demand effects were expected (the key fee framing manipulation is not obvious). The experiment was otherwise very similar to Experiment 1.

Participants. Experiment 1's participants were collected from a general population sample who, despite high average financial literacy scores, may still have been unfamiliar with investing decisions. Therefore, two versions of Experiment 2 were run.

In Experiment 2a, another 503 US-based participants were recruited from Amazon Mechanical Turk (no participant took part in more than one experiment). Participants had a mean age of 34.5 years, 46.7% had at least a college degree, and 57.1% were female. Participants had a mean financial literacy score of 8.3 out of 13. The payment scheme in Experiment 2a was the same as in Experiment 1.

In Experiment 2b, US-based participants were recruited from another online crowdsourcing site, Prolific Academic (N = 501), and were given a baseline payment of just over a dollar. To qualify for this experiment, participants must have previously responded "yes" to the question, "Have you ever made investments (either personal or through your employment) in the common stock or shares of a company?" Participants had a higher mean financial literacy score of 9.6, suggesting that this sample was indeed more representative of individual investors. Participants had a mean age of 33.4 years, 60.1% had at least a college degree, and 37.5% were female. One randomly selected participant was incentivized based on

one randomly selected trial with a simulated 10 year $100 stock market investment. This investment had a normally-distributed annual return of mean (8% - fund fee) and standard deviation of 20%, roughly corresponding to US stock market averages (Dimson, Marsh, & Staunton, 2009).

Materials. Participants received the same materials as in Experiment 1 (with the control disclaimer, "Past performance does not guarantee future results"), but with the following changes. The bonus instruction stated that they would be paid on the basis of one question. The sentence, "Each fund has charged the same annual percentage fee each year." was added at the end of experimental instructions, as this fact is no longer clear in the long-term dollar cost conditions. (This statement could, however, plausibly be misunderstood by participants to mean that each fund charged exactly the same fee. Eliminating any confusion caused by this would likely strengthen the results.)

Table 6 shows the combinations of fees and past performance in each condition. Identical percentage fees were used for funds A-D in every condition. Past performance either varied monotonically from 8% to 2% a year in the low past returns conditions, or from 16% to 2% a year in the high past returns conditions. Notice that doubling past performance from 8% to 16% nearly doubles the 10-year dollar cost impact of a 2% annual fee from $435 to $822. Dollar cost fees were described as "Fees over 10 years*" with clarification below the table, "*Total fees based on $1,000 investment over 10 years."

Table 6

Fund Menus in Each Condition in Experiment 2

Condition	Fund cues	Fund A	Fund B	Fund C	Fund D
Percentages, low returns	Fees	2.0% a year	1.5% a year	1.0% a year	0.5% a year
	Past Performance	8.0% a year	6.0% a year	4.0% a year	2.0% a year
Currency cost, low returns	Fees	$435	$270	$149	$61
	Past Performance	8.0% a year	6.0% a year	4.0% a year	2.0% a year
Percentages, high returns	Fees	2.0% a year	1.5% a year	1.0% a year	0.5% a year
	Past Performance	16.0% a year	8.0% a year	4.0% a year	2.0% a year
Currency cost, high returns	Fees	$822	$319	$149	$61
	Past Performance	16.0% a year	8.0% a year	4.0% a year	2.0% a year

Note: Currency cost fees framed as the, "Total fees based on $1,000 investment over 10 years."

Dollar cost fees were calculated as the simple arithmetic difference in compounding between a fund's past performance and a fund without any fees at all. For example, the fund with past returns of 8% a year and fees of 2% a year was calculated as the difference in final value between $1,000 at 10% a year and 8% a year. This calculation underestimates the impact of fees, since it neglects per-period compounding between the return and the fee (the fee reduces the value of the returns in addition to the principal). True dollar costs will therefore be higher than in Table 6. These conservative calculations were chosen in case

investors attempted to back out the percentage fee in the dollar cost condition. This is also a conservative test of the experimental hypothesis.

Measures. The same demographic questions and financial literacy scale as in Experiment 1 were collected after participants chose between the four mutual funds in each condition.

Results

Table 7 shows the percentage of responses per experimental cell. The number of participants minimizing fees more than doubled from 6.3% to 13.1% across the most extreme high returns conditions as fees were reframed from annual percentages to long-term dollar cost in Experiment 2a. Results were almost identical across Experiments 2a and 2b. A comparison of cell-by-cell responses in Table 7 illustrates that the Experiment 2b's investor sample were actually slightly more likely to irrationally maximize past returns than Experiment 2a's general population sample. Below we analyze the results of Experiment 2b only (investor sample), as the same pattern of results was also found in Experiment 2a.

Table 7

Percentage of Responses Per Experimental Cell in Experiment 2

Response	Low returns		High Returns	
	Percentages	Currency Cost	Percentages	Currency
Fund A (Maximize past performance)	43.6% (50.3%)	32.6% (37.0%)	56.8% (67.9%)	36.9% (38.5%)
Fund B	31.2% (29.6%)	32.4% (32.8%)	26.7% (21.4%)	32.3% (32.6%)
Fund C	15.0% (12.8%)	19.6% (18.3%)	10.2% (7.1%)	17.7% (17.6%)
Fund D (Minimize fees)	10.2% (7.3%)	15.4% (11.9%)	6.3% (3.6%)	13.1% (11.3%)

Note. The first percentage in each cell refers to Experiment 2a (general population sample), and the second percentage in parentheses refers to Experiment 2b (investor sample). Results

across both experiments are very similar, with the investor sample having a slightly higher tendency to irrationally maximize past performance.

An ordinal logistic regression was again used to analyze Experiment 2b's results (see Table 8 for regression model coefficients). Main effects of fee framing (B = 0.86, z = 10.99, p < .001, 95% CI [0.71, 1.01]) and past returns level were significant (B = -0.36, z = -5.92, p < .001, 95% CI [-0.48, -0.24]). Results of a Brant test showed that the parallel lines assumption held for past returns level (χ^2 (2) = 2.43, p = .297) but not for fee framing (χ^2 (2) = 38.17, p < .001). Results of a partial proportional odds model showed that reframing fees as a 10 year currency cost was effective at all levels of fees and past performance (p-values < .001).

The model was improved by the addition of an interaction effect (B = -0.68, z = -5.48, p < .001, 95% CI [-0.92, - 0.44]), meaning that the effect of fee framing became stronger in the high- than low-fees condition (Model 2 in Table 8). The significant interaction term means that there was a larger improvement in fee-sensitivity from long-term dollar cost framing in the high- than low-past returns condition. Long-term dollar cost framing became if anything more effective as past returns increased.

The regression model was rerun with the addition of a control for standardized financial literacy score (Model 3 in Table 8). As in Experiment 1, there was a negative correlation between financial literacy and fee-sensitivity (B = -0.31, z = -4.63, p < .001, 95% CI [-0.44, -0.18]). Significance of this result was again unaffected by inclusion of the other demographic variables (Model 4 in Table 8). Analysis of marginal effects showed that a one standard deviation increase in financial literacy was associated with a seven percentage point increase in probability of choosing the highest-fee fund.

Table 8

Ordinal Logistic Regression Predicting Fee Sensitivity in Mutual Fund Choice

Dependent variable: Mutual fund choice				
	M1	M2	M3	M4
Fees framing	0.86***	0.54***	0.55***	0.55***
Past returns	-0.36***	-0.06	-0.07	-0.07
Fees x past returns		-0.68***	-0.70***	-0.70***
Financial literacy			-0.33***	-0.30***
Age				-0.01
Education				-0.43
				-0.50
				-0.38
Gender				0.18

Significance levels: *p < .05 **p < .01 ***p < .001.

Discussion

This paper adds to a growing experimental literature on potential mutual fund investment disclosure interventions. Two main avenues of investigation in this literature have been the strengthening of disclosure statements (Fisch & Wilkinson-Ryan, 2014; Mercer et al., 2010) and the reframing of mutual fund fees (Choi et al., 2010; Hastings & Tejeda-Ashton, 2008; Newall & Love, 2015), in addition to more general interventions for helping investors (Bateman et al., 2016; Beshears, Choi, Laibson, & Madrian, 2011; Koehler & Mercer, 2009).

This paper demonstrates that social comparison disclaimers, e.g., “Some people invest based on past performance, but funds with low fees have the highest future results”, and the long-term dollar cost framing of fees can prove effective as mutual fund choice architecture features compared to the real world status quo and previously explored alternatives. Future research should explore the inferences that investors draw from disclosure statements. It

could be that the baseline disclaimer, "Past performance does not guarantee future results" nonetheless implies some positive link between past and future performance, and that disclaimers work best by fully breaking this link *and* providing investors with the true link between fees and future returns. That could plausibly explain why the social comparison disclaimer worked best out of those tested. Further investigation of this topic could help to maximize the gains from this relatively simple nudge for mutual fund investors.

Although the effects found in this paper were statistically significant, bias did not disappear even in this simplified choice problem. Selecting mutual funds appears to be a daunting and challenging task for many investors, who are prone to be misled by distracting past returns (Greenwood & Shleifer, 2014; Ippolito, 1992). Naïve diversification has also been put forward as an explanation for this bias (Mauck & Salzsieder, 2017). These experiments show that many investors err even when diversification between different funds is not possible. The low absolute level of financial incentives used in this paper may not be to blame, as participants maximize past returns in the previous literature even with much higher incentives (Choi et al., 2010). The difficulty that participants have in minimizing fees in experimental mutual fund tasks highlights both the importance of further research on this topic, and that real world investors require improved mutual fund choice architecture to choose wisely.

Default options have been much studied as a behavior change tool (Johnson & Goldstein, 2003). Although well-chosen defaults can have many beneficial effects for personal investors (R. H. Thaler & Benartzi, 2004), defaults raise important ethical debates (Smith, Goldstein, & Johnson, 2013). Potential ethical issues surrounding default options are one reason why research should continue to explore a range of methods to help debias investors, as in this paper.

The negative correlation between financial literacy and fee-sensitivity was unexpected. However, similar results have been found before in mutual fund choice tasks (Koehler & Mercer, 2009; Pontari, Stanaland, & Smythe, 2009). Because the financial literacy scale was performed after the main task, it is possible that the main task may have affected latter performance on the scale. However, the negative correlations remain even after removing the three mutual fund questions from the financial literacy scale, suggesting this is not the case. Although this was only correlational evidence, it occurred over both experiments, and should be interpreted in the context of a wider body of evidence showing that interventions for increasing financial literacy have little average effect on improving financial behaviors (Fernandes et al., 2014). This suggests that effective choice architecture manipulations should be cost effective at changing behavior compared to direct financial literacy education campaigns.

Improving access to financial advice is another plausible method for improving financial behavior. However, many financial advisers mistakenly maximize high past returns to an even greater degree than their clients (Linnainmaa, Melzer, & Previtero, 2016). Linnainmaa et al. also found a positive correlation between investing experience and maximizing past returns amongst individual investors. These results are consistent with the findings across Experiments 2a and 2b, where the investor sample in Experiment 2b were slightly more likely to maximize past returns than the general population sample (Table 6). This helps provide reassurance that Experiment 1's results, based on a general population sample, should still be valid for the sub-set of the population currently making active financial decisions.

Choice architecture interventions for helping mutual fund investors to minimize fees are unlikely to be voluntarily enacted by the industry, and so this is an area where

behaviourally-informed public policy may be required (Erta, Hunt, Iscenko, & Brambley, 2013; Sunstein, 2016). The large gaps between normative fee-minimizing behavior and control condition choices in this simplified investment choice task suggest that real world investors are in desperate need of clearer guidance than, “Past performance does not guarantee future results”.

References

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters, 80*(1), 123-129.

Barber, B. M., Odean, T., & Zheng, L. (2005). Out of sight, out of mind: The effects of expenses on mutual fund flows. *The Journal of Business, 78*(6), 2095-2120.

Bateman, H., Dobrescu, L. I., Newell, B. R., Ortmann, A., & Thorp, S. (2016). As easy as pie: How retirement savers use prescribed investment disclosures. *Journal of Economic Behavior & Organization, 121*, 60-76.

Benartzi, S., & Thaler, R. H. (2001). Naive diversification strategies in defined contribution saving plans. *American Economic Review, 91*(1), 79-98.

Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. (2011). How does simplified disclosure affect individuals' mutual fund choices? In D. Wise A. (Ed.), *Explorations in the economics of aging* (pp. 75-96). Chicago: University of Chicago Press.

Buunk, A. P., & Gibbons, F. X. (2007). Social comparison: The end of a theory and the emergence of a field. *Organizational Behavior and Human Decision Processes, 102*(1), 3-21.

Camilleri, A. R., & Larrick, R. P. (2014). Metric and scale design as choice architecture tools. *Journal of Public Policy & Marketing, 33*(1), 108-125.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance, 52*(1), 57-82.

Choi, J. J., Laibson, D., & Madrian, B. C. (2010). Why does the law of one price fail? an experiment on index mutual funds. *The Review of Financial Studies, 23*(4), 1405-1432.

Dimson, E., Marsh, P., & Staunton, M. (2009). *Triumph of the optimists: 101 years of global investment returns*. New Jersey, Princeton: Princeton University Press.

Erta, K., Hunt, S., Iscenko, Z., & Brambley, W. (2013). Applying behavioural economics at the financial conduct authority.

Fernandes, D., Lynch, J., G., & Netemeyer, R., G. (2014). Financial literacy, financial education, and downstream financial behaviors. *Management Science, 60*(8), 1861-1883.

Fisch, J. E., & Wilkinson-Ryan, T. (2014). Why do retail investors make costly mistakes? an experiment on mutual fund choice. *University of Pennsylvania Law Review, 162*(3), 605-648.

Gibbons, F. X., & Buunk, B. P. (1999). Individual differences in social comparison: Development of a scale of social comparison orientation. *Journal of Personality and Social Psychology, 76*(1), 129-142.

Greenwood, R., & Shleifer, A. (2014). Expectations of returns and expected returns. *Review of Financial Studies, 27*(3), 714-746.

Harinck, F., Van Beest, I., Van Dijk, E., & Van Zeeland, M. (2012). Measurement-induced focusing and the magnitude of loss aversion: The difference between comparing gains to losses and losses to gains. *Judgment and Decision Making, 7*(4), 462-471.

Hastings, J. S., & Tejeda-Ashton, L. (2008). Financial literacy, information, and demand elasticity: Survey and experimental evidence from mexico (working paper no. w14538). Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w14538;;

Hauser, D. J., & Schwarz, N. (2016). Attentive turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods, 48*(1), 400-407.

Investment Company Institute. (2015). 2015 investment company fact book. Retrieved from https://www.ici.org/pdf/2015_factbook.pdf

Ippolito, R. A. (1992). Consumer reaction to measures of poor quality: Evidence from the mutual fund industry. *The Journal of Law & Economics, 35*(1), 45-70.

Jain, P. C., & Wu, J. S. (2000). Truth in mutual fund advertising: Evidence on future performance and fund flows. *The Journal of Finance, 55*(2), 937-958.

Johnson, E. J., & Goldstein, D. G. (2003). Do defaults save lives? *Science, 302*, 1338-1339.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263-292.

Koehler, J. J., & Mercer, M. (2009). Selection neglect in mutual fund advertisements. *Management Science, 55*(7), 1107-1121.

Linnainmaa, J. T., Melzer, B. T. & Previtero, A. (2016). The misguided beliefs of financial advisors. Retrieved from http://sfs.org/wp-content/uploads/2016/05/MisguidedBeliefs20160407.pdf

Liu, X. (2015). *Applied ordinal logistic regression using stata*. Thousand Oaks, CA: SAGE Publications.

Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *Journal of Economic Perspectives, 17*(1), 59-82.

Malkiel, B. G. (2016). *A random walk down wall street: The time-tested strategy for successful investing* (11th ed.). New York: WW Norton & Company.

Mauck, N., & Salzsieder, L. (2017). Diversification bias and the law of one price: An experiment on index mutual funds. *Journal of Behavioral Finance, 18*(1), 45-53.

Mercer, M., Palmiter, A. R., & Taha, A. E. (2010). Worthless warnings? testing the effectiveness of disclaimers in mutual fund advertisements. *Journal of Empirical Legal Studies, 7*(3), 429-459.

Newall, P. W. S. (2016). Downside financial risk is misunderstood. *Judgment and Decision Making, 11*(5), 416-423.

Newall, P. W. S., & Love, B., C. (2015). Nudging investors big and small toward better decisions. *Decision, 2*(4), 319-326.

Pontari, B. A., Stanaland, A. J., & Smythe, T. (2009). Regulating information disclosure in mutual fund advertising in the united states: Will consumers utilize cost information? *Journal of Consumer Policy, 32*(4), 333-351.

Ramsey, S. R., Thompson, K. L., McKenzie, M., & Rosenbaum, A. (2016). Psychological research in the internet age: The quality of web-based data. *Computers in Human Behavior, 58*, 354-360.

SEC. (2003). Final rule: Amendments to investment company advertising rules. Retrieved from https://www.sec.gov/rules/final/33-8294.htm

SEC. (2017). Mutual funds and exchange-traded funds (ETFs) – A guide for investors. Retrieved from https://www.sec.gov/reportspubs/investor-publications/investorpubsinwsmfhtm.html#Fees

Sharpe, W. F. (1991). The arithmetic of active management. *The Financial Analysts' Journal, 47*(1), 7-9.

Smith, N. C., Goldstein, D. G., & Johnson, E. J. (2013). Choice without awareness: Ethical and policy implications of defaults. *Journal of Public Policy & Marketing, 32*(2), 159-172.

Sunstein, C. R. (2016). The council of psychological advisers. *Annual Review of Psychology, 67*, 713-737.

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science, 4*(3), 199-214.

Thaler, R. H., & Benartzi, S. (2004). Save more tomorrow™: Using behavioral economics to increase employee saving. *Journal of Political Economy, 112*(S1), S164-S187.

Vlaev, I., Chater, N., & Stewart, N. (2007). Relativistic financial decisions: Context effects on retirement saving and investment risk preferences. *Judgment and Decision Making, 2*(5), 292-311.

Wilcox, R. T. (2003). Bargain hunting or star gazing? investors' preferences for stock mutual funds. *The Journal of Business, 76*(4), 645-663.

Williams, R. (2006). Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *Stata Journal, 6*(1), 58-82.