# EXTREME TIME DISCOUNTERS AND HEALTH, INSURANCE, AND INVESTMENT CHOICES 

Ian Ayres,* Gregory Conyers,** and Darren Segal***


#### Abstract

:

Using a survey taken by a random sample of Discovery customers in South Africa $(\mathrm{N}=3,558)$, we find that time discounting characteristics help to predict heterogeneity in health, insurance, and investment choices. Utilizing a new dataset of insurer-collected data, we find that, controlling for demographic variables, those with higher discount rates are more likely to smoke, are more likely to never exercise, are less involved in Discovery's health promotion program, and purchase less healthy food. They are also less likely to have life insurance, drive less safely, and are more likely to make risky investments. Unlike most previous work, we use discount rate bins to model discounting effects. We find that a dummy for extreme discounters-those who discount at $100 \%$ over a year-does the bulk of the predictive work. We conclude that asking a single question to deduce if an individual is an extreme discounter is nearly as good as more complex elicitation methods.


Keywords: Intertemporal discounting, impatience, present bias, South Africa
JEL Classifications: C, D, I, J

[^0]
## I. Introduction

Variously called "impatience" (Ashraf, Karlan and Yin 2006), "present bias" (Meyer and Sprenger 2010) or having high "discount rates" (Becker and Milligan 1997), the inability to sacrifice today in order to receive a delayed benefit goes by many names in the psychology and behavioral economics literature. The rate at which an individual discounts a benefit they will accrue at some period in the future can help elucidate an individual's ability to put in work or sacrifice a gain today to benefit from that investment at a later time.

We elicit the discount rates of 3,558 South African customers of the insurance company Discovery in an online survey. Discovery customers are an affluent subset of South Africans roughly comparable to any first-world country. ${ }^{1}$ We use a Multiple Price List (MPL) to determine discount rates, following those like Kirby and Bickel (1999) and Chabris et. al. (2008). In a series of questions, respondents are asked to choose between a R15,000 (about $\$ 1,500$ ) reward in a year and a smaller reward immediately. We use the smallest amount they are willing to take today over the later fixed payment in the calculation of their individual discount rate. Using these rates and a robust set of individual covariates, we estimate the predictive power of discount rates on 15 real-world outcomes. Each of these outcomes either directly or indirectly involve decisions about sacrifices today for benefits in the future, including health investments, insurance investments, driving behaviors, and financial investments.

To our knowledge, this study uses the largest sample size to date comparing discount rates from the lab-here, an online survey-to multiple real-world outcomes. Our 3,558 complete responses exceed the number in all previously published surveys, though we do not have a full set of outcomes for more than half outcome variables. Chabris et. al. (2008) compared discount rates to a variety of real-world outcomes including body mass index (BMI), smoking, exercising, and eating healthy foods. In their largest data set $(\mathrm{N}=326)$, they find that discount rates significantly predict only BMI. In their smaller sample ( $\mathrm{N}=126$ ), they find that discount rates predict BMI, smoking, and exercise frequency. Komles, Smith, and Bogin (2004) also predict BMI with discount rates. Other studies have found significant relationships between discount rates and substance use decisions like smoking (Bickel, Odum, and Madden 1999; Mitchell 1999; Harrison, Lau, and Rutstrom 2010), heroin use (Kirby, Petry, and Bickel 1999), and alcohol use (Petry 2001). Sutter et. al. (2011) find that discount rates predict saving, smoking, alcohol consumption, and BMI in German adolescents ( $\mathrm{N}=638$ ). Maier and Sprenger (2010) find that discount rates predict credit card debt for Boston residents ( $\mathrm{N}=606$ ).

This study improves on previous studies in two ways. First, in addition to expanding the sample size for previously studied variables, we also expand the variable

[^1]set under scrutiny. Most notably, we add the decision to purchase life insurance, a driving safety score produced by in-car accelerometers and global positioning systems (GPSs), and investment choices like prospective fund return and fund volatility. We find that, controlling for sex, age, education, employment status, and race, individuals with higher discount rates are much less likely to enroll in life insurance, drive less safely, and make riskier investment decisions.

Second, we explore the possibility of reducing discounting to a binary measure: extreme discounters v . non-extreme discounters. MPLs are clumsy and the intuition behind them relatively obvious, perhaps biasing results. We define extreme discounters to be those who would be willing to take half the money today that they would receive in a year (discount rate $=100 \%$ ). In our MPL, this is R7,500 today (about $\$ 750$ ) versus R15,000 in a year (about $\$ 1,500$ ). We find that most of the negative activities associated with discounting are either only present for extreme discounters or most pronounced for extreme discounters. This opens the possibility of a rough one-time elucidation of discount rates by asking one question: would you rather $\$ \mathrm{X}$ today or $2 * \$ \mathrm{X}$ in a year? This allows a simple and widely applicable determination of discount rate, and one that we show produces significant statistical power. ${ }^{2}$

We understand that uncovering an individual's discount rate using one methodology at a single point in time is an incomplete metric. Relying on the work of many others, Frederick, Loewenstein and O’Donoghue (2002) survey literature showing that many factors affect an individual's response to discount rate questions. The time horizon over which the tradeoff is posed is vitally important. Small monetary rewards are discounted more than large monetary rewards. Zauberman et. al. (2009) find that asking respondents to estimate the time it takes to complete common tasks before answering questions about time discounting lowers discount rates. These important caveats, however, do not lower interest in the predictive power of simple measures of discount rates.

It is important to note that all of these results are correlations, and do not imply causation. However, there are important theoretical reasons to believe that discount rates and choices about distribution of benefits over time are causally related. Grossman (1972) first posited a model for the demand of what he called "good health" over time. Investments now will have payoffs in the form of good health later. Individuals who are less willing to give up something today for a return tomorrow will be less likely to invest in good health, causing lower up-front health investments. This theory suggests that the connection between discount rates and health-related outcomes has a causal component, though there are also clear endogeneity problems related to personal wealth and other factors (Becker and Milligan 1997).

Several of our other non-health related outcomes can also be framed as goods that produce returns over time. The choice to buy life insurance, though perhaps more strongly related to risk preferences, is a stylized form of investment: an individual puts in

[^2]money today for the possibility of a large payoff tomorrow (for dependents or family). However, impatience and risk aversion have been shown to be negatively correlated: the more impatient you are on average-the higher your discount rate-the less willing you are to take risks. ${ }^{3}$ This would suggest that impatient people would be more likely to buy insurance, while we notably find the opposite effect. Driving safely is also a form of investment: by giving up the thrill of driving dangerously or quickly, you improve your long-term health. Investment choices might also be related directly to present bias: impatient people might pick funds with higher return and higher volatility in order to make gains in the short run while putting long-term investment returns at risk.

The paper proceeds as follows. Section II outlines our data and methodology. Section III presents the results of our discount rate correlations. Section IV concludes.

## II. Data and Methodology

During November 2013 we sent 60,000 emails to a random subset of Discovery customers soliciting responses to an online survey in return for the possibility of winning a coffee maker or iPod of equivalent value. 5,593 people clicked through to the survey. The survey exposed them to a Multiple Price List (see Table 1) to elucidate discount rates. In a MPL respondents choose between a theoretical monetary reward today and a larger reward in the future. Our MPL included 11 questions that keep the larger, later reward fixed and vary the reward to be awarded today. The proximate awards vary from R15,000 (the same as the later reward, implying a $0 \%$ discount rate) to R7,500 (half the later reward, implying a $100 \%$ discount rate). The survey then poses a series of questions that went into the construction of four attitudinal indices (impulse control, industriousness, positive affect, and integrity), and finally the respondents are asked to respond to a series of demographic questions. ${ }^{4}$, 1,358 people did not complete the survey, with most dropping out without completing the MPL. Of the people who did complete the survey, 677 responses were discarded because these respondents gave irrational answers on the MPL indicating that they did not take the task seriously. We classified participants as irrational for one of two reasons. First, we classified as irrational those that indicated they would prefer to take R15,000 in a year rather than R15,000 today. We also classified respondents as irrational if they were inconsistent in their answers, switching from column A to B (the later sum of money) and then back to A. 3,558 participants remain.
[Table 1 about here.]

[^3]The possibility of selection bias is inherent in a voluntary survey. For example, those who are more risk-averse-our survey offered the possibility of winning a prize as an incentive-or have more present bias might be less likely to complete the survey. The possibility of this kind of selection effect, however, does not necessarily bias our results. It simply reduces our statistical power with regards to participants with high discount rates.

The demographic controls include sex, age, race, employment status, education, and marital status. Summary statistics are shown in the online appendix's Table A. The four attitudinal indices (impulse control, industriousness, positive affect, and integrity) were constructed from a set of 19 general statements that respondents were asked to agree or disagree with on a 1-10 scale. Examples of these statements include "I spend more money than I would like to" and "As a whole, my work and home environments are stressful." ${ }^{5}$ We use these indices only in regressions presented in the online appendix Table C.

The survey included questions about height and weight which we transformed into Body Mass Index (BMI), as well as smoking habits, workout habits, and saving rates. We expand these traditional, self-reported outcomes of interest by combining objective data collected by Discovery with these self-reported outcomes. Unfortunately, we cannot cross-reference these self-reported outcomes because they do not overlap with the objective outcomes. These Discovery health-related outcomes are wide-ranging. They include the extent to which customers are involved in Discovery's Vitality program, which seeks to help customers live more healthy lives through a variety of behavioral economics-influenced incentive processes. ${ }^{6}$ For those who are enrolled in a healthy foods program that offers a rebate for healthy food purchases, we know the proportion of food that they buy that is healthy based off of grocery store receipt data automatically collected by Discovery. Healthy foods include fruits and vegetables and non-fat dairy products. Unhealthy foods include those that are high in saturated fats, trans-fatty acids, sugar, salt, or refined starch (Sturm et. al. 2013). Customers can also enroll in a gym membership at reduced cost through Discovery, and we know if they have elected to do so or not. These Discovery-provided health-related outcomes supplement the selfreported health outcomes.

Discovery has also provided us with anonymous data on the insurance, credit, and driving decisions of survey participants. We know if participants have purchased life

[^4]insurance from Discovery. ${ }^{7}$ We know if participants have a credit card with Discovery, and if they do if they have ever defaulted on their card. For a subset of customers, Discovery has installed a telematics device in their cars that provides an objective measure of driving quality, a Driver Score, from 0 to 750 with 750 being the best score. ${ }^{8}$ Finally, Discovery has provided us with the investment choices of those customers with investments. We know the relative return, the volatility, and the Total Expense Ratio (TER), a measure of fund cost, for survey participants. All of these numbers are available to customers when they are choosing funds, so, like the other variables, we can directly model the decision to pick one fund over another. ${ }^{9}$

Table 2 presents the summary statistics for the discount rates and outcome variables for our sample. The number of observations deviates from 3,552 for the outcome variables where the client does not have the respective product with Discovery.
[Table 2 about here.]
Importantly, we use a series of dummies for participants' discount rates rather than a continuous measure. The average discount rate for our sample is $29 \%$ over a year. This is consistent with previous studies that have used MPLs. Harrison, Lau, and Rutstrom (2010), for example, find average discount rates to be between 23 and $30 \%$ for 252 Danish participants. We do this for four reasons. First, preliminary regressions rejected a linearity constraint in the relationship between discount rates and most outcomes. A quadratic discount rate term is often significant. Second, specifications with indicator variables for bins of four discount rate ranges capture a greater proportion of the outcome variance than continuous discount rates. Third, we find that the significant results are clustered around respondents we call "extreme" discounters-those who are willing to take R7,500 rand today rather than R15,000 rand in a year. Extreme discounters constitute about $10 \%$ of our sample. Fourth, following Chabris et. al. (2008), who transform their discount rates into percentiles, we do not want the long right-hand tail to bias our regression results by violating Ordinary Least Squares (OLS) assumptions of normality. Kirby, Petry, and Bickel (1999) log discount rates to deal with this problem, while we chose a bin methodology that allows for scrutiny of high and extreme discounter predictions.

[^5]Our regression models for determining the relationship of discount rates to the outcomes are simple. All are Ordinary Least Squares. Because all regressions are at the individual respondent level, we omit subscripts in the models below. We show the first regression to showcase the predictive power of the single dummy Extreme Discounter:
(1) Outcome $Y=\alpha_{0}+\alpha_{1}$ Extreme Discounter $+\varepsilon$
where Outcome $Y$ is the outcome of interest. The second, more controlled, regression is as follows:
(2) Outcome $Y=\alpha_{0}+\alpha_{1}$ Extreme Discounter $+\alpha_{2}$ High Discounter $+\alpha_{3}$ Medium

Discounter $+\alpha_{4}$ Medium-Low Discounter $+X+\varepsilon$
where $X$ is a full set of individual-level demographic controls including sex, age, race, employment status, education, and marital status. ${ }^{10}$ We also run a third model, model (3), including the four attitudinal indices (see Table C in the online appendix). The inclusion of the four indices does not materially affect the coefficient estimates, though several are significant.

## III. Results

Results for each of our 15 outcomes are presented in the following order by group: self-reported health outcomes, Discovery-captured health outcomes, insurance, credit, and driving-device outcomes, and investment outcomes.

Table 3 shows the results of models (1) and (2) for the self-reported variables signaling obesity ( $\mathrm{BMI}>30$ ), smoking, never going to the gym, and saving less that $5 \%$ of income. After controlling for demographic factors, we find that extreme discounters are 6 percentage points more likely to smoke than low discounters, 6 percentage points more likely to never go to the gym, and extreme, high, and medium discounters are much more likely to save less than $5 \%$ of their income than low discounters. On its own, the extreme discounting dummy is highly predictive of smoking, never going to the gym, and saving less than $5 \%$. The discounting variables never significantly predict obesity, failing to replicate the findings of previous studies like Chabris et. al. (2008) and Komles, Smith, and Bogin (2004).
[Table 3 about here.]

[^6]More consistently significant results are found for the objective, Discoverycaptured measures of health, perhaps because respondents with bad habits are reticent about reporting them on the survey (Boniface and Shelton 2013). Table 4 presents regression results for Vitality status-a measure of engagement in a wellness program, 15 with 5 being the most engaged-, whether Vitality members enroll in a health-food program, proportion of healthy food bought at the grocery store, and whether Vitality members sign up for subsidized gym memberships. Extreme and high discounters are less involved in the Vitality program. Extreme discounters are 7 percentage points less likely to enroll in the healthy foods program and, if they are enrolled, buy 5 percentage points less healthy food than low discounters. And high discounters are 5 percentage points less likely to take advantage of the subsidized gym membership. Though the point estimate for extreme discounters is also around -6 percentage points for gym membership, it is not significant. These results are consistent with the theory of Grossman (1972): high discounters are less likely to make investments in their long-term health.
[Table 4 about here]
We observe several significant relationships between discount rates and Discovery-captured outcomes that do not deal directly with day-to-day health choices, but instead with insurance and credit decisions and the unique Driver Score variable. Table 5 presents these results. We find that extreme discounters are both less likely to buy life insurance and less likely to obtain a Discovery credit card. These results are very robust to the addition of demographic controls, and are consistent with the theory that an up-front cost (either the cost of insurance or of applying for a credit card) are being avoided at the cost of a later benefit (a payout if the individual happens to die or the ability to utilize credit to smooth costs over time). ${ }^{11}$

We also find the surprising result that extreme discounters, not controlling for demographic factors, are 2 percentage points more likely to default on their credit card than non-extreme discounters. This is remarkable given that everyone else defaults at only a $1 \%$ rate, suggesting that there is an over $200 \%$ increase in risk for high discounters. However, this coefficient is not significant.

Our regressions concerning Driver Score based on in-car devices find that Drivers with higher discount rates have a lower driver safety score. This result is significant at the $95 \%$ confidence level for high discounters. The coefficient on extreme discounters is never significant, but that may be because only $6.5 \%$, or 52 , of the 805 respondents with telematics devices in their cars are extreme discounters (while $15.0 \%$ are high discounters). However, the regressions provide compelling evidence that discount rates can predict driver quality.
[Table 5 about here.]

[^7]Finally, Table 6 presents evidence that extreme discounters make riskier and perhaps sub-optimal investment decisions. We find that extreme discounters pick funds with higher Relative Return-however, Relative Return is a metric, 1-5, provided by Discovery that correlate best with historical volatility. ${ }^{12}$ Not surprisingly, we also find that extreme discounters choose more volatile funds, choosing funds with 3-year daily return standard deviation .38 higher than low discounters. These are twin results, as Relative Return and Volatility are highly correlated. High discounters also choose funds with $.16 \%$ higher fees than low discounters. These arguably sub-optimal investment decisions are consistent with the theory that present-biased, high discount rate respondents are seeking short-term returns by taking on greater risk.
[Table 6 about here]
Taken together, Tables 3-6 show the pervasive predictive power of discount rates. They also suggest, in aggregate, that the extreme discounter indicator variable does the lion's share of the predictive work. In 9 of the 13 demographic-controlled regressions that have significant discounting coefficients, extreme discounter is a statistically significant predictor. This suggests that there is significant merit to asking a single question to determine if a respondent is an extreme or non-extreme discounter.

Our results also point to this group being important because of how different extreme discounters are than high discounters. In five of the 13 regressions with significant coefficients, the point estimate is more than twice as large as the coefficient on high discounter (Never Go To The Gym, Health Food Program, Health Food Ratio, Credit Card, Relative Return). In another four it is substantially larger (Smoke, Vitality Status, Life Insurance, and Volatility). In just four it is equal to or smaller than high discounter (Save Less Than 5\%, Gym Member, Driver Score, and TER). These results suggest that there is great value to identifying the right-hand tail of the discounting distribution over differences near the median, as that is where the bulk of predictive power lies.

Knowing whether someone is an extreme discounter or not has important economic implications. A researcher or firm can ask a single question to elicit whether someone would rather $\$ X$ today or $2 * \$ X$ in a year. A preference for half as much today indicates that the respondent is an extreme discounter. Identifying the approximately $10 \%$-or perhaps more, given that high discounters may have disproportionately opted out of the survey-of respondents who are extreme discounters might help insurance companies predict the highest-risk applicants for health, driving, credit, and investment benefits. This experimental result suggests that this low-cost proxy for individual discount rates may help researchers and actuaries understand the risk profiles of subjects and clients. This question could be added easily to any demographic questionnaire-or

[^8]might be elicited sending subjects an offer of \$X today or $2 * \$ \mathrm{X}$ in a year-without resort to a MPL or more complicated procedure.

## IV. Conclusion

This paper provides evidence that discount rates - especially extreme discount rates - are a robust predictor of real-world outcomes. We add objectively measured new variables to the literature: an in-vehicle measure of driver quality, the decision to purchase life insurance, and the decision to make risky investments. The predictive power of laboratory-elicited discount rates are concentrated at the right-hand tail of the discounting distribution. In situations in which a full MPL or even more complicated elucidation of discount rates are impossible or infeasible, a single question can determine if someone is an extreme discounter. Alternative, individuals could be offered to choose between being given a certain amount immediately or waiting a year for twice as much. Their choice to receive more tomorrow might qualify them for lower credit-card interest rates or insurance premiums.

Overall, we find that highly present-biased respondents make a less positive set of health, driving, credit, saving, and insurance decisions. We find that, compared to low discounters, high and extreme discounters are more likely to never go to the gym, more likely to save less than $5 \%$ of their income, less involved in the Vitality program, less likely to be involved in the HealthyFoods program, buy less healthy food if they are enrolled, and are less likely to avail themselves of Discovery's subsidized gym program. They are also less likely to have life insurance and a credit card and drive less safely. They are also more likely to make risky investments with higher fund costs. Presentbiased activity in the laboratory is a robust predictor of many negative activities in the field.

## References

Anderson, Steffen, Glenn W Harrison, Morten I Lau, and E Elisabet Rutstrom. 2008. Eliciting Risk and Time Preferences. Econometrica 76(3): 583-618.

Asraf, Nava, Dean Karlan, and Wesley Yin. 2006. Tying Odysseus to the Mast: Evidence From a Commitment Savings Product in the Philippines. Quarterly Journal of Economics 121(2): 635-672.

Becker, Gary and Casey Mulligan. 1997. The Endogenous Determination of Time Preference. The Quarterly Journal of Economics 112(3): 729-758.

Bickel, Warren, Amy Odum, Gregory Madden. 1999. Impulsivity and Cigarette Smoking: delay discounting in current, never, and ex-smokers. Pychopharmacology 146: 447-554.

Boniface, Sadie and Nicola Shelton. 2013. How is alcohol consumption affected if we account for under-reporting? A hypothetical scenario. The European Journal of Public Health 23(6): 1076 - 1081.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde. 2010. "Are Risk Aversion and Impatience Related to Cognitive Ability?" American Economic Review, 100: 1238-1260.

Frederick, Shane, George Loewenstein, and Ted O’Donoghue. 2002. Time Discounting and Time Preference: A Critical Review. Journal of Economic Literature Vol. 40(2): 351-401.

Harrison, G. W., Lau, M. I. and Rutström, E. E. 2010. Individual Discount Rates and Smoking: Evidence From a Field Experiment in Denmark. Journal of Health Economics 29: 708-717.

Grossman, M. 1972. On the concept of health capital and the demand for health. Journal of Political Economy 80(2): 223-255.

Kirby, Kris, Nancy Petry, and Warren Bickle. 1999. Heroin Addicts Have Higher Discount Rates for Delayed Rewards than Non-Drug-Using Controls. Journal of Experimental Psychology 128: 78-87.

Komlos, John, Patricia Smith, and Barry Bogin. 2004. Obesity and the rate of time preference: is there a connection? Journal of Biosocial Science 36: 209-219.

Maier, Stephan, and Charles Sprenger. 2010. Present-Biased Preferences and Credit Card Borrowing. American Economic Journal: Applied Economics 2: 193-210.

Mitchell, Suzanne. 1999. Measures of impulsivity in cigarette smokers and nonsmokers. Psychopharmacology 146:455-464.

Patel, Deepak, Craig Nossel, Eleanore Alexander, Derek Yach. 2013. Promotion in SubSaharan Africa: Progress and Persisting Challenges. Progress in Cardiovascular Diseases 56(3): 356-362.

Petry, N. M. 2001. Delay discounting of money and alcohol in actively using alcoholics, currently abstinent alcoholics, and controls. Psychopharmacology 154: 243-250.

Sturm, Roland, Ruopeng An, Darren Segal, Deepak Patel. 2013. A Cash-Back Rebate Program for Healthy Food Purchases in South Africa. American Journal of Preventative Medicine 44(6): 567-572.

Sutter, Matthias, Martin Kocher, Daniela Rützler and Stefan Trautmann. 2013. Impatience and uncertainty: Experimental decisions predict adolescents' field behavior. American Economic Review 103: 510-531.

Zauberman, Gal, B. Kyu Kim, Selin Malkoc, and James Bettman. 2009. Discounting Time and Time Discounting: Subjective Time Perception and Intertemporal Preferences. Journal of Marketing Research 46(4): 543-556.

Table 1. Multiple Price List

| Payoff <br> Alternative | Option A <br> (Today), rand | Option B (1 <br> Year), rand | Implied <br> Annual <br> Discount <br> Rate |
| :---: | :---: | :---: | :---: |
| 1 | 15,000 | 15,000 | 0.0\% |
| 2 | 14,250 | 15,000 | 5.3\% |
| 3 | 13,500 | 15,000 | 11.1\% |
| 4 | 12,750 | 15,000 | 17.6\% |
| 5 | 12,000 | 15,000 | 25.0\% |
| 6 | 11,250 | 15,000 | 33.3\% |
| 7 | 10,500 | 15,000 | 42.9\% |
| 8 | 9,750 | 15,000 | 53.8\% |
| 9 | 9,000 | 15,000 | 66.7\% |
| 10 | 8,250 | 15,000 | 81.8\% |
| 11 | 7,500 | 15,000 | 100.0\% |
| Note: The exchange rate in 2014 for South African rand:USD is approximately 10:1. |  |  |  |

Table 2. Summary Statistics

| Variable | Obs | Mean |  | Std. Dev. Min |  | Max |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| Discount Rates |  |  |  |  |  |  |
| Discount Rate | 3558 | 0.29 | 0.28 | 0 | 1 |  |
| dis_extreme $(100 \%+$ ) | 3558 | 0.10 | 0.30 | 0 | 1 |  |
| dis_high (40-99\%) | 3558 | 0.15 | 0.36 | 0 | 1 |  |
| dis_med (20-39\%) | 3558 | 0.24 | 0.43 | 0 | 1 |  |
| dis_med_low (10-19\%) | 3558 | 0.28 | 0.45 | 0 | 1 |  |
| dis_low (0-9\%) | 3558 | 0.22 | 0.42 | 0 | 1 |  |
| Self-Reported Outcomes |  |  |  |  |  |  |
| Obese | 3401 | 0.22 | 0.42 | 0 | 1 |  |
| Smoke | 3558 | 0.13 | 0.34 | 0 | 1 |  |
| Never Go To Gym | 3558 | 0.35 | 0.48 | 0 | 1 |  |
| Save Less Than 5\% | 3216 | 0.26 | 0.44 | 0 | 1 |  |
| Discovery Outcomes |  |  |  |  |  |  |
| Vitality Status | 2916 | -- | 1.51 | 1 | 5 |  |
| Health Food Program | 2916 | 0.24 | 0.43 | 0 | 1 |  |
| Health Food Ratio | 571 | 0.26 | 0.13 | 0 | 1 |  |
| Gym Member | 2916 | 0.55 | 0.50 | 0 | 1 |  |
| Life Insurance | 3558 | 0.54 | 0.50 | 0 | 1 |  |
| Credit Card | 3558 | 0.43 | 0.50 | 0 | 1 |  |
| Credit Card Default | 1533 | 0.01 | 0.10 | 0 | 1 |  |
| Driver Score | 804 | -- | 202.39 | 10 | 750 |  |
| Relative Return | 1780 | 2.91 | 0.61 | 0.3827459 | 5 |  |
| Volatilty | 1744 | 7.95 | 1.87 | 1.6 | 16.9 |  |
| TER | 1362 | 1.23 | 0.58 | 0.0 | 2.8 |  |

Note: We have omitted the values marked with a "--" to comply with our Discovery agreement not to share this proprietary information.

Table 3. Self-reported health outcomes

| VARIABLES | (1) <br> Obese | (2) Obese | (3) Smoke | (4) Smoke | (5) <br> Never Go To Gym | (6) <br> Never Go To Gym | (7) <br> Save Less Than 5\% | (8) <br> Save Less Than 5\% |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| dis_extreme | $\begin{gathered} 0.0220 \\ (0.0253) \end{gathered}$ | $\begin{gathered} 0.0130 \\ (0.0283) \end{gathered}$ | $\begin{gathered} 0.0596^{* * *} \\ (0.0215) \end{gathered}$ | $\begin{gathered} 0.0569 * * \\ (0.0238) \end{gathered}$ | $\begin{gathered} 0.0795 * * * \\ (0.0276) \end{gathered}$ | $\begin{aligned} & 0.0568^{*} \\ & (0.0306) \end{aligned}$ | $\begin{gathered} 0.0804 * * * \\ (0.0285) \end{gathered}$ | $\begin{gathered} 0.0796^{* * *} \\ (0.0306) \end{gathered}$ |
| dis_high |  | $\begin{gathered} 0.0151 \\ (0.0234) \end{gathered}$ |  | $\begin{gathered} 0.0293 \\ (0.0188) \end{gathered}$ |  | $\begin{aligned} & 0.00730 \\ & (0.0259) \end{aligned}$ |  | $\begin{gathered} 0.0796^{* * *} \\ (0.0259) \end{gathered}$ |
| dis_med |  | $\begin{gathered} 0.0211 \\ (0.0209) \end{gathered}$ |  | $\begin{aligned} & 0.00418 \\ & (0.0158) \end{aligned}$ |  | $\begin{gathered} 0.000316 \\ (0.0233) \end{gathered}$ |  | $\begin{gathered} 0.0512 * * \\ (0.0218) \end{gathered}$ |
| dis_med_low |  | $\begin{aligned} & -0.00574 \\ & (0.0195) \end{aligned}$ |  | $\begin{gathered} 0.0138 \\ (0.0148) \end{gathered}$ |  | $\begin{aligned} & -0.0123 \\ & (0.0219) \end{aligned}$ |  | $\begin{gathered} 0.0330 \\ (0.0205) \end{gathered}$ |
| Demographic Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Constant | $\begin{aligned} & 0.222 * * * \\ & (0.00748) \end{aligned}$ | $\begin{gathered} 0.232 * * * \\ (0.0306) \end{gathered}$ | $\begin{aligned} & 0.125^{* * *} \\ & (0.00584) \end{aligned}$ | $\begin{gathered} 0.245 * * * \\ (0.0275) \end{gathered}$ | $\begin{aligned} & 0.338^{* * *} \\ & (0.00836) \end{aligned}$ | $\begin{gathered} 0.386^{* * *} \\ (0.0341) \end{gathered}$ | $\begin{aligned} & 0.254 * * * \\ & (0.00806) \end{aligned}$ | $\begin{gathered} 0.315 * * * \\ (0.0338) \end{gathered}$ |
| Observations | 3,401 | 3,401 | 3,558 | 3,558 | 3,558 | 3,558 | 3,216 | 3,216 |
| R -squared | 0.000 | 0.035 | 0.003 | 0.045 | 0.002 | 0.055 | 0.003 | 0.065 |

Note: Robust standard errors in parentheses. Asterisks indicate significance ( $*=.10$, $* *=.05, * * *=.01$ ). Coefficients on demographic controls can be found in Online Appendix Table C.

Table 4. Discovery-captured health outcomes

|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| VARIABLES | Vitality Status | Vitality Status | Health Food Program | Health <br> Food Program | Health <br> Food <br> Ratio | Health <br> Food <br> Ratio | Gym Member | Gym Member |
| dis_extreme | $\begin{gathered} -0.448 * * * \\ (0.0928) \end{gathered}$ | $\begin{gathered} -0.396^{* * *} \\ (0.106) \end{gathered}$ | $\begin{gathered} 0.0655^{*} * \\ (0.0254) \end{gathered}$ | $\begin{gathered} - \\ 0.0673 * * \\ (0.0294) \end{gathered}$ | $\begin{aligned} & -0.0316 \\ & (0.0244) \end{aligned}$ | $\begin{aligned} & -0.0471^{*} \\ & (0.0258) \end{aligned}$ | $\begin{aligned} & -0.0525 \\ & (0.0329) \end{aligned}$ | $\begin{gathered} -0.0568 \\ (0.0371) \end{gathered}$ |
| dis_high |  | $\begin{aligned} & -0.224^{* *} \\ & (0.0914) \end{aligned}$ |  | $\begin{aligned} & -0.0177 \\ & (0.0259) \end{aligned}$ |  | $\begin{aligned} & -0.00230 \\ & (0.0188) \end{aligned}$ |  | $\begin{gathered} -0.0520^{*} \\ (0.0303) \end{gathered}$ |
| dis_med |  | $\begin{aligned} & -0.0734 \\ & (0.0800) \end{aligned}$ |  | $\begin{aligned} & -0.00556 \\ & (0.0234) \end{aligned}$ |  | $\begin{aligned} & -0.0182 \\ & (0.0147) \end{aligned}$ |  | $\begin{gathered} -0.0261 \\ (0.0268) \end{gathered}$ |
| dis_med_low |  | $\begin{gathered} 0.0403 \\ (0.0776) \end{gathered}$ |  | $\begin{aligned} & -0.0119 \\ & (0.0225) \end{aligned}$ |  | $\begin{gathered} -0.0254^{*} \\ (0.0144) \end{gathered}$ |  | $\begin{gathered} -0.0238 \\ (0.0256) \end{gathered}$ |
| Demographic Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Constant | $\begin{gathered} --* * * \\ (0.0293) \end{gathered}$ | $\begin{gathered} --* * * \\ (0.115) \end{gathered}$ | $\begin{aligned} & 0.243 * * * \\ & (0.00832) \end{aligned}$ | $\begin{gathered} 0.0168 \\ (0.0289) \end{gathered}$ | $\begin{aligned} & 0.262 * * * \\ & (0.00541) \end{aligned}$ | $\begin{gathered} 0.304 * * * \\ (0.0577) \end{gathered}$ | $\begin{aligned} & 0.558 * * * \\ & (0.00963) \end{aligned}$ | $\begin{gathered} 0.567 * * * \\ (0.0405) \end{gathered}$ |
| Observations | 2,916 | 2,916 | 2,916 | 2,916 | 571 | 571 | 2,916 | 2,916 |
| R-squared | 0.007 | 0.064 | 0.002 | 0.039 | 0.004 | 0.092 | 0.001 | 0.033 |

Note: Robust standard errors in parentheses. Asterisks indicate significance ( $*=.10$, $* *=.05, * * *=.01$ ). Coefficients on demographic controls can be found in Online Appendix Table C. We have omitted the constants marked with a "--" to comply with our Discovery agreement not to share this proprietary information.

Table 5. Insurance, credit, and driving-device outcomes

|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| VARIABLES | Life Insurance | Life Insurance | Credit <br> Card | Credit Card | Credit Card Default | Credit Card <br> Default | Driver Score | Driver Score |
| dis_extreme | $\begin{gathered} - \\ 0.122^{* * *} \\ (0.0278) \end{gathered}$ | $\begin{gathered} -0.112^{* * *} \\ (0.0305) \end{gathered}$ | $\begin{gathered} - \\ 0.165 * * * \\ (0.0256) \end{gathered}$ | $\begin{gathered} -0.121^{* * *} \\ (0.0293) \end{gathered}$ | $\begin{gathered} 0.0205 \\ (0.0174) \end{gathered}$ | $\begin{gathered} 0.0122 \\ (0.0194) \end{gathered}$ | $\begin{gathered} 1.419 \\ (30.03) \end{gathered}$ | $\begin{aligned} & -34.85 \\ & (31.94) \end{aligned}$ |
| dis_high |  | $\begin{gathered} 0.0768^{-} * * \\ (0.0265) \end{gathered}$ |  | $\begin{aligned} & -0.0119 \\ & (0.0265) \end{aligned}$ |  | $\begin{aligned} & -0.00670 \\ & (0.00914) \end{aligned}$ |  | $\begin{gathered} -45.84^{* *} \\ (22.93) \end{gathered}$ |
| dis_med |  | $\begin{gathered} 0.0192 \\ (0.0232) \end{gathered}$ |  | $\begin{gathered} 0.0171 \\ (0.0237) \end{gathered}$ |  | $\begin{aligned} & -0.00650 \\ & (0.00835) \end{aligned}$ |  | $\begin{gathered} -44.02 * * \\ (21.03) \end{gathered}$ |
| dis_med_low |  | $\begin{aligned} & -0.00844 \\ & (0.0224) \end{aligned}$ |  | $\begin{gathered} 0.0127 \\ (0.0228) \end{gathered}$ |  | $\begin{aligned} & -0.00921 \\ & (0.00752) \end{aligned}$ |  | $\begin{gathered} -31.84^{*} \\ (19.27) \end{gathered}$ |
| Demographic Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Constant | $\begin{aligned} & 0.554 * * * \\ & (0.00878) \end{aligned}$ | $\begin{gathered} 0.358 * * * \\ (0.0346) \end{gathered}$ | $\begin{aligned} & 0.449 * * * \\ & (0.00879) \end{aligned}$ | $\begin{gathered} 0.275 * * * \\ (0.0329) \end{gathered}$ | $\begin{gathered} 0.00976 * * * \\ (0.00260) \end{gathered}$ | $\begin{aligned} & 0.0326^{*} \\ & (0.0167) \end{aligned}$ | (7.366) | (34.22) |
| Observations | 3,558 | 3,558 | 3,558 | 3,558 | 1,533 | 1,533 | 804 | 804 |
| R-squared | 0.005 | 0.109 | 0.010 | 0.089 | 0.002 | 0.022 | 0.000 | 0.111 |

Note: Robust standard errors in parentheses. Asterisks indicate significance (*=.10, $* *=.05, * * *=.01$ ). Coefficients on demographic controls can be found in Online Appendix Table C. We have omitted the constants marked with a "--" to comply with our Discovery agreement not to share this proprietary information.

Table 6. Investment outcomes

| VARIABLES | (1) <br> Relative Return | (2) <br> Relative Return | (5) <br> Volatilty | (6) <br> Volatilty | (7) <br> TER | (8) <br> TER |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| dis_extreme | $\begin{aligned} & 0.122^{* *} \\ & (0.0490) \end{aligned}$ | $\begin{aligned} & 0.124^{* *} \\ & (0.0556) \end{aligned}$ | $\begin{gathered} 0.303 * * \\ (0.150) \end{gathered}$ | $\begin{aligned} & 0.309^{*} \\ & (0.168) \end{aligned}$ | $\begin{aligned} & -0.0261 \\ & (0.0539) \end{aligned}$ | $\begin{aligned} & -0.00425 \\ & (0.0623) \end{aligned}$ |
| dis_high |  | $\begin{gathered} 0.0158 \\ (0.0464) \end{gathered}$ |  | $\begin{gathered} 0.164 \\ (0.151) \end{gathered}$ |  | $\begin{gathered} 0.161^{* * *} \\ (0.0522) \end{gathered}$ |
| dis_med |  | $\begin{gathered} -0.0354 \\ (0.0420) \end{gathered}$ |  | $\begin{gathered} -0.223^{*} \\ (0.125) \end{gathered}$ |  | $\begin{gathered} 0.0631 \\ (0.0454) \end{gathered}$ |
| dis_med_low |  | $\begin{aligned} & -0.0210 \\ & (0.0415) \end{aligned}$ |  | $\begin{aligned} & -0.0702 \\ & (0.128) \end{aligned}$ |  | $\begin{gathered} 0.0149 \\ (0.0452) \end{gathered}$ |
| Demographic Controls | No | Yes | No | Yes | No | Yes |
| Constant | $\begin{gathered} 2.898 * * * \\ (0.0152) \end{gathered}$ | $\begin{gathered} 2.800 * * * \\ (0.0674) \end{gathered}$ | $\begin{gathered} 7.923 * * * \\ (0.0471) \end{gathered}$ | $\begin{gathered} 7.596^{* * *} \\ (0.207) \end{gathered}$ | $\begin{gathered} 1.236 * * * \\ (0.0164) \end{gathered}$ | $\begin{gathered} 1.203 * * * \\ (0.0716) \end{gathered}$ |
| Observations | 1,780 | 1,780 | 1,744 | 1,744 | 1,362 | 1,362 |
| R-squared | 0.003 | 0.014 | 0.002 | 0.026 | 0.000 | 0.034 |

Note: Robust standard errors in parentheses. Asterisks indicate significance ( $*=.10$, $* *=.05, * * *=.01$ ). Coefficients on demographic controls can be found in Online Appendix Table C. When choosing investment options, Discovery customers can see all of the fund-level variables above at https://www.discovery.co.za/portal/individual/invest-fund-compare. TER is a Total Expense Ratio.


[^0]:    *William K. Townsend Professor at Yale Law School. **Research Specialist at Yale Law School. ${ }^{* * *}$ Research and Development Actuary at Discovery Limited. For helpful comments, thanks to Emile Stipp and Derek Yach. For help with the survey, thanks to Eitan Dubb and Lauren Hinckley. For project funding and data access, thanks to Discovery. Comments and questions can be directed to darrense@discovery.co.za.

[^1]:    ${ }^{1}$ For example, the median income for people over 18 in our sample with credit cards is 360,000 rand, or about $\$ 36,000(\mathrm{~N}=1,024)$. The median income for men and women over 15 in the US in 2012 was $\$ 28,000$ (number from Census Table P-8AR https://www.census.gov/hhes/www/income/data/historical/people/).

[^2]:    ${ }^{2}$ Ashraf, Karlan, and Yin (2006) use a binary methodology for measures of impatience in their regression analysis for savings account take-up in the Philippines. We go further to define the optimal cutoffs for the binary more precisely.

[^3]:    ${ }^{3}$ Dohmen et. al. (2010) link these two variables to measures of intelligence and in the process establish this correlation. Anderson et. al. (2008) find the same relationship.
    ${ }^{4}$ The survey introduced a randomly assigned prompt before the MPL. The goal of this prompt was to see if asking participants to estimate the time it takes to complete everyday tasks lowers their discount rates, following Zauberman et. al. (2008). However, because attrition rates were much higher for our control (estimation of the number of calories in common food) than the treatment, we have elected to omit the results of this randomized experiment.

[^4]:    ${ }^{5}$ We group the statements by theme and constructed indices where a higher number suggests more positive activities (for example, responses for both questions above were reversed by preforming the simple calculation: $x=11$ - response) by averaging the responses to the questions in that theme. See Table A of the online appendix for summary statistics on demographics and attitudinal indices. For a full list of survey questions and a more complete explanation of how the indices are constructed, please see Table B of the online appendix.
    ${ }^{6}$ For an overview of the program, its history, and its goals, see Patel et. al. (2013). This number 1-5 is based off of holistic involvement in the program. This video shows how a Discovery client can improve their Vitality Status: https://www.youtube.com/watch?v=93ebyDe0vIU\#t=74.

[^5]:    ${ }^{7}$ Because our sample only consists of Discovery customers, this variable compares Discovery customers who have life insurance to Discovery customers who do not. This does not mean that those customers who have not purchased life insurance from Discovery have not purchased it from another provider, but it does provide an excellent proxy for the decision to purchase life insurance. However, Discovery's life insurance product is high quality, and offered at a premium to other insurer's products. This may deflect high discounter's business, confounding our results.
    ${ }^{8}$ The telematics devices have a GPS and an accelerometer that monitor various parameters including time of driving, acceleration, cornering, breaking and speeding. Frequency of harsh events determines the score. Higher scores indicate safer driving.
    ${ }^{9}$ The funds are sold through brokers, which might confound causal results if participant selection of brokers is systematically related to their personal characteristics. However, because this is a predictive project we need not worry about this effect.

[^6]:    ${ }^{10}$ Because we use the same regression specification for several dependent variables, we run a robustness test using Seemingly Unrelated Regression (SUR) specifications to account for correlated residuals in the models for those variables for which overlapping samples make it feasible ( 7 of 15 ). These regressions produce nearly identical results as the individual model results presented below. See Table D in the online appendix for results for regressions involving Smoke, Never Go To The Gym, Credit Card, and Life Insurance $(\mathrm{N}=3558)$ and Vitality Status, Health Food Program, and Gym Member ( $\mathrm{N}=2917$ ).

[^7]:    ${ }^{11}$ This result remains significant when we change the right-hand-side variable from having a credit card to having applied for a credit card, suggesting that denial of credit by Discovery does not bias our results.

[^8]:    ${ }^{12}$ The correlation coefficient between Discovery's "Relative Return" variable and 3-year historical return is -0.0049 . The correlation between "Relative Return" and volatility-the standard deviation daily returns over three years-however is 0.6734 .

