Spending profiles predict savings*

Fabian Gunzinger Neil Stewart

October 31, 2022

Abstract

We show that spending patterns are predictive of the frequency with which individual's make transfers into their "emergency savings". We capture the variation in spending profiles using entropy, and define emergency savings as transfers into savings accounts. The effect of spending entropy on savings is statistically significant and economically large, and robust across a number of different specifications and entropy measures based on different product and merchant categories. The direction of the effect depends, however, on the handling of categories in which individuals make no transactions in a given month. While our results give some indication as to how we can explain this result, further research is needed for a more complete understanding.

1 Introduction

This paper documents variation in spending profiles – the way spending transactions are allocated across product of merchant categories – for a large set of users of a financial management app and shows that spending profiles predict "emergency savings" – short-term savings intended as a buffer against unexpected financial shocks.

Emergency savings are distinct from long-term savings aimed to build up funds for retirement, either through individually owned pension and investment vehicles or employer-linked pension schemes. Both kinds of saving are important for financial wellbeing (MPS 2018, CFPB 2017), yet while there is a large literature on pension savings – particularly on the effect on auto-enrollment on savings plan

*We are grateful to Redzo Mujcic and and Zvi Safra for helpful comments. The research was supported by Economic and Social Research Council grant number ES/V004867/1. WBS ethics code: E-414-01-20. Gunzinger: Warwick Business School, fabian.gunzinger@warwick.ac.uk; Stewart: Warwick Business School, neil.stewart@wbs.ac.uk.

participation – little is known about how individuals save for the short-term.[1] Apart from Beshears, Choi, Iwry, et al. (2020), who study design options for automated contributions to emergency saving pots, and Phillips et al. (2021), who test such a design in the UK, there is no research on how we can help individuals build such emergency savings. Furthermore, around a quarter of adults in the UK and the US are unable to cover irregular expenses like car and medical bills: in the UK, 25 percent of adults would be unable to cover an unexpected bill of £300 (Phillips et al. 2021), while in the US, about 30 percent would be unable to cover a $400 bill (Governors of the Federal Reserve System 2022). Research from the UK suggests, however, that even having £1000 in savings reduces by more than half a household's chances of falling into debt that leads to financial problems (Phillips et al. 2021).

There is a large literature on the determinants of financial behaviour, and some of the main factors that have been identified to influence have individuals and households make financial decisions are: cognitive limitations and financial literacy (Agarwal, Driscoll, et al. 2009, Agarwal and Mazumder 2013, Korniotis and Kumar 2011, Agarwal, Amromin, et al. 2010, Fernandes et al. 2014, Jørring 2020); time-preferences and self-control (Frederick et al. 2002, Read et al. 2018, Ericson and Laibson 2019, Cohen et al. 2020); attitude towards money and spending (Rick, Cryder, et al. 2008, Rick, Small, et al. 2011); ones perceived locus of control (Perry and Morris 2005), degree of optimism (Puri and Robinson 2007), ability to frame decisions broadly rather than narrowly (Kumar and Lim 2008), and propensity to gamble (Kumar 2009); ones social network (Bailey et al. 2018, Kuchler and Stroebel 2021); the degree of ones financial planning (Ameriks et al. 2003); and habits (Blumenstock et al. 2018, Schaner 2018, De Mel et al. 2013).[2]

In this paper, we use simple summary statistics of individuals' spending profiles and test whether differences in spending profiles are predictive of the frequency with which individuals make transactions into their savings accounts. To summarise spending profiles, we calculate the entropy of an individual's spending profile in a given period, which captures the predictability of spending transactions in said period. We focus on spending profiles for three reasons: first, our understanding of how individuals spend their money is largely based on survey data from a relatively small number of individuals. Large-scale transaction data of the type used in this paper has become available to researchers only in recent years and has not, to the best of our knowledge, been used to explore patterns in

[1]For main contributions to the to literature on auto-enrollment into savings plans, see Madrian and Shea (2001), Choi et al. (2002), Choi et al. (2004), and Beshears, Choi, Laibson, et al. (2009).

[2]For two thorough reviews, see Agarwal, Chomsisengphet, et al. (2017) and Greenberg and Hershfield (2019).

consumer spending behaviour.

Second, research suggests that spending profiles might reflect circumstances in an individual's life that are also related to their saving behaviour. For instance, Muggleton et al. (2020) find that higher spending entropy predicts financial distress, and hypothesise that this is because disorder in ones personal life might simultaneously be reflected in ones spending behaviour and make it harder to deal with financial obligations, based on a literature in psychology that suggests that disorder is associated with a range of negative outcomes such as impaired executive function (Vernon-Feagans et al. 2016), lower cognitive inhibition (Mittal et al. 2015), and activation of anxiety-related neural circuits (Hirsh et al. 2012). In addition to financial outcomes, high entropy behaviour has also been found to be associated with higher calorie intake (Skatova et al. 2019) and with a higher number of visits to and higher spend in supermarkets (Guidotti et al. 2015).

Finally, we think of our focus on spending profile as a proof of concept for the broader agenda of testing whether we can extract simple summary statistics from large-scale transaction data that capture relevant information about individuals' life circumstances. If successful, such information could be used in the design of financial management apps to provide timely and customised assistance to users when they need it most.

Our results show that spending entropy is predictive of the frequency of savings transactions, with an effect size similar to that of an increase in monthly income income of between £1,000 and £2,000. This holds true if we control for simple component parts of entropy, suggesting that the non-linear way in which entropy combines these components captures something about spending profiles that is of relevance. The results are also largely consistent across entropy measures calculated based on different product and merchant categories. However, they direction of the effect changes for different types of entropy that differ in the way in which we treat product or merchant categories in which an individual makes no transactions in a given period. Exploratory research suggests that the number of such zero count categories, together with the variation of counts for categories with a positive number of transactions, goes some way in explaining the outcome. But more work is necessary to more fully understand this outcome, which we leave for future research.

Apart from the small literature on emergency savings mentioned above, our work contributes to a broader literature that uses financial-transaction data from banks or financial aggregator apps to understand consumer financial behaviour. For instance, Kuchler and Pagel (2020) use data from a financial aggregator app to estimate time preferences. Similar data has been used to show that consumer spending varies across the pay cycle (Gelman et al. 2014, Olafsson and

Pagel 2018), to test the consumer spending response to exogenous shocks (Baker 2018, Baugh et al. 2014), and to better understand the generational differences in financial platform usage patterns (Carlin et al. 2019). Some researchers use transaction-data directly provided by banks. Ganong and Noel (2019) show that consumer spending drops sharply after the predictable income drop from exhausting unemployment insurance benefits, Meyer and Pagel (2018) analyse how individuals reinvest realised capital gains and losses, and Muggleton et al. (2020) show that chaotic spending behaviour is a harbinger of financial distress.[3]

The remainder of this paper is organised as follows: Section 2 discusses data pre-processing and our methodological approach, Section 3 presents the main results and further exploratory analysis, and Section 4 concludes.

2 Methods

2.1 Dataset description

We use data from Money Dashboard (MDB), a financial management app that allows its users to link accounts from different banks to obtain an integrated view of their finances.[4] The full dataset contains more than 500 million transactions made between 2012 and June 2020 by about 270,000 users, and provides information such as date, amount, and description about the transaction as well as account and user-level information.

The main advantages of the data for the study of consumer financial behaviour are its high frequency, that it can be cheaply collected for a very large number of users, that collection is automatic and thus less prone to errors and unaffected by biases that bedevil survey measures, and that it offers a view of consumers' entire financial life across all their accounts, rather than just a view of their accounts held at a single bank, provided they added all their accounts to MDB. The main limitation is the non-representativeness of the sample relative to the population as a whole. Financial management apps are known to be used disproportionally by men, younger people, and people of higher socioeconomic status (Carlin et al. 2019). Also, as pointed out in Gelman et al. (2014), a willingness to share financial information with a third party might not only select on demographic characteristics, but also for an increased need for financial management or a higher degree of financial sophistication. Because our analysis does not rely on representativeness, we do not address this.[5] Another major limitation is imperfect

[3]For a comprehensive review of the literature using financial transaction data, see Baker and Kueng (2022).

[4]https://www.moneydashboard.com.

[5]For an example of how re-weighing can be used to mitigate the non-representative issue,

transaction labelling. Money Dasbhoard automatically classifies transaction into different categories based on transaction descriptions and merchant information. While this classification is quite mature and reliable for transactions with major retailers, it is imprecise and at times incorrect for less common transactions. We discuss this further in the next section.

2.2 Preprocessing and sample selection

We use the dataset described above for a number of projects, and perform a number of steps to create a minimally cleaned version of the dataset that is the basis for all such projects. These steps are performed in a dedicated data repository and not run as part of this project.[6] Here, we briefly describe the main cleaning steps and their rationale. We drop all transactions with a missing description string because these cannot be categorised, and all transactions that are not automatically categorised by the app. Dropping these transactions makes it likely that we will underestimate amounts spent and saved, but minimises the risk of incorrectly classified transactions. We also group transactions into transfer, spend, and income subgroups, following Muggleton et al. (2020) to define spend subgroups and Hacıoğlu-Hoke et al. (2021) to define income subgroups.[7] Finally, we classify as duplicates and drop transactions with identical user ID, account ID, date, amount, and transaction description. This will drop some genuine transactions, such as when a user buys two identical cups of coffees at the same coffee shop on the same day. However, data inspection suggests that in most cases, we remove genuine duplicates.[8]

To minimise the influence of outliers, we winsorise all variables at the 1 percent level or – if we winsorise on both ends of the distribution – at the 0.5 percent level.[9] We rely on winsorisation (replacing top values with percentile values) instead of trimming (replacing top values with missing values) because data inspection suggests that, in most cases, very large (absolute) values are not the result of data errors, which would call for trimming, but reflect genuine outcomes, which makes winsorising appropriate because it leaves these observations in the data while lowering their leverage to influence results.

The overall aim of the sample selection is to restrict our sample to users for whom we can observe a regular income, can be reasonably sure that they

see Bourquin et al. (2020).

[6]The module with all cleaning functions is available on Github.

[7]The precise list used to classify transactions is available on Github.

[8]While standing order transactions into savings accounts are unlikely to be related to entropy in the short-run, we do not exclude such transactions since, best we can tell, they only account for a small fraction of total transactions.

[9]The code that performs the winsorisation are available on Github.

have added all their bank accounts to MDB, and for whom we observe at least six months of data. Table 1 summarises the sample selection steps we applied, associated data losses, and the size of our final sample.

Table 1: Sample selection

	Users	User-months	Txns	Txns (m£)
Raw sample	271,856	7,948,520	662,112,975	124,573
Drop first and last month	265,760	7,406,482	643,851,490	121,098
At least 6 months of data	231,888	7,318,202	638,056,402	120,155
At least one savings account	144,309	4,788,799	445,037,084	90,541
At least one current account	141,514	4,723,402	439,698,170	89,629
At least £5,000 of annual income	54,248	1,636,887	172,146,281	33,921
At least 10 spend txns each month	40,596	1,186,503	133,867,104	24,998
At least 4 grocery txns each month	18,468	472,656	60,646,753	9,472
At least £200 of monthly spend	18,254	467,639	60,189,168	9,435
Complete demographic information	14,704	394,377	50,895,612	7,923
Working age	14,539	389,790	50,410,659	7,774
Final sample	14,539	389,790	50,410,659	7,774

Notes: Number of users, user-months, transactions, and transaction volume in millions of British Pounds left in our sample after each sample selection step.

We start by dropping the first and last month of data for each user because we are unlikely to observe users' complete data in these months, which might affect selection for the criteria below. Next, we keep only users whom we observe for at least 6 months to ensure that we have a suitable amount of data for each user. We also require users to have at least one current and one savings accounts to be able to observe their spending and savings behaviour. The next few steps are all aimed at selecting users about whom we can be reasonably sure that they have added all their financial accounts. We thus select on criteria we would expect to see for all such users: an annual income of at least £5,000, at least 10 monthly spending transaction, at least 4 monthly grocery transactions, and a total monthly spend of at least £200. Finally, we keep only users for whom we observe all relevant demographic characteristics, and who are between 18 and 65 years old. The reason for the last step is that younger users and retires will plausibly have different financial objectives than people in their working age, which is the population we are interested in.[10]

2.3 Dependent variables

Our outcome variable is a binary indicator for whether of not a user made a payment into any of their savings account in a given month. We classify as payments into savings accounts all savings account credits of £5 or more.

[10]The code that implements the selection criteria is available on Github.

Such transactions will include those that are intended for purposes other than building to deal with unexpected financial shocks. For instance, users might transfer money to their savings accounts to save for a particular goal, such as new car, a holiday, or a wedding. While Money Dashboard allows users to create custom tags for transactions that might, in principle, be helpful to distinguish between these different types of savings transactions, only a very small number of transactions actually have such tags, so that we cannot use them to identify emergency savings more precisely. As a result, we refer to all observed transactions into savings accounts as emergency savings.

The reason we focus on a binary indicator is twofold: first, because we hypothesis that chaotic or difficult life circumstances that are also reflected in spending entropy might make it harder to remember to save, for which the accurate test is to see whether spending entropy is related to the likelihood of making a savings transaction. Second, research by the UK's Money and Pension Service suggests that to build up sufficient emergency funds over time forming a habit of saving regularly is more important than the specific amounts saved month to month (MPS 2018).

2.4 Spending profiles

We define a user's spending profile as the distribution of the number of spending transactions across different spend categories. To summarise these distributions, we calculate spending entropy, based on the formula proposed by Shannon (1948), who defines entropy as $H = -\sum p_i log(p_i)$, which sums, for all possible events, the product of the probability of an event i occurring with the logarithm of that probability.[11] The base of the logarithm is often chosen to be 2, though other choices are possible. Entropy is a cornerstone of information theory, where it measures the amount of information contained in an event. In the behavioural sciences, behavioural entropy has recently been shown to predict the frequency of grocery visits and the per-capita spend per visit (Guidotti et al. 2015), the amount of calories consumed (Skatova et al. 2019), and the propensity for financial distress (Muggleton et al. 2020). In our context, we define the entropy of a user's spending profile in a particular period as (we omit individual and time subscripts to keep the notation simpler):

$$H = -\sum_{c \in \mathcal{C}} p_c log(p_c), \tag{1}$$

where $\mathcal{C}$ is the set of all spending categories, p_c the probability that an individual

[11] Shannon entropy is customarily denoted as H following Shannon's own naming after Ludwig Boltzman's 1872 H-theorem in statistical mechanics, to which it is analogous.

makes a purchase in spending category c, and *log* the base 2 logarithm, which expresses entropy in bits.

Higher entropy means that transactions are more equal across different spending categories, which makes it hard to predict the next transaction, whereas low entropy profiles have the bulk of transactions in a few dominant categories (such as groceries and transportation) and have relatively few transactions in other categories.[12] For simpler interpretation of our regression coefficients below, we standardise entropy scores to have a mean of 0 and a standard deviation of 1.

One limitation introduced by the imperfect transaction labelling in the MDB data is that entropy scores for high-entropy individuals will be biased downwards. This happens because unlabelled transactions tend to be transactions that are rare (i.e. not grocery or Amazon purchases), and it is high-entropy individuals that are more likely to engage in rare transactions. Because our analysis mainly relies on relative entropy levels, this is not of major consequence and we do not pursue this further.

We calculate entropy based on three sets of spend categories. The first measure is based on 9 spending categories used by Muggleton et al. (2020). The second measure is based on our own, more fine-grained, categorisation into 48 different categories.[13] The third measure is based on merchant names, as labelled by Money Dashboard.

We also calculate spending category probabilities in two different ways. To calculate what we call "unsmoothed" entropy scores, we calculate the p_cs in Equation 1 as simple frequentist probabilities

$$p_c = \frac{f_c}{F}, \tag{2}$$

where f_c is the number of transactions in spend category c (the frequency with which c occurs) and $F = \sum_{c \in \mathcal{C}} f_c$ the total number of spending transactions. To avoid taking the log of zero for categories with zero transactions, the sum in Equation 1 is taken over categories with positive transaction counts only.[14] To calculate "smoothed" entropy scores, we apply additive smoothing to calculate probabilities as

$$p_c^s = \frac{f_c + 1}{F + |\mathcal{C}|}, \tag{3}$$

[12]For further discussion on how to interpret Equation 1, see Appendix A.

[13]The precise mapping from MDB transaction tags into 9 and 48 categories is available on Github here and here, respectively. While we group all transaction tags available in the Money Dashboard dataset into 48 categories, there are no "mortgage release" transactions in our final analysis sample, so that, effectively, all calculations are based on 47 categories.

[14]This is automatically handled by the entropy implementation of Python's SciPy package, which is what we use to calculate entropy scores.

where the size of set $\mathcal{C}$, $|\mathcal{C}|$, is the number of unique spending categories. Hence, we simply increment all frequency counts by one and then calculate probabilities as in Equation 2. Because categories with a zero transaction count will have a numerator of 1, the sum in Equation 1 will be taken over all categories.

2.5 Summary statistics

Table 2 provides summary statistics and Figure 1 distributions for the key variables used in our analysis as well as for some sample demographics. Income and spending distributions are being compared to those of the Living Cost and Food Survey, which is the ONS's main survey on household spending. We can see that the sample distributions broadly match the UK-wide ones but that we tend to underestimate incomes and overestimate spend. As discussed above, our sample is biased towards high-earning individuals, so that we would expect both income and spend to be higher than across the whole population. The main reason this is not the case is likely due to the imperfect identification of income transactions in the Money Dasbhoard data.

Table 2: Summary statistics

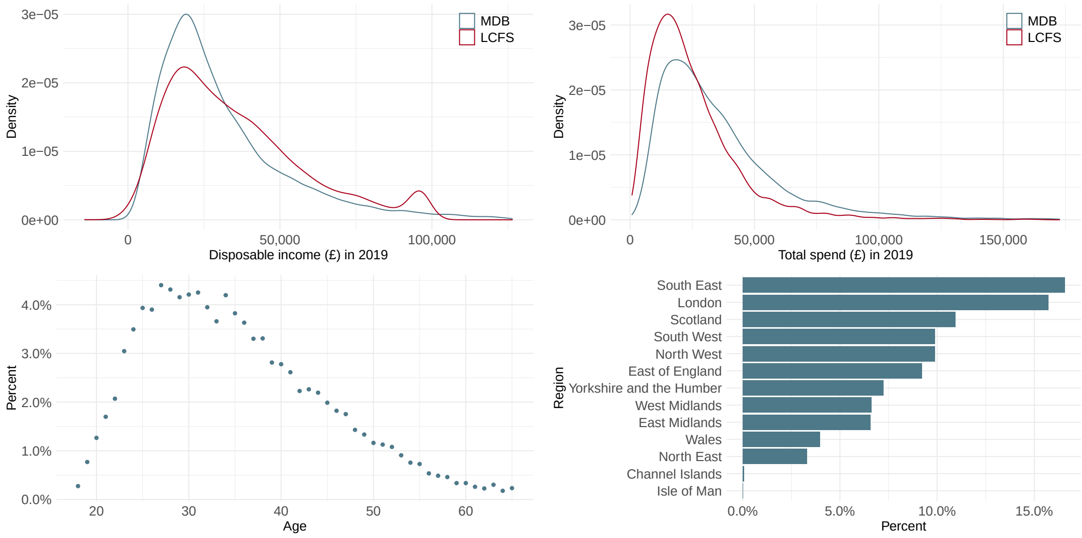
Statistic	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Month income (£'000)	2.77	2.23	0.00	1.45	2.18	3.43	13.69
Has income in month	0.98	0.13	0	1	1	1	1
Makes savings txns.	0.50	0.50	0	0	1	1	1
Month spend (£'000)	2.90	2.50	0.20	1.37	2.20	3.49	16.05
Age (years)	35.72	9.74	18	28	34	42	65
Female	0.43	0.49	0	0	0	1	1
Urban	0.85	0.36	0	1	1	1	1
Unique categories (9)	7.84	1.05	1	7	8	9	9
Unique categories (48)	16.54	4.13	1	14	16	19	35
Unique categories (Merch.)	26.78	9.35	2	20	26	33	85

Notes: Summary statistics for main variables used in analysis and user characteristics. Unique categories refer to the number of distinct spending categories that individuals made purchases in within a given month for spending transaction categorisation based on 9 categories, 48 categories, and merchant names.

Individuals in our data make at least one savings transaction in half of all observed periods, and have at least some income in 98 percent of those periods. As discussed above, our sample is skewed towards men in their 20s, 30s, and 40s, who live in urban areas, mainly the South East of England. Mean monthly spend is £2,000 and the median is £2,200, and the average individual spends money across 27 different merchants in 17 different categories (of the 48 category classification).

Figure 2 shows the distributions of the unsmoothed and smoothed version of entropy based on 9 and 48 spending categories, in the top row, the associated average spending profiles for individuals in the first, third and fifth quintiles in

Figure 1: Demographic characteristics of Money Dashboard users



Notes: The top left and top right panels show the distribution of disposable income and total spending in 2019, respectively, benchmarked against the 2018/19 wave of the ONS Living Cost and Food Survey (LCFS). The bottom left panel shows the distribution of age, the bottom right panel that of the regions.

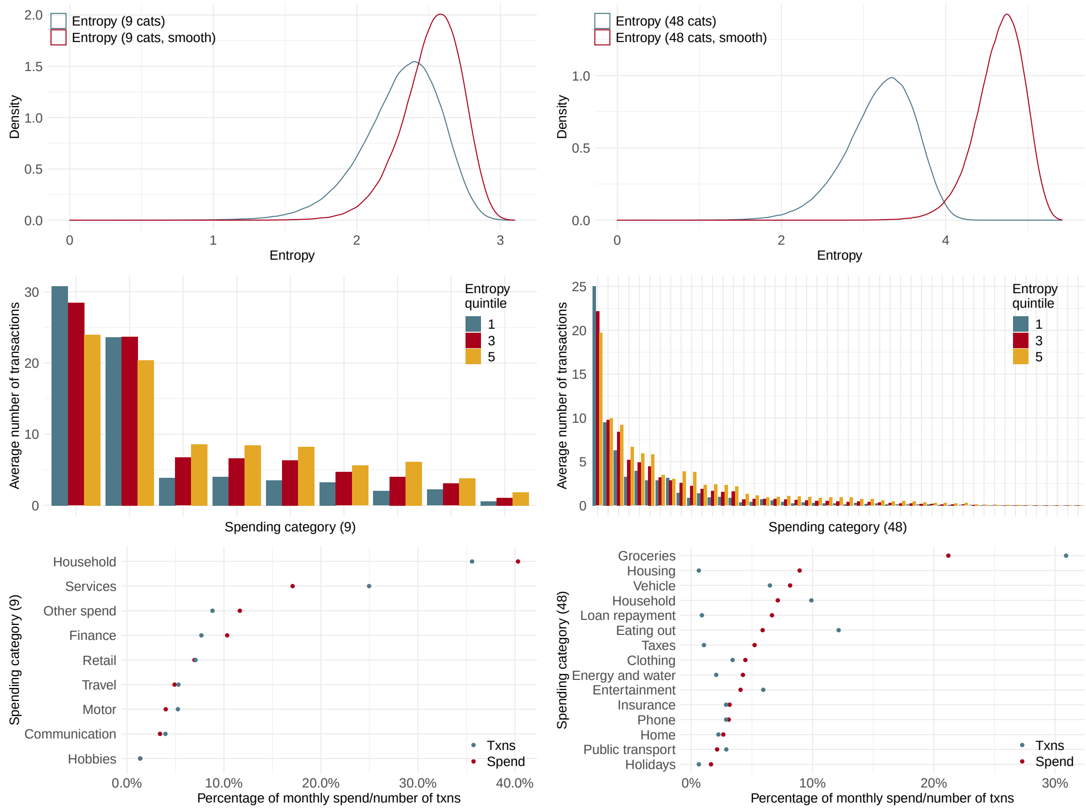
the middle row, and an overview of the main categories individuals spend on in the bottom row. The entropy distributions are bounded on the right by the maximum entropy value, given by the value for a discrete uniform distribution with 9 or 48 distinct values,[15] and on the left by zero, the value for cases where an individual makes all spending transactions in a single category. The different effect of smoothing on the two variables is a result of the different number of categories with counts of zero.

The spending profiles in the middle row show that higher entropy individuals make both fewer transactions in high-count categories and more transactions in low-count categories. The spend profile based on the 48 category classification also shows that spend counts follow approximately a power law distributions, with very few of the 48 categories accounting for the large majority of transactions.

The relative ordering of the top spending categories based on the 48 category classification is about as expected, as are the relationship between the number of transactions and spend: there are, for instance, a large number of relatively low value transactions and meals eaten out (which includes snacks and coffee) and a very small number of high value transactions for housing. The proportion of spend that goes towards housing, less than 10%, is clearly an underestimate, however, and reflects the challenge and imperfect nature of automatic transaction labelling in our data that we discussed above.

[15]For instance, the upper bound for the 9 category variable is $9 \times \left(\frac{1}{9}\right) log_2 \left(\frac{1}{9}\right) \approx 3.17$.

Figure 2: Entropy distributions



Notes: The top row shows distributions of unsmoothed and smoothed versions of entropy. The middle row shows associated spending profiles for individuals in the first, third, and fifth quintile of the entropy distribution. Spending category labels are deliberately omitted to not overcrowd the plots and to focus on the distribution of transaction counts. The bottom row shows the distributions of spend and number of transactions that goes to the top categories. The plot on the left shows all 9 categories, that on the right the top 15 out of 48 categories. The left column shows data for the 9 category classification, the right column data for the 48 categories classification.

2.6 Estimation

To estimate the relationship between spending entropy and savings behaviour we estimate fixed-effect models of the form:

$$y_{i,t} = \alpha_i + \lambda_t + \beta H_{i,t} + x_{i,t}'\delta + \epsilon_{i,t}, \tag{4}$$

where $y_{i,t}$ is an indicator variable equal to one if individual i made at least one transfers to any of their savings account in year-month period t and zero otherwise, H_{it} is i's spending entropy in year-month period t, $x_{i,t}$ a vector of control variables, α_i an individual fixed effect, λ_t a year-month fixed effect, and $\epsilon_{i,t}$ the error term.

The vector of controls includes month spend, month income, an indicator for whether a user had positive income in a given month, and income variability, calculated as the standard deviation of month income over the previous 12 months.

Note that while we might in principle be worried about reverse causality, since making payments into savings accounts might lead to a non-zero count in an additional spend category and thus change entropy, this is not a concern here. As discussed in Section 2.3 and Section 2.4, we define savings as inflows into savings accounts and define entropy based on the classification of spend transactions on current accounts. If a user pays money from their current into one of their savings account, such a transaction will usually be labelled in their current account as a transfer rather than a spending transaction, and thus not enter the calculation of their entropy score.

3 Spending profiles predict emergency savings

Table 3 shows the effect of entropy on the probability of building emergency savings in a given month. Columns (1)-(3) show results for unsmoothed entropy based on 9 categories, 48 categories, and merchant names, respectively. Columns (4)-(6) show results for smoothed entropy based on the same variables. All models include user and year-month fixed effects, and standard errors are clustered at the user-level. 95% confidence intervals are shown in brackets.[16]

Results for unsmoothed entropy suggest that a one standard-deviation increase in spending entropy is associated with an increase in the probability of a user making at least one transfer into their savings accounts of between 1.1 and 2.3 percentage points – an effect equal to an increase in monthly income of between £1000 and £2000. Combined, the results in the three columns present very strong evidence that entropy captures something about users's spending distribution that is related to their likelihood for making payments into their savings accounts. Furthermore, entropy variables defined based on a larger number of unique categories, that allow for a more precise segmentation of spending behaviour, capture features of spending behaviour that are as strongly or more strongly related to savings behaviour as monthly income.

The results for smoothed entropy are similar but a tend to be smaller in magnitude, and – together – also provide strong evidence that smoothed entropy captures features of spending behaviour that is related to savings behaviour in a statistically significant and economically relevant way. The main difference to the results for unsmoothed entropy is the reversal of the sign of the coefficients: across all three measures, an increase in entropy is estimated to be associated with a decrease in the likelihood of any savings transactions. The effect of the measure based on 9 categories is not meaningfully different from zero, but the estimates

[16] In line with an emerging consensus for how to avoid over-reliance on statistical significance, we deliberately do not show significance stars in our regression tables and instead emphasise the implications of the bounds of the confidence intervals when discussing our results. For two recent discussions, see Imbens (2021) and Romer (2020).

for the measures based on 48 categories and merchants are almost identical and suggest that the magnitude of the effect is between 1.2 and 2.1 percentage points – equalling the effect of a reduction of monthly income of between £1000 and £2000.

Table 3: Effect of entropy on P(savings transactions)

	(1)	(2)	(3)	(4)	(5)	(6)
Entropy (9 cats)	0.011					
	[0.005; 0.017]					
Entropy (48 cats)		0.023				
		[0.017; 0.029]				
Entropy (merchant)			0.017			
			[0.011; 0.023]			
Entropy (9 cats, smooth)				-0.004		
				[-0.007; -0.000]		
Entropy (48 cats, smooth)					-0.017	
					[-0.021; -0.012]	
Entropy (merchant, smooth)						-0.017
						[-0.021; -0.013]
Month spend (£'000)	0.010	0.009	0.010	0.010	0.009	0.008
	[0.009; 0.011]	[0.008; 0.011]	[0.008; 0.011]	[0.009; 0.011]	[0.008; 0.010]	[0.006; 0.009]
Month income (£'000)	0.012	0.012	0.012	0.012	0.012	0.012
	[0.011; 0.014]	[0.011; 0.014]	[0.011; 0.014]	[0.011; 0.014]	[0.010; 0.014]	[0.010; 0.013]
Has income in month	0.086	0.085	0.086	0.087	0.086	0.086
	[0.070; 0.102]	[0.069; 0.101]	[0.069; 0.102]	[0.071; 0.103]	[0.070; 0.102]	[0.070; 0.102]
Income variability	0.001	0.001	0.001	0.001	0.001	0.001
	[0.000; 0.002]	[0.000; 0.002]	[0.000; 0.002]	[0.000; 0.002]	[0.000; 0.002]	[0.000; 0.002]
Observations	389,790	389,790	389,790	389,790	389,790	389,790
R^2	0.46824	0.46838	0.46830	0.46820	0.46844	0.46870
Within R^2	0.00621	0.00648	0.00633	0.00615	0.00659	0.00709
User fixed effects	✓	✓	✓	✓	✓	✓
Year-month fixed effects	✓	✓	✓	✓	✓	✓

Notes: Results from estimating Equation 4. The dependent variable in all columns is a dummy variable indicating whether a user made at least one transaction into any of their savings accounts in a given period. Terms in brackets denote the upper and lower ends of a 95% confidence interval. Standard errors are clustered at the user-level.

As discussed in Section 2.6, these results are not due to reverse causality. While we might think that making a savings transactions might change some or all of the components of entropy discussed in Section 2.4 – the number of unique spending categories with positive frequency count, the standard deviation of these counts, and the total number of spend transactions – and thus change entropy, this is not the case because of the way we define entropy and savings, and the way spending transactions are categorised. We define entropy based on all current account debits that are identified as spends, while we define savings transactions as the sum of all savings accounts credits. If a user transfers money from their current account to their savings account, this will be identified as a savings transaction, but be identified as a transfer on their current account and thus not considered when calculating their entropy score.

The estimates of our control variables are largely as expected, with the exception of monthly spend, which one might have expected to be negatively correlated with savings. Also, it is evident that the strongest predictor among the included controls for whether a user makes any savings transfer is whether they receive any income in that month. Income variability, in contrast, is not correlated with savings behaviour in any economically significant way, suggesting that people with variable incomes do not build savings cushions for periods where they have no income.

Overall, the effect of entropy in spending profiles is statistically and economically significant, and robust across different definitions. In other words, the scores seem to pick up a feature of the spending distribution that is predictive of savings behaviour. The obvious question raised by the results is why smoothing entropy scores flips the direction of the effect of entropy. We address this next.

3.1 Why does smoothing flip the direction of the effect

One way to think about the sign change in Table 3 is to realise that it implies that at least for some individuals, the relative rank of smoothed and unsmoothed entropy must differ considerably – there must be some individuals that have low unsmoothed entropy but high smoothed entropy or some that have high unsmoothed entropy and low smoothed entropy or both. Understanding who those individuals are might thus help us understand the sign flip.

To understand rank differences between unsmoothed and smoothed entropy scores it is useful to rewrite Equation 1 in a way that makes it easy to see its component parts. Remember from Section 2.4 that f_c is the number of transactions made by a user in a given period in spending category c, $\mathcal{C}$ the set of all spending categories, and F the total number of transactions made by a given user in a given period. Additionally, let $\mathcal{C}^+ = \{c : f_c > 0\}$ be the set of all spending categories with positive frequency counts (i.e. with at least one transaction) and $\mathcal{C}^0 = \{c : f_c = 0\}$ the set of all spending categories with a zero frequency count, so that $\mathcal{C} = \mathcal{C}^0 \cup \mathcal{C}^+$. Then, using our definitions of unsmoothed and smoothed probabilities, we can write unsmoothed entropy as

$$H = -\sum_{c \in \mathcal{C}^+} \left(\frac{f_c}{F}\right) log \left(\frac{f_c}{F}\right), \tag{5}$$

and smoothed entropy as:

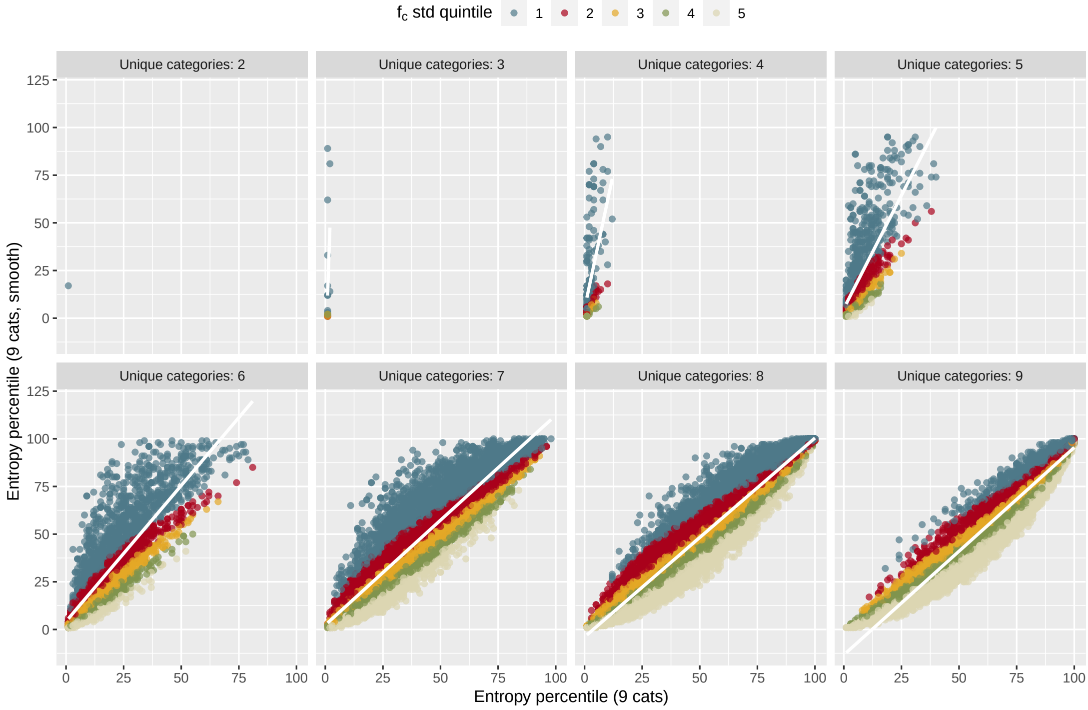
$$H^s = -\sum_{c \in \mathcal{C}^+} \left(\frac{f_c + 1}{F + |\mathcal{C}|}\right) log \left(\frac{f_c + 1}{F + |\mathcal{C}|}\right) - |\mathcal{C}^0| \left(\frac{1}{F + |\mathcal{C}|}\right) log \left(\frac{1}{F + |\mathcal{C}|}\right), \tag{6}$$

where the size of set $\mathcal{C}^0$, $|\mathcal{C}^0|$, is the number of all spending categories in which a user makes no transactions in the period. These expressions make clear that, by definition, unsmoothed entropy is a function of frequency counts of categories with positive counts only while smoothed entropy has two parts: the sum over all additively smoothed frequency counts of categories with positive counts, plus the same sum for additively smoothed probabilities of categories with zero counts, which reduces to a constant term that is multiplied by the number of such categories.

The expressions also make transparent the three main components of both types of entropy that are determined by user behaviour. The first is the number of spending categories with a non-zero frequency count, $|\mathcal{C}^+|$, which determines the number of elements summed over in Equation 5, and partitions the categories into either contributing to the sum on the left hand side of Equation 6 or – given that for an exogenously fixed $|\mathcal{C}|$, $|\mathcal{C}^0| = |\mathcal{C} \backslash \mathcal{C}^+|$ – to the constant term on the right hand side. The second component is the variation of the frequency counts, f_c, and the third component is the number of total spend transactions, $F = \sum_{\mathcal{C}} f_c$. Together, these two elements determine the probabilities associated with a spend transaction occurring in a given category. The number of total spending categories, $|\mathcal{C}|$, also determines smoothed entropy and, implicitly, also unsmoothed entropy since it "scales" the number of categories with a positive frequency count, $|\mathcal{C}^+|$, as a given number of spending transactions are categorised into finer or coarser categories. But it is exogenously given and does not depend on user behaviour.

These components help us reason about rank differences between unsmoothed and smoothed entropy. From equations 5 and 6 we can see that the first part of smoothed entropy that sums over all spending categories with positive fre-

Figure 3: Effect of smoothing on entropy



Notes: Percentile ranks of 9-category-based unsmoothed and smoothed entropy separated by the number of categories with positive frequency counts. White reference lines indicate equal percentile ranks; colours, frequency count standard deviation quintiles. Figure is based on a 10% sample of the dataset used for analysis. There are only 8 panels since there are cases where a user has spends in only a single category.

quency counts is very similar to the entire expression of unsmoothed entropy – it is that same expression but with additively smoothed probabilities. Hence, all else equal, the higher the number of categories with positive counts, the more smoothed entropy is determined by that first part, and the more similar it will be to unsmoothed entropy. As a result, we would expect to find large (rank) differences between entropy scores among cases with few positive-counts categories. Furthermore, given that unsmoothed entropy is increasing in the number of categories with positive counts, we expect these cases to have relatively low unsmoothed entropy.

Given that we expect high rank difference cases to make spends across a small number of categories, and expect these cases to have low unsmoothed entropy, high rank differences will occur among the subset of those cases that have high smoothed entropy. Remember from Section 2.4 that entropy is higher the more equal the spending category probabilities are. Hence, for a given number of zero-count categories, smoothed entropy will be higher if the (additively smoothed) probabilities of all positive-count categories are close to the (additively smoothed) probabilities of the zero-count categories, which will be the case (i) if there are few overall transactions, such that frequency counts (f_c) are close to zero and (ii) if the counts are similar.

Figure 3 visualises this entire line of reasoning for our 9-category based entropy variable: it shows scatter plots of the percentile ranks of unsmoothed and smoothed entropy with a reference line indicating identical rank, separated by the number of categories with positive frequency counts and coloured based on the quintile of the frequency count standard deviation.[17] First, ignoring the colouring and focusing on the shape of the dots only we can see that, as expected, the relationship between the two entropy measures is tighter the higher the number of non-zero spending categories is. Cases with large entropy rank differences are thus to be found among cases with fewer positive spend categories. Among these, the colouring makes clear that, as expected, the cases with the largest rank differences – those with the furthest vertical difference to the reference line – have low variation in their frequency counts. However, it is also clear that the reverse is not true: there are cases with low count frequency variation that experience little or even a negative rank difference. Furthermore, among cases with a higher number of categories with positive counts, cases with a very high variation in frequency counts can also experience quite large rank differences, albeit in the opposite direction. Hence, while counts variation is some help in identifying cases with high entropy rank differences, it does not do so perfectly.

One reason the relationship is not perfect is that while cases with few total transactions that are evenly spread across a small number of categories will have high smoothed entropy, such will also have high unsmoothed entropy. What we are looking for, more precisely, are cases where this pattern holds broadly, but that also have a small number of frequency counts that dominate the others but are still relatively close to zero. This would lead to low unsmoothed entropy (because of the dominant counts) but high smoothed entropy (because all counts are still relatively close to zero, so that the additively smoothed distribution resembles a uniform distribution). Further research will be necessary to more fully classify such cases, and to inquire what it is about their behaviour that might be related to savings behaviour.

3.2 Is entropy informative beyond its component parts?

Another question that arises once we think of entropy as comprising three main component parts is whether its non-linear nature captures anything about the relationship between spending and saving behaviour that is not captured already by the three simpler components.

[17] Colouring based on the quintile of the total number of transactions leads to a very similar plot and is thus not shown.

Table 4: Controlling for components

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Entropy (9 cats)	0.011	0.004						
	[0.005; 0.017]	[-0.003; 0.011]						
Entropy (9 cats, smooth)			-0.004	-0.004				
			[-0.007; -0.000]	[-0.009; 0.001]				
Entropy (48 cats)					0.023	0.020		
					[0.017; 0.029]	[0.005; 0.034]		
Entropy (48 cats, smooth)							-0.017	-0.021
							[-0.021; -0.012]	[-0.031; -0.010]
Unique categories ($\lvert\mathcal{C}^+\rvert$)		0.004		0.004		0.002		0.004
		[0.002; 0.005]		[0.003; 0.005]		[0.001; 0.004]		[0.003; 0.005]
Category counts std.		-0.001		-0.003		0.004		-0.010
		[-0.005; 0.003]		[-0.008; 0.001]		[-0.002; 0.010]		[-0.016; -0.005]
Number of spend txns (F)		0.001		0.001		0.000		0.001
		[0.000; 0.001]		[0.000; 0.001]		[-0.000; 0.001]		[0.000; 0.001]
Month spend	0.010	0.006	0.010	0.006	0.009	0.006	0.009	0.006
	[0.009; 0.011]	[0.005; 0.007]	[0.009; 0.011]	[0.005; 0.007]	[0.008; 0.011]	[0.005; 0.007]	[0.008; 0.010]	[0.005; 0.007]
Month income	0.012	0.011	0.012	0.011	0.012	0.011	0.012	0.011
	[0.011; 0.014]	[0.009; 0.013]	[0.011; 0.014]	[0.009; 0.013]	[0.011; 0.014]	[0.009; 0.013]	[0.010; 0.014]	[0.009; 0.013]
Has income in month	0.086	0.082	0.087	0.082	0.085	0.081	0.086	0.081
	[0.070; 0.102]	[0.066; 0.098]	[0.071; 0.103]	[0.066; 0.098]	[0.069; 0.101]	[0.065; 0.097]	[0.070; 0.102]	[0.065; 0.097]
Income variability	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	[0.000; 0.002]	[0.000; 0.002]	[0.000; 0.002]	[0.000; 0.002]	[0.000; 0.002]	[0.000; 0.002]	[0.000; 0.002]	[0.000; 0.002]
Observations	389,790	389,790	389,790	389,790	389,790	389,790	389,790	389,790
R^2	0.46824	0.46908	0.46820	0.46908	0.46838	0.46910	0.46844	0.46914
Within R^2	0.00621	0.00778	0.00615	0.00779	0.00648	0.00783	0.00659	0.00790
User fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Year-month fixed effects	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Terms in brackets denote the upper and lower ends of a 95% confidence interval. Standard errors are clustered at the user-level.

We can test this by checking whether the relationship between spending entropy and the probability of making a savings transaction remains economically and statistically significant once we control for the three components. Columns (1) and (3) in Table 4 replicate the results for the 48-category-based unsmoothed and smoothed entropy measures presented in Table 3 for reference. In columns (2) and (4) we additionally control for the three entropy components. Including these components has some effect: the coefficients change slightly – decreasing in absolute magnitude in the case of unsmoothed entropy, increasing in the case of smoothed entropy – while the width of the confidence intervals about double in both cases, reflecting the strong collinearity amount the component and entropy. However, both coefficients remain statistically significant and their confidence intervals cover values that are also economically significant. Hence, the results make clear that the results in Table 3 cannot be attributed simply to the effect of one or more of entropy's simple components.

4 Conclusion

In this paper, we calculate simple summary statistics for spending profiles for a large number of users of a UK-based financial aggregator app and show that the predictability of spending, captured using entropy, is predictive of the frequency with which users make transactions into their savings accounts. The effect of spending entropy on savings is statistically significant and economically large, and robust across a number of different specifications and entropy measures based on different product and merchant categories. The direction of the effect depends, however, on the handling of categories in which individuals make no transactions in a given month. While our results give some indication as to how we can explain this result, further research is needed.

As part of a fuller understanding of said sign reversal, more work generally is needed on understanding to what extend entropy is useful in capturing meaningful information about individuals' lives, and what particular ways of calculating entropy are most appropriate and informative in different circumstances. For instance, while, in this paper, we have calculated entropy based on the number of transactions in a given category, there are a number of alternatives. We could calculate profiles based on the distribution of transaction values rather than counts. We could also calculate profiles based on inter-temporal rather than intra-temporal distributions, focusing on consistency of purchasing behaviour over time rather than on predictability at any given time (Krumme et al. 2013). Further, we could focus on time-based rather than category-based measures, focusing, for instance, on whether purchases of the same type tend to occur on the same day of the week (Guidotti et al. 2015). Finally, one could also create composite measures based on principal component analysis, an approach used in Eagle et al. (2010).

Finally, work like ours inherently risks portraying low savings mainly as a result of behavioural biases, implying that if can help people overcome these biases, they will be able to save more. However, as convincingly argued in Chater and Loewenstein (2022), for problems where the main solution lies not in individual behaviour change but public policies, such an approach runs the risk of moving attention and resources away from required society-wide action. Research by the Money and Pension Service in the UK (MPS 2018) and the Consumer Financial Protection Bureau (CFPB 2017) suggests that while low levels of savings are – unsurprisingly – often concentrated among very low income individuals or households, individual behaviour still plays a large role in explaining differences in savings levels across income groups. Hence, addressing low levels of savings is a matter for both public policy and individual behaviour change. But it is important to ensure that focusing on the latter does not distract from the former.

References

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas D Evanoff (2010). "Learning to cope: Voluntary financial education and loan performance during a housing crisis". In: *American Economic Review* 100.2, pp. 495–500.

Agarwal, Sumit, Souphala Chomsisengphet, and Cheryl Lim (2017). "What shapes consumer choice and financial products? A Review". In: *Annual Review of Financial Economics* 9, pp. 127–146.

Agarwal, Sumit, John C Driscoll, Xavier Gabaix, and David Laibson (2009). "The age of reason: Financial decisions over the life cycle and implications for regulation". In: *Brookings papers on Economic activity* 2009.2, pp. 51–117.

Agarwal, Sumit and Bhashkar Mazumder (2013). "Cognitive abilities and household financial decision making". In: *American Economic Journal: Applied Economics* 5.1, pp. 193–207.

Ameriks, John, Andrew Caplin, and John Leahy (2003). "Wealth accumulation and the propensity to plan". In: *The Quarterly Journal of Economics* 118.3, pp. 1007–1047.

Bailey, Michael, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel (2018). "The economic effects of social networks: Evidence from the housing market". In: *Journal of Political Economy* 126.6, pp. 2224–2276.

Baker, Scott R (2018). "Debt and the response to household income shocks: Validation and application of linked financial account data". In: *Journal of Political Economy* 126.4, pp. 1504–1557.

Baker, Scott R and Lorenz Kueng (2022). "Household financial transaction data". In: *Annual Review of Economics* 14, pp. 47–67.

Baugh, Brian, Itzhak Ben-David, and Hoonsuk Park (2014). "Disentangling financial constraints, precautionary savings, and myopia: household behavior surrounding federal tax returns". Tech. rep. National Bureau of Economic Research.

Beshears, John, James J Choi, J Mark Iwry, David C John, David Laibson, and Brigitte C Madrian (2020). "Building emergency savings through employer-sponsored rainy-day savings accounts". In: *Tax Policy and the Economy* 34.1, pp. 43–90.

Beshears, John, James J Choi, David Laibson, and Brigitte C Madrian (2009). "The importance of default options for retirement saving outcomes: Evidence from the United States". In: *Social security policy in a changing environment.* University of Chicago Press, pp. 167–195.

Blumenstock, Joshua, Michael Callen, and Tarek Ghani (2018). "Why do defaults affect behavior? Experimental evidence from Afghanistan". In: *American Economic Review* 108.10, pp. 2868–2901.

Bourquin, Pascale, Isaac Delestre, Robert Joyce, Imran Rasul, and Tom Walters (2020). "The effects of coronavirus on household finances and financial distress". In: *IFS Briefing Note BN298*.

Carlin, Bruce, Arna Olafsson, and Michaela Pagel (2019). "Generational Differences in Managing Personal Finances". In: *AEA Papers and Proceedings*. Vol. 109, pp. 54–59.

CFPB, Consumer Financial Protection Bureau (2017). "Financial Well-being in America". Tech. rep. URL: https://www.consumerfinance.gov/data-research/research-reports/financial-well-being-america/.

Chater, Nick and George Loewenstein (2022). "The i-frame and the s-frame: How focusing on the individual-level solutions has led behavioral public policy astray". In: *Available at SSRN 4046264*.

Choi, James J, David Laibson, Brigitte C Madrian, and Andrew Metrick (2002). "Defined contribution pensions: Plan rules, participant choices, and the path of least resistance". In: *Tax policy and the economy* 16, pp. 67–113.

— (2004). "For better or for worse: Default effects and 401 (k) savings behavior". In: *Perspectives on the Economics of Aging*. University of Chicago Press, pp. 81–126.

Cohen, Jonathan, Keith Marzilli Ericson, David Laibson, and John Myles White (2020). "Measuring time preferences". In: *Journal of Economic Literature* 58.2, pp. 299–347.

De Mel, Suresh, Craig McIntosh, and Christopher Woodruff (2013). "Deposit collecting: Unbundling the role of frequency, salience, and habit formation in generating savings". In: *American Economic Review* 103.3, pp. 387–92.

Eagle, Nathan, Michael Macy, and Rob Claxton (2010). "Network diversity and economic development". In: *Science* 328.5981, pp. 1029–1031.

Ericson, Keith Marzilli and David Laibson (2019). "Intertemporal choice". In: *Handbook of Behavioral Economics: Applications and Foundations 1*. Vol. 2. Elsevier, pp. 1–67.

Fernandes, Daniel, John G Lynch Jr, and Richard G Netemeyer (2014). "Financial literacy, financial education, and downstream financial behaviors". In: *Management Science* 60.8, pp. 1861–1883.

Frederick, Shane, George Loewenstein, and Ted O'Donoghue (2002). "Time discounting and time preference: A critical review". In: *Journal of economic literature* 40.2, pp. 351–401.

Ganong, Peter and Pascal Noel (2019). "Consumer spending during unemployment: Positive and normative implications". In: *American Economic Review* 109.7, pp. 2383–2424.

Gelman, Michael, Shachar Kariv, Matthew D Shapiro, Dan Silverman, and Steven Tadelis (2014). "Harnessing naturally occurring data to measure the response of spending to income". In: *Science* 345.6193, pp. 212–215.

Governors of the Federal Reserve System, Board of (2022). "Economic Well-Being of U.S. Households in 2021". Tech. rep.

Greenberg, Adam Eric and Hal Hershfield (2019). "Financial decision making". In: *Consumer Psychology Review* 2.1, pp. 17–29.

Guidotti, Riccardo, Michele Coscia, Dino Pedreschi, and Diego Pennacchioli (2015). "Behavioral entropy and profitability in retail". In: *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. IEEE, pp. 1–10.

Hacıoğlu-Hoke, Sinem, Diego R Känzig, and Paolo Surico (2021). "The distributional impact of the pandemic". In: *European Economic Review* 134, p. 103680.

Hirsh, Jacob B, Raymond A Mar, and Jordan B Peterson (2012). "Psychological entropy: a framework for understanding uncertainty-related anxiety." In: *Psychological review* 119.2, p. 304.

Imbens, Guido W (2021). "Statistical significance, p-values, and the reporting of uncertainty". In: *Journal of Economic Perspectives* 35.3, pp. 157–74.

Jørring, Adam (2020). "Financial sophistication and consumer spending". Tech. rep. Working Paper.

Korniotis, George M and Alok Kumar (2011). "Do older investors make better investment decisions?" In: *The Review of Economics and Statistics* 93.1, pp. 244–265.

Krumme, Coco, Alejandro Llorente, Manuel Cebrian, Alex Pentland, and Esteban Moro (2013). "The predictability of consumer visitation patterns". In: *Scientific reports* 3.1, pp. 1–5.

Kuchler, Theresa and Michaela Pagel (2020). "Sticking To Your Plan: The Role of Present Bias for Credit Card Debt Paydown". In: *Journal of Financial Economics, forthcoming.*

Kuchler, Theresa and Johannes Stroebel (2021). "Social finance". In: *Annual Review of Financial Economics* 13, pp. 37–55.

Kumar, Alok (2009). "Who gambles in the stock market?" In: *The Journal of Finance* 64.4, pp. 1889–1933.

Kumar, Alok and Sonya Seongyeon Lim (2008). "How do decision frames influence the stock investment choices of individual investors?" In: *Management science* 54.6, pp. 1052–1064.

Madrian, Brigitte C and Dennis F Shea (2001). "The power of suggestion: Inertia in 401 (k) participation and savings behavior". In: *The Quarterly journal of economics* 116.4, pp. 1149–1187.

Meyer, Steffen and Michaela Pagel (2018). "Fully closed: Individual responses to realized capital gains and losses". Tech. rep. Working paper.

Mittal, Chiraag, Vladas Griskevicius, Jeffry A Simpson, Sooyeon Sung, and Ethan S Young (2015). "Cognitive adaptations to stressful environments: When childhood adversity enhances adult executive function." In: *Journal of personality and social psychology* 109.4, p. 604.

MPS, Money and Pension Service (2018). "Building the Financial Capability of UK Adults". Tech. rep. URL: https://moneyandpensionsservice.org.uk/2019/02/06/adult-financial-capability-building-the-financial-capability-of-uk-adults-survey/.

Muggleton, Naomi K, Edika G Quispe-Torreblanca, David Leake, John Gathergood, and Neil Stewart (2020). "Evidence from mass-transactional data that chaotic spending behaviour precedes consumer financial distress". Tech. rep. DOI: 10.31234/osf.io/qabgm. URL: psyarxiv.com/qabgm.

Olafsson, Arna and Michaela Pagel (2018). "The liquid hand-to-mouth: Evidence from personal finance management software". In: *The Review of Financial Studies* 31.11, pp. 4398–4446.

Perry, Vanessa G and Marlene D Morris (2005). "Who is in control? The role of self-perception, knowledge, and income in explaining consumer financial behavior". In: *Journal of consumer affairs* 39.2, pp. 299–313.

Phillips, Jo, Annick Kuipers, and Will Sandbrook (2021). "Supporting emergency savings: early learnings of the employee experience of workplace sidecar savings". Tech. rep.

Puri, Manju and David T Robinson (2007). "Optimism and economic choice". In: *Journal of financial economics* 86.1, pp. 71–99.

Read, Daniel, Rebecca McDonald, and Lisheng He (2018). "Intertemporal choice: Choosing for the future". In: *The Cambridge handbook of psychology and economic behavior*, pp. 167–197.

Rick, Scott I, Cynthia E Cryder, and George Loewenstein (2008). "Tightwads and spendthrifts". In: *Journal of consumer research* 34.6, pp. 767–782.

Rick, Scott I, Deborah A Small, and Eli J Finkel (2011). "Fatal (fiscal) attraction: Spendthrifts and tightwads in marriage". In: *Journal of Marketing Research* 48.2, pp. 228–237.

Romer, David (2020). "In praise of confidence intervals". In: *AEA Papers and Proceedings*. Vol. 110, pp. 55–60.

Schaner, Simone (2018). “The persistent power of behavioral change: Long-run impacts of temporary savings subsidies for the poor”. In: *American Economic Journal: Applied Economics* 10.3, pp. 67–100.

Shannon, Claude Elwood (1948). “A mathematical theory of communication”. In: *The Bell system technical journal* 27.3, pp. 379–423.

Skatova, Anya, Neil Stewart, Edward Flavahan, and James Goulding (2019). “Those Whose Calorie Consumption Varies Most Eat Most”. In:

Vernon-Feagans, Lynne, Michael Willoughby, and Patricia Garrett-Peters (2016). “Predictors of behavioral regulation in kindergarten: Household chaos, parenting, and early executive functions.” In: *Developmental psychology* 52.3, p. 430.

A Interpreting entropy

To see how we can interpret entropy as the predictability of a user's spendin that we discussed above behaviour, it is useful to have a more complete understanding of Equation 1. The building blocks of entropy is the information content of a single event. The key intuition Shannon (1948) aimed to capture was that learning of the occurrence of a low-probability event is more informative than learning of the occurrence of a high-probability event. The information of an event $I(E)$ is thus inversely proportional to is probability $p(E)$. One way to capture this would be to define the information of event E as $I(E) = \frac{1}{p(E)}$. Yet this implied that an event that is certain to occur had information 1, when it would make sense to have information 0. To remedy this (and also satisfy additional desireable characteristics of an information function), Shannon proposed using the log of the expression. Hence, the information of event E, often called *Shannon information*, *self-information*, or just *information*, is defined as:

$$I(E) = log\left(\frac{1}{p(E)}\right) = -log(p(E)). \tag{7}$$

The choice of the base for the logarithm varies by application and determines the units: base 2 means that information is expressed in bits; the natural logarithm, another popular choice, expresses information in *nats*.

Entropy, often called *Information entropy*, *Shannon entropy*, or just *entropy*, is the information of a random variable, X, and captures the expected amount of information of an event drawn at random from the probability distribution of the random variable. It is calcualted as:

$$H(X) = -\sum_{x} p(x) \times log(p(x)) = \sum_{x} p(x)I(x) = \mathbb{E}I(x). \tag{8}$$

For a single event, the key intution was that the less likely an event, the more information is conveyed when it occurs. The related idea for distributions is similar: the less skewed a distribution of a random variable, the less certain the realised value of a single draw from the distribution, the higher is entropy - the maximum entropy distribution is the uniform distribution.

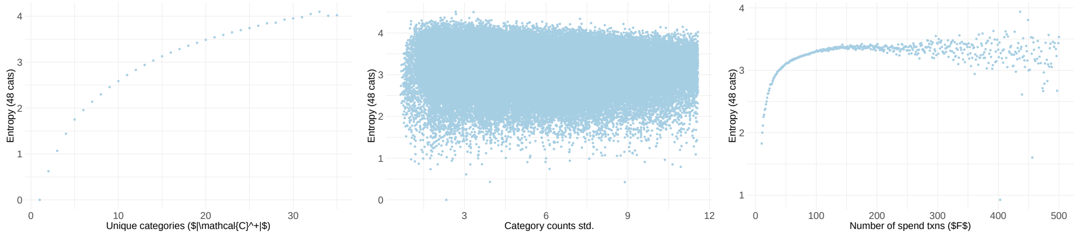
B Entropy components

Figure 4 shows the empirical relationship with our 48-categories-based unsmoothed entropy variable and these three components.[18] We can see that for the values

[18]To highlight the main features of the relationships we have trimmed the component values at the 95th percentile.

we observe in the dataset, entropy increases monotonically in the number of unique spending categories with positive frequency counts, has no clear relationship with the standard deviation of those counts, and increases in the number of total spending transactions up to about 175 transaction, before being increasingly determined by other elements thereafter.

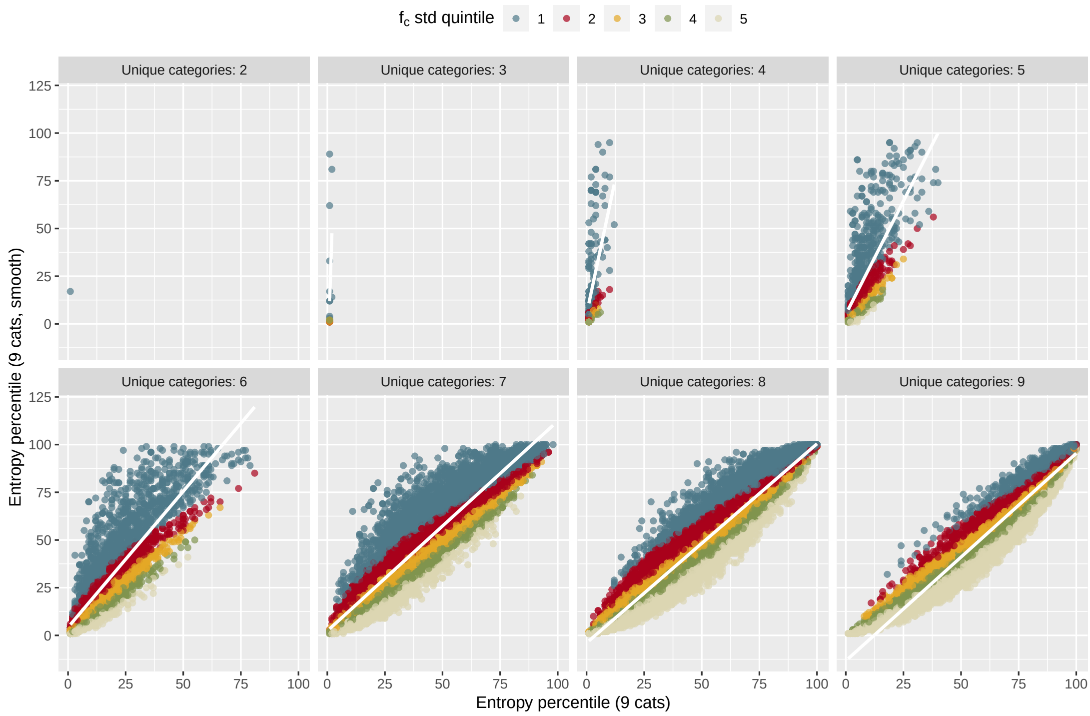
Figure 4: Correlation of entropy with its components



Notes: Correlation of 48-categories-based unsmoothed entropy with its three main components: the number of unique spending categories with positive frequency counts (left), the standard deviation of those frequency counts (middle), and the number of total spend transactions (left).

C Effect of smoothing on entropy

Figure 5: Effect of smoothing on entropy



Notes: Percentile ranks of 9-category-based unsmoothed and smoothed entropy separated by the number of categories with positive frequency counts. White reference lines indicate equal percentile ranks. Colours indicate the quintile of the total number of spending transactions.