

Spending profiles predict savings*

Fabian Gunzinger

Neil Stewart

September 14, 2022

Contents

1	Introduction	2
2	Methods	3
2.1	Dataset description	3
2.2	Preprocessing and sample selection	4
2.3	Dependent variables	5
2.4	Spending profiles	5
2.5	Summary statistics	6
2.6	Model specification	8
3	Spending profiles	9
4	Savings patterns	11
5	Spending profiles predict emergency savings	11
5.1	Effect of financial resilience	12
6	Discussion	14
A	Understanding entropy	17
B	Additional results	17
B.1	Endogeneity	17
B.2	Further results by financial resilience	18

*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.

1 Introduction

This paper documents variations in spending profiles and payments into emergency savings for a large set of users of a financial management app and shows that spending profiles predict emergency savings.

We define emergency savings as inflows into savings accounts. These savings will be made up of savings for particular goals – a new car, a holiday, a wedding – and savings directed towards building up a buffer for financial emergencies. Because in our data we cannot distinguish between these two cases, we refer to all of them as emergency savings.¹ These short-term 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 well-being, yet while there is a large literature on pension savings, little is known about how people save for the short-term.²

Studying the determinants of emergency savings is important because 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 (Philipps 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). Similarly, research in the UK has shown that having £1000 in savings reduces by more than half a household’s chances of falling into debt that leads to financial problems (Philipps et al. 2021).

Studying spending profiles is of interest because:

- Our understanding of how people spend their money is based on survey data.
- Large-scale transaction-level data offers the possibility to study spending behaviour based on real-time data that are automatically collected for a large number of users. Such data has only become available very recently and have not, thus far, been used to investigate systematically how people spend their money.
- Research in psychology suggests that disorder is maladaptive and 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 the study of human behaviour, more chaotic behaviour has been found to predict the a higher number of visits to and higher spend in supermarkets (Guidotti et al. 2015), higher calorie intake (Skatova et al. 2019) and financial distress (Muggleton et al. 2020).

¹MDB allows users to create custom tags and some users use them to indicate the intended use for their savings transactions (e.g. “wedding”, “holidays”). But only a very small number of transactions have such tags, and we do not pursue this further.

²Well-documented behavioural biases that help explain undersaving for pensions are, among others, present bias (Laibson 1997, Laibson and Marzilli-Ericson 2019), inertia (Madrian and Shea 2001), over-extrapolation (Choi et al. 2009), and limited self-control and willpower (Thaler and Shefrin 1981, Benhabib and Bisin 2005, Fudenberg and Levine 2006, Loewenstein and O’Donoghue 2004, Gul and Pesendorfer 2001). One danger of viewing low savings mainly as a result of behavioural biases is that while these biases likely do play some role and designing environments and tools to help correct them are thus part of the solution, it is at least conceivable that this is an area where the focus on behaviour-level solutions distracts from an effort to find more effective society-level solutions, a danger inherent in behavioural science research convincingly highlighted in Chater and Loewenstein (2022): if the main problem is that many people are unable to earn enough to save, then the effectiveness of helping them manage their low incomes more effectively pales in comparison with efforts to help them earn more.

We hypothesise that less predictable spending patterns are associated with a lower probability for making payments into emergency savings accounts. Possible channels:

- Disorder (personal life or environment): leads to more impulsive shopping behaviour and makes forgetting to save more likely.
- Scarcity: life challenges focus attention away from deliberate shopping, causing more impulse purchases, and make forgetting to save more likely.

What we do:

- Systematically documenting emergency savings patterns.
- Systematically documenting variation in spending profiles.
- Showing that unpredictability in spending profiles is associated with lower emergency savings.

Contribution to literatures:

- Understanding emergency savings behaviour (nest, aspen reports), (Sabat and Gallagher 2019) for sources on short-term savings literature, Colby and Chapman (2013) for lit on savings goals. See Colby and Chapman (2013) has useful literature review on short-term savings and suggests that subgoals can increase willingness to forego short-amounts in the present because they move the reference point in a prospect-theory framework. Philipps et al. (2021) present results from an employer-linked initiative that offers employees to have a portion of their salary automatically transferred into a savings pot. Policy literature: (CAN 2019, CFPB 2017, MPS 2018). Older literature:Savings lit: Lunt and Livingstone (1991) and Oaten and Cheng (2007)
- Understanding effect of behavioural entropy - eliciting useful personality characteristics from large-scale data.
- Use of high-frequency transaction data (itself a sub-literature of use of newly available large-scale datasets).

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.³ The dataset contains more than 500 million transactions made between 2012 and June 2020 by about 250,000 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 is automatically collected and updated and thus less prone to errors and unaffected

³<https://www.moneydashboard.com>.

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 disproportionately 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.⁴

Data issues Bourquin et al. (2020) argue that because some of the accounts in the data will be joint accounts, units of observations should be thought of as "households" rather than "users". We do not agree that this is the most prudent approach. The validity of thinking of units as households depends on the proportion of users in the data who add joint accounts and on the proportion of transactions – out of a user's total number of transactions – additionally observed as a result. Given that the sample is skewed towards younger individuals we think it is unlikely that a majority of them has added joint accounts. Furthermore, it seems reasonable to assume that in most cases, joint accounts are mainly used for common household expenditures similar that are similar to those of a single user (albeit in higher amounts), and are thus unlikely to alter the observed spending profile much. Thus, we think of units of observations as individuals, not households.

Some accounts might be business accounts. Using versions of the algorithms used by Bourquin et al. (2020) to identify such accounts showed, however, that such accounts only make up a tiny percentage of overall accounts and would not influence our results. We thus do not exclude them.

2.2 Preprocessing and sample selection

We restrict our sample to users for whom we can observe a regular income, can be reasonably sure that they have added all their bank account to MDB, and for whom we observe at least six months of data. Table 1 summarises the sample selection steps we applied to a 1 percent sample of the raw data, associated data losses, and the size of our final sample.

⁴For an example of how re-weighting can be used to mitigate the non-representative issue, see Bourquin et al. (2020).

Table 1: Sample selection

	Users	User-months	Txns	Txns (m£)
Raw sample	27,175	795,338	65,972,558	12,527
Drop first and last month	26,565	741,170	64,157,932	12,179
At least 6 months of data	23,238	732,499	63,592,713	12,074
At least one savings account	14,315	473,814	43,983,467	8,898
At least one current account	14,028	466,827	43,408,017	8,792
At least £5,000 of annual income	5,335	159,663	16,815,335	3,339
At least 10 txns each month	4,789	142,602	15,328,431	3,023
At least £200 of monthly spend	4,175	122,174	13,592,387	2,713
Complete demographic information	3,417	104,660	11,689,245	2,299
Drop test users	3,410	104,307	11,638,106	2,292
Working age	3,345	102,302	11,465,704	2,241
Final sample	3,345	102,302	11,465,704	2,241

2.3 Dependent variables

Identifying savings transactions: We classify as payments into savings accounts all savings account credits of £5 or more that are not identified as interest payments or automated "save the change" transfers (similarly for debits).⁵

Dummy for savings txn in current month. Motivation: MPS (2018) finds that saving habit is often more important than amount saved.

2.4 Spending profiles

We define a user's spending profile as the distribution of the number of spending transactions across different spend categories.⁶ To summarise these distributions, we calculate spending entropy, based on the formula proposed by Shannon (1948), which defines entropy as:⁷

$$H = - \sum p_i \log(p_i), \quad (1)$$

where p_i is the probability that an individual makes a purchase in spending category i , and \log is the base 2 logarithm.⁸

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

⁵While standing order transactions are unlikely to be related to entropy in the short-run, we do not exclude such transactions since, best we can tell, the only account for a small fraction of total transactions.

⁶There are a number of alternative ways to characterise spend profiles. 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). We leave these extensions for future research.

⁷Shannon 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.

⁸The choice of the base for the logarithm varies by application and determines the units of $I(E)$. Base 2 means that information is expressed in bits. The natural logarithm, another popular choice, expresses information in *nats*.

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 the context of spending profiles, 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.⁹

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).¹⁰ The second measure is based on our own, more fine-grained, categorisation into 48 different categories.¹¹ The third measure is based on merchant names, as labelled by Money Dashboard.

We also handle categories with zero transaction counts in two different ways. To calculate what we call “unsmoothed” entropy scores, we calculate the p_i s in Equation 1 as simple frequentist probabilities

$$p_i^{us} = \frac{T_i}{\sum_{i=1}^N T_i}, \quad (2)$$

where T_i is the number of transactions in spend category i and N the number of categories. 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.¹² To calculate “smoothed” entropy scores, we apply additive smoothing to calculate probabilities as

$$p_i^s = \frac{T_i + 1}{\sum_{i=1}^N T_i + N}. \quad (3)$$

That is: we add 1 to each category’s transaction count and divide by the sum of the total number of transactions and the number of categories. Because categories with a zero transaction count will have a numerator of 1, the sum in Equation 1 will be taken over all categories.

One slight 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.

2.5 Summary statistics

Table 2 provides summary statistics.

⁹For further discussion on how to interpret Equation 1, see Appendix ??.

¹⁰The precise mapping from MDB transaction tags into these 9 categories is available on [Github](#).

¹¹The precise mapping from MDB transaction tags into these 48 categories is available on [Github](#).

¹²This is automatically handled by the entropy [implementation](#) of Python’s SciPy package, which is what we use to calculate entropy scores.

Table 2: Summary statistics

Statistic	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Year income	31.34	44.02	0.43	14.43	22.81	37.05	4,229.21
Income variability	1.23	4.87	0.00	0.21	0.54	1.16	559.67
Has income in month	0.98	0.15	0	1	1	1	1
Has savings	0.48	0.50	0	0	0	1	1
Month spend	2.70	2.65	0.20	1.16	1.91	3.19	17.44
Age	35.14	10.27	18	27	33	42	65
Female	0.39	0.49	0	0	0	1	1
Urban	0.85	0.35	0	1	1	1	1
Unique categories (9)	7.51	1.26	1	7	8	8	9
Unique categories (48)	15.15	4.43	1	12	15	18	35
Unique categories (Merchants)	22.99	9.77	0	16	22	29	85

Notes: Income and spend variables in '000s of Pounds, number of unique categories for spend transaction classification based on 9 categories, 48 categories, and merchant names.

Figure 1

Figure 1: Demographic characteristics of Money Dashboard users

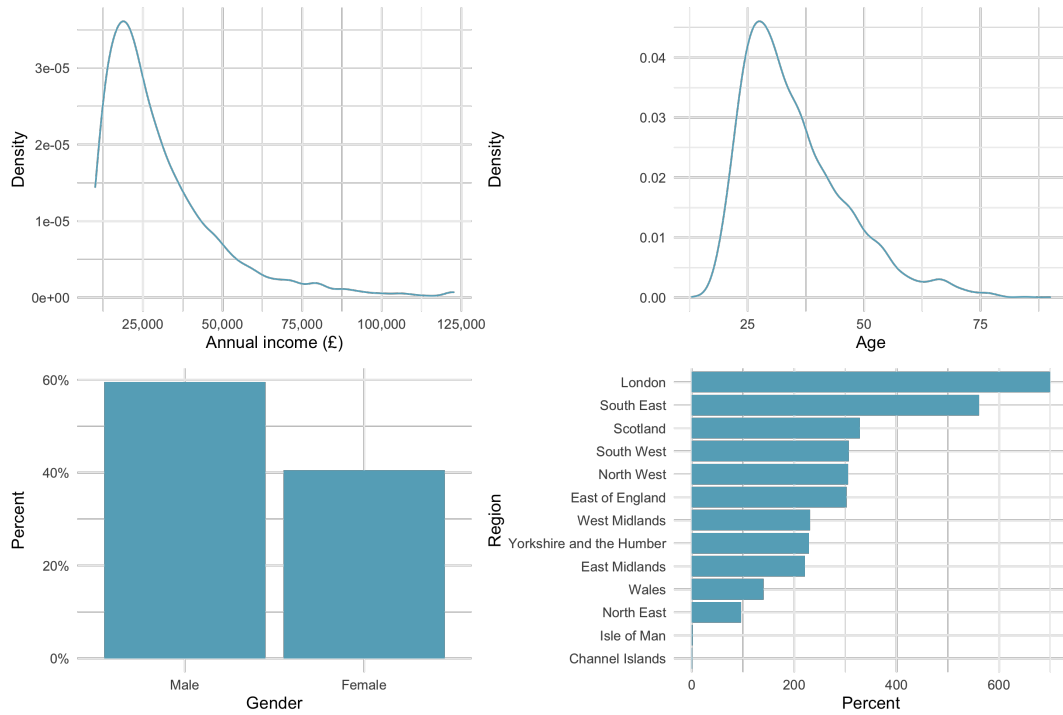
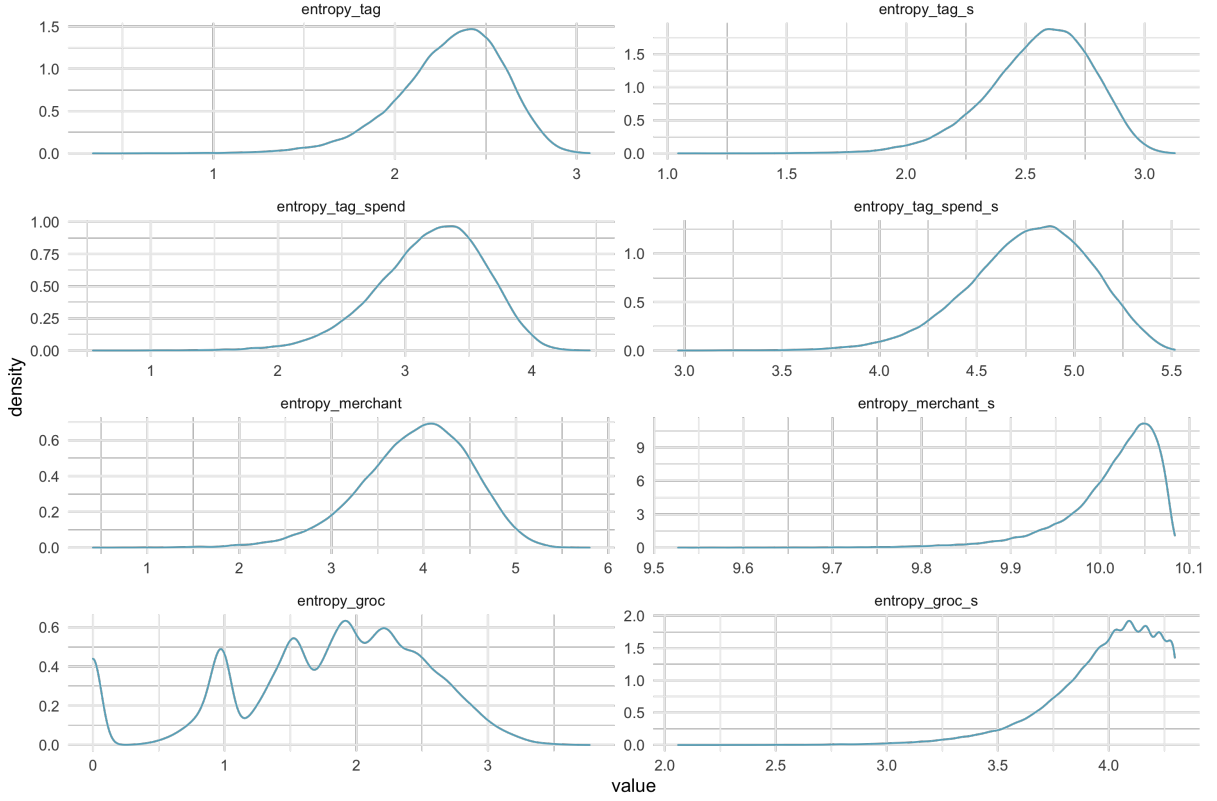


Figure 2

Figure 2: Entropy distributions



Notes:

2.6 Model specification

We estimate models of the form:

$$y_{i,t} = \alpha_i + \lambda_t + \beta H_{i,t} + x'_{i,t} \delta + \epsilon_{i,t}, \quad (4)$$

where $y_{i,t}$ is an indicator variable equal to one if individual i made one or more transfers to any of their savings account in year-month period t and zero otherwise, $H_{i,t}$ 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.

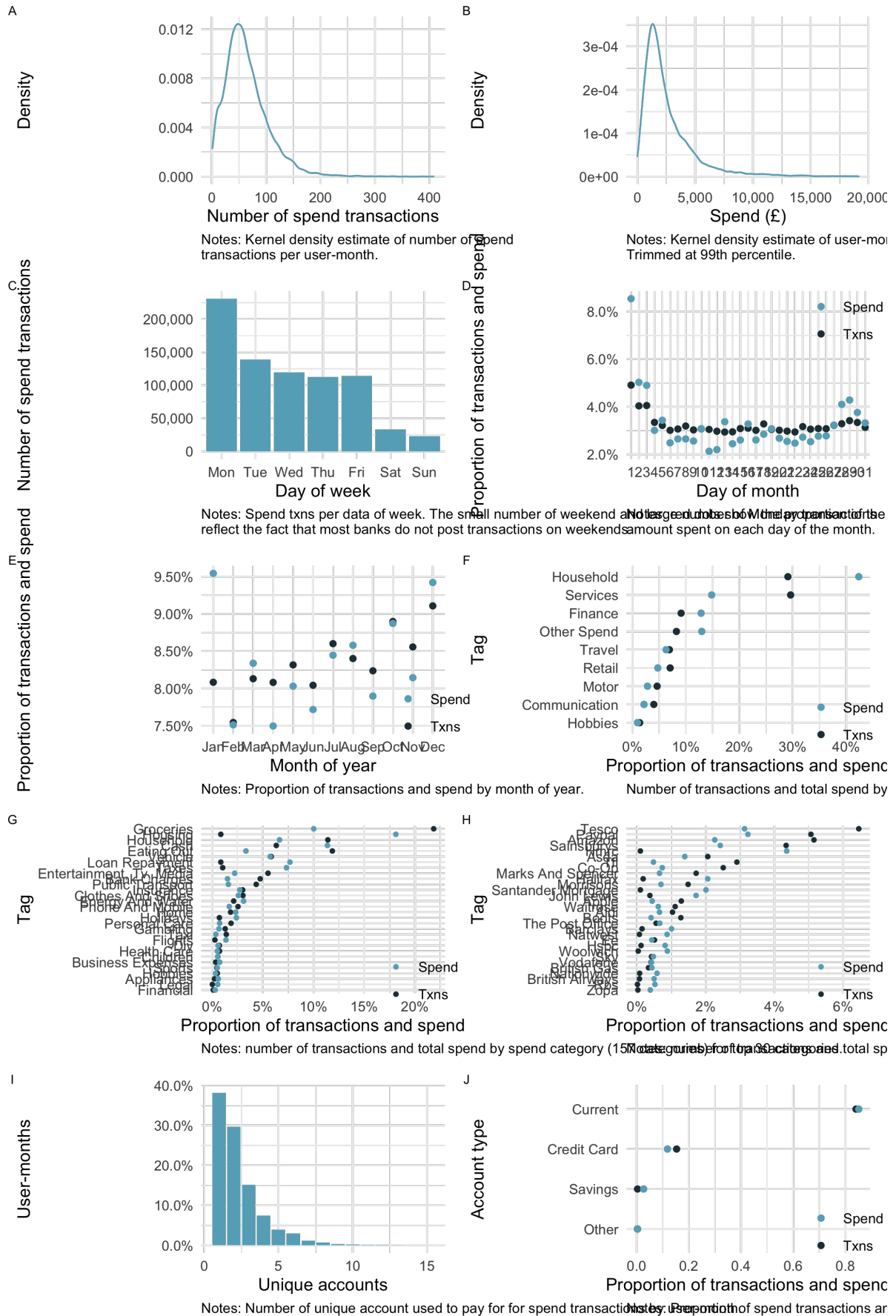
Note that while we might in principle be worried about reverse causality, since making payments into savings accounts might lead to a non-zero count in an additional spend category and thus change entropy, this is not a concern here. As discussed in Section 2.3 and Section 2.4, we define savings as inflows 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, this will usually be labelled in their current account as a transfer and not enter the calculation of their entropy score. In Appendix B.1, we provide robustness checks using lagged entropy scores and controlling for the number of non-zero categories, which broadly leave

the results unchanged.

3 Spending profiles

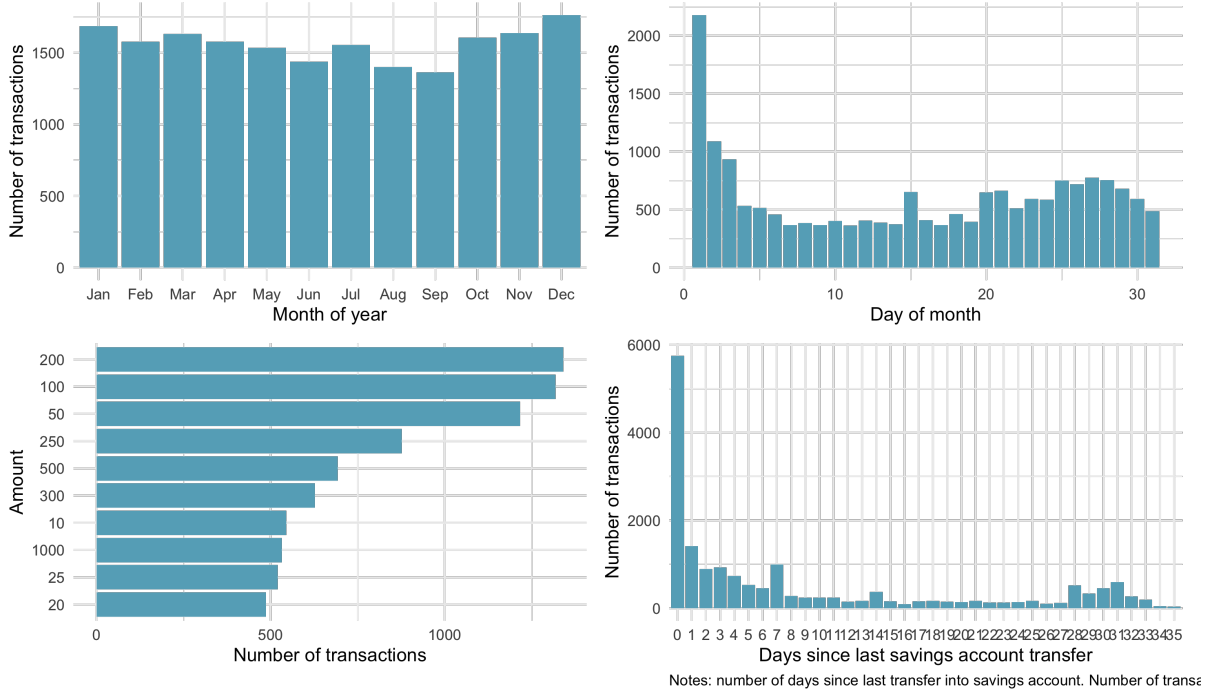
Figure 3

Figure 3: Spending behaviour



4 Savings patterns

Figure 4: Savings behaviour



5 Spending profiles predict emergency savings

Table 3 shows the effect of entropy on the probability of having emergency savings. Columns (1)-(3) show results for unsmoothed entropy based on 9 categories, 48 categories, and merchant names, respectively. Columns (4)-(6) 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. Confidence intervals are shown in brackets.

Results for unsmoothed entropy suggest that a one unit increase in entropy is associated with an increase in the probability of a user making at least one transfer into their savings accounts of between 1.5 and 2.7 percentage points – an effect up to two times larger than that of a £1000 increase in monthly income.

Conversely, the effect for unsmooth entropy is smaller in magnitude but runs in the reverse direction: a one-unit increase in the smoothed entropy score is associated with a reduction in the probability of transferring money into savings account of between 0.4 and 1.6 percentage points – an effect that, in absolute magnitude, is about equal to that of a £1000 increase in monthly income.

Overall, then, 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.

But how can we account for the opposite sign for smoothed and unsmoothed entropy scores? We don't have the answer to this yet...

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

Model:	(1)	(2)	(3)	(4)	(5)	(6)
<i>Variables</i>						
Entropy (9 cats)	0.016*** [0.013; 0.019]					
Entropy (48 cats)		0.029*** [0.025; 0.033]				
Entropy (merchant)			0.032*** [0.029; 0.036]			
Entropy (9 cats, smooth)				-0.008*** [-0.010; -0.006]		
Entropy (48 cats, smooth)					-0.023*** [-0.025; -0.020]	
Entropy (merchant, smooth)						-0.019*** [-0.021; -0.016]
Month spend	0.009*** [0.009; 0.010]	0.009*** [0.008; 0.009]	0.008*** [0.008; 0.009]	0.009*** [0.009; 0.010]	0.008*** [0.007; 0.009]	0.007*** [0.007; 0.008]
Month income	0.012*** [0.011; 0.013]	0.012*** [0.011; 0.013]	0.012*** [0.011; 0.013]	0.012*** [0.011; 0.013]	0.011*** [0.011; 0.012]	0.011*** [0.010; 0.012]
Has income in month	0.086*** [0.077; 0.094]	0.084*** [0.075; 0.092]	0.083*** [0.074; 0.091]	0.087*** [0.079; 0.096]	0.085*** [0.076; 0.093]	0.086*** [0.078; 0.095]
Income variability	0.001* [-0.000; 0.001]	0.001* [-0.000; 0.001]	0.001* [-0.000; 0.001]	0.001* [-0.000; 0.001]	0.001* [-0.000; 0.001]	0.000 [-0.000; 0.001]
<i>Fixed-effects</i>						
User	Yes	Yes	Yes	Yes	Yes	Yes
Year-month	Yes	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>						
Observations	1,043,727	1,043,727	1,043,416	1,043,727	1,043,727	1,043,416
R ²	0.45368	0.45395	0.45410	0.45363	0.45415	0.45410
Within R ²	0.00719	0.00768	0.00807	0.00709	0.00805	0.00808

Clustered (User) co-variance matrix, 95% confidence intervals in brackets

Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

5.1 Effect of financial resilience

If we think that it is scarcity that causes the relationship between unpredictable spending behaviour and savings, then we would expect the effect to be strongest for people with the lowest incomes, since they are more prone to face financial shocks they find difficult to meet. The results below support such an interpretation: the effect of both unsmoothed and smoothed entropy on the probability of making a savings transaction is largest in magnitude for users in the first income quintile.

In these regressions, we do not control for whether a user has income in in a given month, since for all but the group in the first income quintile, this will always be the case.

Table 4: Effect of entropy on P(has savings) by income quintile

Model:	(1)	(2)	(3)	(4)	(5)
Income quintile:	1	2	3	4	5
<i>Variables</i>					
Entropy (48 cats)	0.033*** [0.028; 0.039]	0.019*** [0.013; 0.025]	0.017*** [0.011; 0.023]	0.016*** [0.010; 0.022]	0.018*** [0.012; 0.025]
Month spend	0.010*** [0.009; 0.011]	0.007*** [0.006; 0.008]	0.008*** [0.006; 0.009]	0.007*** [0.006; 0.008]	0.006*** [0.005; 0.007]
Month income	0.015*** [0.012; 0.019]	0.010*** [0.004; 0.017]	0.012*** [0.005; 0.019]	0.013*** [0.008; 0.019]	0.008*** [0.007; 0.010]
Income variability	-0.000 [-0.001; 0.001]	0.001 [-0.000; 0.002]	0.002*** [0.001; 0.004]	0.001** [0.000; 0.002]	-0.000 [-0.001; 0.001]
<i>Fixed-effects</i>					
User	Yes	Yes	Yes	Yes	Yes
Year-month	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>					
Observations	223,462	215,151	208,417	201,319	195,378
R ²	0.59138	0.58949	0.58354	0.58212	0.57162
Within R ²	0.00535	0.00180	0.00211	0.00191	0.00271

Clustered (User) co-variance matrix, 95% confidence intervals in brackets

Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 5: Effect of entropy on P(has savings) by income quintile

Model:	(1)	(2)	(3)	(4)	(5)
Income quintile:	1	2	3	4	5
<i>Variables</i>					
Entropy (48 cats, smooth)	-0.026*** [-0.030; -0.023]	-0.020*** [-0.024; -0.016]	-0.017*** [-0.021; -0.013]	-0.015*** [-0.019; -0.011]	-0.017*** [-0.021; -0.012]
Month spend	0.009*** [0.008; 0.011]	0.006*** [0.005; 0.007]	0.007*** [0.006; 0.008]	0.007*** [0.005; 0.008]	0.006*** [0.005; 0.007]
Month income	0.015*** [0.012; 0.019]	0.010*** [0.004; 0.017]	0.012*** [0.004; 0.019]	0.013*** [0.007; 0.019]	0.008*** [0.007; 0.010]
Income variability	-0.000 [-0.001; 0.001]	0.001 [-0.000; 0.002]	0.002*** [0.001; 0.004]	0.001** [0.000; 0.002]	-0.000 [-0.001; 0.001]
<i>Fixed-effects</i>					
User	Yes	Yes	Yes	Yes	Yes
Year-month	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>					
Observations	223,462	215,151	208,417	201,319	195,378
R ²	0.59152	0.58970	0.58369	0.58223	0.57175
Within R ²	0.00570	0.00230	0.00246	0.00216	0.00302

Clustered (User) co-variance matrix, 95% confidence intervals in brackets
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 6: Effect of entropy on P(has savings) by income variability quintile

Model:	(1)	(2)	(3)	(4)	(5)
Income variability quintile:	1	2	3	4	5
<i>Variables</i>					
Entropy (48 cats)	0.020*** [0.014; 0.026]	0.014*** [0.008; 0.020]	0.022*** [0.016; 0.028]	0.028*** [0.022; 0.034]	0.021*** [0.015; 0.027]
Month spend	0.008*** [0.007; 0.009]	0.008*** [0.007; 0.009]	0.007*** [0.006; 0.008]	0.007*** [0.006; 0.008]	0.007*** [0.006; 0.008]
Month income	0.014*** [0.011; 0.017]	0.010*** [0.008; 0.012]	0.011*** [0.010; 0.013]	0.010*** [0.009; 0.012]	0.010*** [0.009; 0.011]
Has income in month	0.092*** [0.073; 0.111]	0.076*** [0.056; 0.095]	0.064*** [0.047; 0.081]	0.055*** [0.040; 0.070]	0.067*** [0.053; 0.081]
Income variability	-0.001 [-0.003; 0.001]	0.006** [0.001; 0.011]	-0.000 [-0.004; 0.003]	0.000 [-0.001; 0.002]	-0.004** [-0.007; -0.001]
<i>Fixed-effects</i>					
User	Yes	Yes	Yes	Yes	Yes
Year-month	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>					
Observations	223,462	215,151	208,417	201,319	195,378
R ²	0.61166	0.59951	0.59062	0.59379	0.63519
Within R ²	0.00558	0.00398	0.00484	0.00579	0.00781

Clustered (User) co-variance matrix, 95% confidence intervals in brackets
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 7: Effect of entropy on P(has savings) by income variability quintile

Model:	(1)	(2)	(3)	(4)	(5)
Income variability quintile:	1	2	3	4	5
<i>Variables</i>					
Entropy (48 cats, smooth)	-0.019*** [-0.024; -0.015]	-0.020*** [-0.024; -0.016]	-0.017*** [-0.021; -0.013]	-0.022*** [-0.026; -0.018]	-0.024*** [-0.028; -0.019]
Month spend	0.007*** [0.006; 0.008]	0.007*** [0.006; 0.008]	0.006*** [0.005; 0.008]	0.006*** [0.005; 0.007]	0.006*** [0.005; 0.007]
Month income	0.014*** [0.011; 0.017]	0.010*** [0.008; 0.012]	0.011*** [0.009; 0.013]	0.010*** [0.009; 0.012]	0.010*** [0.009; 0.011]
Has income in month	0.093*** [0.074; 0.112]	0.075*** [0.055; 0.094]	0.064*** [0.047; 0.081]	0.056*** [0.040; 0.071]	0.066*** [0.052; 0.081]
Income variability	-0.001 [-0.003; 0.001]	0.005** [0.001; 0.010]	-0.000 [-0.004; 0.003]	0.000 [-0.001; 0.002]	-0.004** [-0.006; -0.001]
<i>Fixed-effects</i>					
User	Yes	Yes	Yes	Yes	Yes
Year-month	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>					
Observations	223,462	215,151	208,417	201,319	195,378
R ²	0.61181	0.59979	0.59070	0.59394	0.63550
Within R ²	0.00597	0.00466	0.00502	0.00614	0.00864

Clustered (User) co-variance matrix, 95% confidence intervals in brackets
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

6 Discussion

References

- Benhabib, Jess and Alberto Bisin (2005). “Modeling internal commitment mechanisms and self-control: A neuroeconomics approach to consumption–saving decisions”. In: *Games and economic Behavior* 52.2, pp. 460–492.
- 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*.
- CAN, Commonwealth Bank of Australia (2019). “Improving the Financial Wellbeing of Australians”. Tech. rep. URL: https://www.commbank.com.au/content/dam/caas/newsroom/docs/CWM0375_Financial%20Wellbeing%20Report_v4.pdf.
- 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 (2009). “Reinforcement learning and savings behavior”. In: *The Journal of finance* 64.6, pp. 2515–2534.
- Colby, Helen and Gretchen B Chapman (2013). “Savings, subgoals, and reference points”. In: Eagle, Nathan, Michael Macy, and Rob Claxton (2010). “Network diversity and economic development”. In: *Science* 328.5981, pp. 1029–1031.
- Fudenberg, Drew and David K Levine (2006). “A dual-self model of impulse control”. In: *American economic review* 96.5, pp. 1449–1476.
- 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.
- 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.
- Gul, Faruk and Wolfgang Pesendorfer (2001). “Temptation and self-control”. In: *Econometrica* 69.6, pp. 1403–1435.
- 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.
- 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.
- Laibson, David (1997). “Golden eggs and hyperbolic discounting”. In: *The Quarterly Journal of Economics* 112.2, pp. 443–478.

- Laibson, David and Keith Marzilli-Ericson (2019). “Intertemporal choice”. In: *Handbook of Behavioral Economics* 2.
- Loewenstein, George and Ted O’Donoghue (2004). “Animal spirits: Affective and deliberative processes in economic behavior”. In: *Available at SSRN* 539843.
- Lunt, Peter K and Sonia M Livingstone (1991). “Psychological, social and economic determinants of saving: Comparing recurrent and total savings”. In: *Journal of economic Psychology* 12.4, pp. 621–641.
- 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.
- Mittal, Chiraag, Vladas Griskevicius, Jeffrey 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](https://doi.org/10.31234/osf.io/qabgm). URL: psyarxiv.com/qabgm/.
- Oaten, Megan and Ken Cheng (2007). “Improvements in self-control from financial monitoring”. In: *Journal of Economic Psychology* 28.4, pp. 487–501.
- Philipps, Jo, Annick Kuipers, and Will Sandbrook (2021). “Supporting emergency savings: early learnings of the employee experience of workplace sidecar savings”. Tech. rep.
- Sabat, Jorge and Emily Gallagher (2019). “Rules of thumb in household savings decisions: Estimation using threshold regression”. In: *Available at SSRN* 3455696.
- 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:
- Thaler, Richard H and Hersh M Shefrin (1981). “An economic theory of self-control”. In: *Journal of political Economy* 89.2, pp. 392–406.
- 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 Understanding entropy

To see how we can interpret entropy as the predictability of a user’s spending 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 its 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 desirable characteristics of an information function), we can use the log of the expression. Hence, the information of event E , often called *Shannon information*, *self-information*, or just *information*, is defined as:

$$I(E) = \log\left(\frac{1}{p(E)}\right) = -\log(p(E)). \quad (5)$$

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 calculated as:

$$H(X) = -\sum_x p(x) \times \log(p(x)) = \sum_x p(x) I(x) = \mathbb{E}I(x). \quad (6)$$

For a single event, the key intuition 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 Additional results

B.1 Endogeneity

As discussed in Section ??, one concern one might have about our results in Section ?? is reverse causality: transferring money into savings accounts might change the distribution of spend categories and thus change entropy. As noted previously, this is not a major concern because of the way we calculate savings and spend profiles: savings are calculated as the sum of inflows into savings accounts, while spend profiles are based on the classification of spend transactions in current accounts, and transfers from current to savings accounts are labelled as such and not treated as spend transactions.

However, because transaction labelling is imperfect, it is possible that some transfers are misclassified as spends and included in the calculation of entropy scores. Two ways to deal with this is to use lagged entropy scores and to explicitly control for the number of non-zero spend categories used to calculate entropy scores. Here, we show that our results remain qualitatively unchanged with both of these approaches.

Table 9 presents results similar to the main results in the main text, but using entropy lagged

by one year-month period as the independent variable of interest, while Table 8 adds the number of non-zero spend categories as an additional control.

Table 8: Effect of entropy on P(has savings)

Model:	(1)	(2)	(3)	(4)	(5)	(6)
<i>Variables</i>						
Entropy (9 cats)	-0.006*** [-0.010; -0.002]					
Entropy (48 cats)		-0.007*** [-0.011; -0.002]				
Entropy (merchant)			-0.004 [-0.009; 0.002]			
Entropy (9 cats, smooth)				-0.012*** [-0.015; -0.010]		
Entropy (48 cats, smooth)					-0.017*** [-0.020; -0.014]	
Entropy (merchant, smooth)						-0.011*** [-0.014; -0.009]
Month spend	0.008*** [0.008; 0.009]	0.007*** [0.006; 0.007]	0.007*** [0.006; 0.007]	0.008*** [0.007; 0.009]	0.006*** [0.005; 0.007]	0.006*** [0.005; 0.006]
Month income	0.012*** [0.011; 0.013]	0.011*** [0.010; 0.012]	0.011*** [0.010; 0.012]	0.012*** [0.011; 0.013]	0.011*** [0.010; 0.012]	0.011*** [0.010; 0.012]
Has income in month	0.082*** [0.073; 0.090]	0.079*** [0.071; 0.088]	0.081*** [0.072; 0.090]	0.081*** [0.072; 0.089]	0.078*** [0.069; 0.086]	0.081*** [0.073; 0.090]
Income variability	0.001* [-0.000; 0.001]	0.000* [-0.000; 0.001]	0.001* [-0.000; 0.001]	0.001* [-0.000; 0.001]	0.000* [-0.000; 0.001]	0.000 [-0.000; 0.001]
Unique categories	0.016*** [0.015; 0.018]	0.007*** [0.007; 0.008]	0.003*** [0.003; 0.004]	0.016*** [0.015; 0.018]	0.006*** [0.005; 0.007]	0.003*** [0.002; 0.003]
<i>Fixed-effects</i>						
User	Yes	Yes	Yes	Yes	Yes	Yes
Year-month	Yes	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>						
Observations	1,043,727	1,043,727	1,043,416	1,043,727	1,043,727	1,043,416
R ²	0.45421	0.45467	0.45460	0.45441	0.45499	0.45479
Within R ²	0.00815	0.00898	0.00898	0.00851	0.00958	0.00933

Clustered (User) co-variance matrix, 95% confidence intervals in brackets
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

B.2 Further results by financial resilience

The Tables below show results presented in Section 5.1 for alternative entropy variables.

Table 9: Effect of entropy on P(has savings)

Model:	(1)	(2)	(3)	(4)	(5)	(6)
<i>Variables</i>						
Entropy lag (9 cats)	0.015*** [0.012; 0.019]					
Entropy lag (48 cats)		0.024*** [0.021; 0.028]				
Entropy lag (merchant)			0.027*** [0.023; 0.030]			
Entropy lag (9 cats, smooth)				-0.004*** [-0.006; -0.002]		
Entropy lag (48 cats, smooth)					-0.016*** [-0.018; -0.013]	
Entropy lag (merchant, smooth)						-0.013*** [-0.016; -0.011]
Month spend	0.009*** [0.009; 0.010]	0.009*** [0.008; 0.010]	0.009*** [0.008; 0.010]	0.009*** [0.009; 0.010]	0.009*** [0.008; 0.009]	0.008*** [0.008; 0.009]
Month income	0.012*** [0.011; 0.013]	0.012*** [0.011; 0.013]	0.012*** [0.011; 0.013]	0.012*** [0.011; 0.013]	0.012*** [0.011; 0.013]	0.011*** [0.010; 0.012]
Has income in month	0.083*** [0.074; 0.092]	0.082*** [0.073; 0.091]	0.082*** [0.073; 0.091]	0.084*** [0.075; 0.093]	0.083*** [0.074; 0.092]	0.084*** [0.075; 0.092]
Income variability	0.001* [-0.000; 0.001]	0.001* [-0.000; 0.001]	0.001* [-0.000; 0.001]	0.001* [-0.000; 0.001]	0.001* [-0.000; 0.001]	0.001* [-0.000; 0.001]
<i>Fixed-effects</i>						
User	Yes	Yes	Yes	Yes	Yes	Yes
Year-month	Yes	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>						
Observations	1,009,843	1,009,843	1,009,554	1,009,843	1,009,843	1,009,554
R ²	0.45785	0.45802	0.45809	0.45775	0.45803	0.45800
Within R ²	0.00691	0.00722	0.00748	0.00671	0.00724	0.00731

Clustered (User) co-variance matrix, 95% confidence intervals in brackets
Signif. Codes: ***, 0.01, **, 0.05, *, 0.1

Table 10: Effect of entropy on P(has savings) by income quintile

Model:	(1)	(2)	(3)	(4)	(5)
Income quintile:	1	2	3	4	5
<i>Variables</i>					
Entropy (9 cats)	0.022*** [0.017; 0.026]	0.006** [0.001; 0.011]	0.009*** [0.004; 0.015]	0.007** [0.002; 0.013]	0.006** [0.000; 0.012]
Month spend	0.011*** [0.010; 0.012]	0.007*** [0.006; 0.009]	0.008*** [0.007; 0.009]	0.007*** [0.006; 0.008]	0.007*** [0.005; 0.008]
Month income	0.016*** [0.013; 0.019]	0.011*** [0.004; 0.017]	0.012*** [0.005; 0.020]	0.013*** [0.008; 0.019]	0.008*** [0.007; 0.010]
Income variability	0.000 [-0.001; 0.001]	0.001 [-0.000; 0.002]	0.002*** [0.001; 0.004]	0.001** [0.000; 0.002]	-0.000 [-0.001; 0.001]
<i>Fixed-effects</i>					
User	Yes	Yes	Yes	Yes	Yes
Year-month	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>					
Observations	223,462	215,151	208,417	201,319	195,378
R ²	0.59111	0.58937	0.58347	0.58205	0.57151
Within R ²	0.00469	0.00151	0.00193	0.00175	0.00246

Clustered (User) co-variance matrix, 95% confidence intervals in brackets
Signif. Codes: ***, 0.01, **, 0.05, *, 0.1

Table 11: Effect of entropy on P(has savings) by income quintile

Model:	(1)	(2)	(3)	(4)	(5)
Income quintile:	1	2	3	4	5
<i>Variables</i>					
Entropy (9 cats, smooth)	-0.008*** [-0.012; -0.005]	-0.009*** [-0.012; -0.006]	-0.006*** [-0.009; -0.002]	-0.006*** [-0.009; -0.002]	-0.008*** [-0.012; -0.005]
Month spend	0.011*** [0.010; 0.012]	0.007*** [0.006; 0.008]	0.008*** [0.007; 0.009]	0.007*** [0.006; 0.008]	0.006*** [0.005; 0.007]
Month income	0.017*** [0.013; 0.020]	0.011*** [0.004; 0.017]	0.012*** [0.005; 0.020]	0.013*** [0.008; 0.019]	0.008*** [0.007; 0.010]
Income variability	0.000 [-0.001; 0.001]	0.001 [-0.000; 0.002]	0.002*** [0.001; 0.004]	0.001** [0.000; 0.002]	-0.000 [-0.001; 0.001]
<i>Fixed-effects</i>					
User	Yes	Yes	Yes	Yes	Yes
Year-month	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>					
Observations	223,462	215,151	208,417	201,319	195,378
R ²	0.59095	0.58944	0.58347	0.58207	0.57158
Within R ²	0.00431	0.00168	0.00193	0.00179	0.00260

Clustered (User) co-variance matrix, 95% confidence intervals in brackets
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 12: Effect of entropy on P(has savings) by income variability quintile

Model:	(1)	(2)	(3)	(4)	(5)
Income variability quintile:	1	2	3	4	5
<i>Variables</i>					
Entropy (9 cats)	0.011*** [0.006; 0.016]	0.005* [-0.001; 0.010]	0.012*** [0.006; 0.018]	0.016*** [0.010; 0.022]	0.011*** [0.006; 0.017]
Month spend	0.008*** [0.007; 0.009]	0.008*** [0.007; 0.009]	0.007*** [0.006; 0.008]	0.007*** [0.006; 0.009]	0.007*** [0.006; 0.008]
Month income	0.015*** [0.012; 0.017]	0.010*** [0.008; 0.012]	0.011*** [0.010; 0.013]	0.010*** [0.009; 0.012]	0.010*** [0.009; 0.011]
Has income in month	0.094*** [0.075; 0.113]	0.077*** [0.057; 0.096]	0.065*** [0.048; 0.082]	0.057*** [0.041; 0.072]	0.068*** [0.054; 0.083]
Income variability	-0.001 [-0.003; 0.002]	0.006** [0.001; 0.011]	-0.000 [-0.004; 0.003]	0.001 [-0.001; 0.002]	-0.004*** [-0.007; -0.001]
<i>Fixed-effects</i>					
User	Yes	Yes	Yes	Yes	Yes
Year-month	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>					
Observations	223,462	215,151	208,417	201,319	195,378
R ²	0.61156	0.59945	0.59051	0.59361	0.63509
Within R ²	0.00532	0.00383	0.00456	0.00534	0.00753

Clustered (User) co-variance matrix, 95% confidence intervals in brackets
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 13: Effect of entropy on P(has savings) by income variability quintile

Model:	(1)	(2)	(3)	(4)	(5)
Income variability quintile:	1	2	3	4	5
<i>Variables</i>					
Entropy (9 cats, smooth)	-0.007*** [-0.011; -0.004]	-0.009*** [-0.013; -0.006]	-0.005*** [-0.009; -0.002]	-0.007*** [-0.010; -0.003]	-0.010*** [-0.014; -0.007]
Month spend	0.008*** [0.007; 0.009]	0.008*** [0.007; 0.009]	0.007*** [0.006; 0.008]	0.007*** [0.006; 0.009]	0.007*** [0.006; 0.008]
Month income	0.015*** [0.012; 0.018]	0.010*** [0.008; 0.012]	0.012*** [0.010; 0.013]	0.011*** [0.009; 0.012]	0.010*** [0.009; 0.011]
Has income in month	0.095*** [0.076; 0.115]	0.077*** [0.057; 0.096]	0.066*** [0.049; 0.083]	0.058*** [0.042; 0.073]	0.069*** [0.055; 0.083]
Income variability	-0.001 [-0.003; 0.001]	0.006** [0.001; 0.011]	-0.000 [-0.004; 0.003]	0.000 [-0.001; 0.002]	-0.004*** [-0.007; -0.001]
<i>Fixed-effects</i>					
User	Yes	Yes	Yes	Yes	Yes
Year-month	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>					
Observations	223,462	215,151	208,417	201,319	195,378
R ²	0.61156	0.59954	0.59048	0.59355	0.63514
Within R ²	0.00533	0.00404	0.00448	0.00519	0.00767

Clustered (User) co-variance matrix, 95% confidence intervals in brackets
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 14: Effect of entropy on P(has savings) by income quintile

Model:	(1)	(2)	(3)	(4)	(5)
Income quintile:	1	2	3	4	5
<i>Variables</i>					
Entropy (merchant)	0.032*** [0.027; 0.037]	0.027*** [0.022; 0.033]	0.024*** [0.019; 0.030]	0.017*** [0.011; 0.023]	0.021*** [0.015; 0.027]
Month spend	0.010*** [0.009; 0.011]	0.006*** [0.005; 0.008]	0.007*** [0.006; 0.008]	0.007*** [0.006; 0.008]	0.006*** [0.005; 0.007]
Month income	0.015*** [0.012; 0.019]	0.010*** [0.004; 0.017]	0.012*** [0.005; 0.019]	0.013*** [0.008; 0.019]	0.008*** [0.007; 0.010]
Income variability	-0.000 [-0.001; 0.001]	0.001 [-0.000; 0.002]	0.002*** [0.001; 0.004]	0.001** [0.000; 0.002]	-0.000 [-0.001; 0.001]
<i>Fixed-effects</i>					
User	Yes	Yes	Yes	Yes	Yes
Year-month	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>					
Observations	223,357	215,087	208,364	201,280	195,328
R ²	0.59137	0.58963	0.58366	0.58211	0.57163
Within R ²	0.00544	0.00229	0.00245	0.00198	0.00288

Clustered (User) co-variance matrix, 95% confidence intervals in brackets
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 15: Effect of entropy on P(has savings) by income quintile

Model:	(1)	(2)	(3)	(4)	(5)
Income quintile:	1	2	3	4	5
<i>Variables</i>					
Entropy (merchant, smooth)	-0.025*** [-0.029; -0.022]	-0.017*** [-0.021; -0.013]	-0.014*** [-0.018; -0.011]	-0.013*** [-0.018; -0.009]	-0.015*** [-0.019; -0.012]
Month spend	0.009*** [0.008; 0.010]	0.006*** [0.005; 0.007]	0.007*** [0.005; 0.008]	0.006*** [0.005; 0.007]	0.005*** [0.004; 0.006]
Month income	0.015*** [0.012; 0.019]	0.010*** [0.003; 0.017]	0.011*** [0.004; 0.018]	0.013*** [0.007; 0.019]	0.008*** [0.007; 0.010]
Income variability	-0.000 [-0.001; 0.001]	0.001 [-0.000; 0.002]	0.002*** [0.001; 0.004]	0.001** [0.000; 0.002]	-0.000 [-0.001; 0.001]
<i>Fixed-effects</i>					
User	Yes	Yes	Yes	Yes	Yes
Year-month	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>					
Observations	223,357	215,087	208,364	201,280	195,328
R ²	0.59153	0.58963	0.58368	0.58225	0.57180
Within R ²	0.00581	0.00230	0.00249	0.00232	0.00327

Clustered (User) co-variance matrix, 95% confidence intervals in brackets
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 16: Effect of entropy on P(has savings) by income variability quintile

Model:	(1)	(2)	(3)	(4)	(5)
Income variability quintile:	1	2	3	4	5
<i>Variables</i>					
Entropy (merchant)	0.025*** [0.020; 0.031]	0.020*** [0.014; 0.025]	0.025*** [0.019; 0.031]	0.036*** [0.030; 0.042]	0.026*** [0.020; 0.031]
Month spend	0.007*** [0.006; 0.009]	0.008*** [0.006; 0.009]	0.007*** [0.005; 0.008]	0.006*** [0.005; 0.007]	0.007*** [0.005; 0.008]
Month income	0.014*** [0.011; 0.017]	0.010*** [0.008; 0.012]	0.011*** [0.010; 0.013]	0.010*** [0.009; 0.012]	0.010*** [0.009; 0.011]
Has income in month	0.091*** [0.072; 0.111]	0.075*** [0.055; 0.094]	0.063*** [0.046; 0.080]	0.054*** [0.039; 0.069]	0.066*** [0.052; 0.080]
Income variability	-0.001 [-0.003; 0.001]	0.006** [0.001; 0.010]	-0.000 [-0.004; 0.003]	0.000 [-0.001; 0.002]	-0.004** [-0.007; -0.001]
<i>Fixed-effects</i>					
User	Yes	Yes	Yes	Yes	Yes
Year-month	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>					
Observations	223,369	215,092	208,349	201,276	195,330
R ²	0.61172	0.59956	0.59064	0.59404	0.63529
Within R ²	0.00592	0.00422	0.00506	0.00644	0.00814

Clustered (User) co-variance matrix, 95% confidence intervals in brackets
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 17: Effect of entropy on P(has savings) by income variability quintile

Model:	(1)	(2)	(3)	(4)	(5)
Income variability quintile:	1	2	3	4	5
<i>Variables</i>					
Entropy (merchant, smooth)	-0.018*** [-0.021; -0.014]	-0.018*** [-0.022; -0.014]	-0.016*** [-0.019; -0.012]	-0.017*** [-0.020; -0.013]	-0.016*** [-0.021; -0.012]
Month spend	0.007*** [0.006; 0.008]	0.006*** [0.005; 0.008]	0.006*** [0.005; 0.007]	0.006*** [0.005; 0.007]	0.006*** [0.005; 0.007]
Month income	0.014*** [0.011; 0.017]	0.010*** [0.007; 0.012]	0.011*** [0.009; 0.013]	0.010*** [0.009; 0.011]	0.010*** [0.009; 0.011]
Has income in month	0.094*** [0.075; 0.113]	0.075*** [0.056; 0.095]	0.065*** [0.048; 0.082]	0.057*** [0.042; 0.072]	0.068*** [0.054; 0.082]
Income variability	-0.001 [-0.003; 0.001]	0.005** [0.000; 0.010]	-0.000 [-0.004; 0.003]	0.000 [-0.001; 0.002]	-0.004** [-0.006; -0.001]
<i>Fixed-effects</i>					
User	Yes	Yes	Yes	Yes	Yes
Year-month	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>					
Observations	223,369	215,092	208,349	201,276	195,330
R ²	0.61177	0.59980	0.59070	0.59387	0.63537
Within R ²	0.00607	0.00480	0.00520	0.00602	0.00836

Clustered (User) co-variance matrix, 95% confidence intervals in brackets
Signif. Codes: ***: 0.01, **: 0.05, *: 0.1