Spending profiles predict savings*

Fabian Gunzinger Neil Stewart

September 26, 2022

Contents

*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 individuals 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, resarch 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 individuals 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 individuals. Such data has only become available ver recently and have not, thus far, been used to investigate systematically how individuals spend their money.
- Research in psychology suggests that disorder is maladaptive and associated with

[1] MDB 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.

[2] Well-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 individuals 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.

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).

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 hore impulse purchases, and make fortetting 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 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.[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 tought 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 algorightms 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.

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

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

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.

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.

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 transactionas across different spend categories. To summarise these distributions, we calculate spending entropy, based on the formula proposed by Shannon (1948), who defines entropy

[5]While 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.

as $H = -\sum p_i log(p_i)$, which sums, for all possible events, the product of the probability of an event i occuring with the logarithm of that probability.[6] The base 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 makes a purchase in spending category c, and log the base 2 logarithm. 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.[7] For simpler interpretation of our regression coefficients below, we standardise entropy scores to have a mean of 0 and a standard deviation of 1.

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.[8] 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.[9] To calculate "smoothed" entropy scores,

[6]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 analagous.

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

[8]The precise mapping from MDB transaction tags into 9 and 48 categories is available on Github here and here, respectively.

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

we apply additive smoothing to calculate propabilities as

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

where the size of set $\mathcal{C}$, $|\mathcal{C}|$, is the number of unique spending categories. Hence, additive smoothing simply adds one to the numerator and the number of unqiue spending categories to the denominator of the unsmoothed probabilities. Because categories with a zero transaction count will have a numerator of 1, the sum in Equation 1 will be taken over all categories.

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 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), 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:

$$I(E) = log\left(\frac{1}{p(E)}\right) = -log(p(E)). \tag{4}$$

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:

$$H(X) = -\sum_x p(x) \times log(p(x)) = \sum_x p(x)I(x) = \mathbb{E}I(x). \tag{5}$$

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.

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

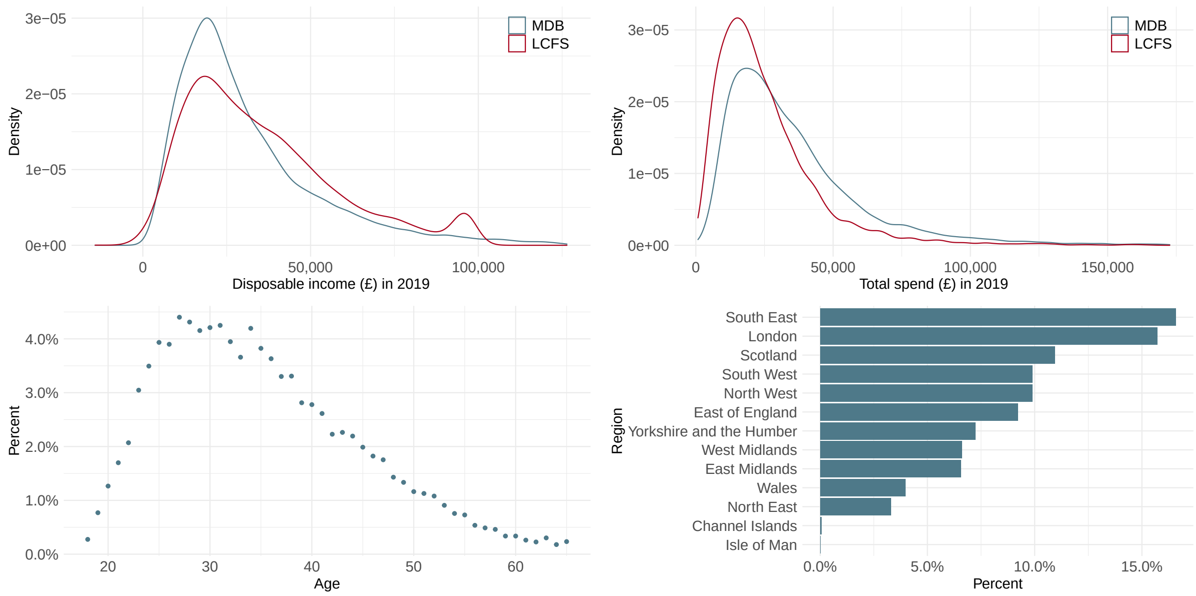
Table 2 provides summary statistics.

Table 2: Summary statistics

Statistic	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Month income	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
Has savings	0.50	0.50	0	0	1	1	1
Month spend	2.90	2.50	0.20	1.37	2.20	3.49	16.05
Age	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 (Merchants)	26.78	9.35	2	20	26	33	85

Figure 1 shows sample characteristics.

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.

2.6 Estimation

We estimate models of the form:

$$y_{i,t} = \alpha_i + \lambda_t + \beta H_{i,t} + x'_{i,t}\delta + \epsilon_{i,t}, \tag{6}$$

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_{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. In Appendix B.1, we provide robustness checks using lagged entropy scores, which produces very similar results.

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) 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 brakets.

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 1.1 and 2.3 percentage points – an effect up to two times larger than that of a £1000 increase in monthly income. Conversely, the effect for unsmooth entropy tends to be somewhat 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 1.7 percentage points.

As discussed in Section 2.6, these results are not a results of 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

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	0.010 [0.009; 0.011]	0.009 [0.008; 0.011]	0.010 [0.008; 0.011]	0.010 [0.009; 0.011]	0.009 [0.008; 0.010]	0.008 [0.006; 0.009]
Month income	0.012 [0.011; 0.014]	0.012 [0.011; 0.014]	0.012 [0.011; 0.014]	0.012 [0.011; 0.014]	0.012 [0.010; 0.014]	0.012 [0.010; 0.013]
Has income in month	0.086 [0.070; 0.102]	0.085 [0.069; 0.101]	0.086 [0.069; 0.102]	0.087 [0.071; 0.103]	0.086 [0.070; 0.102]	0.086 [0.070; 0.102]
Income variability	0.001 [0.000; 0.002]	0.001 [0.000; 0.002]	0.001 [0.000; 0.002]	0.001 [0.000; 0.002]	0.001 [0.000; 0.002]	0.001 [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 6. 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.

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 unsmoohed entropy but high smoothed entropy or 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 $\mathcal{C}$ is the set of all spending categories, and 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 $\mathcal{C} = \mathcal{C}^0 \cup \mathcal{C}^+$. Remember also that f_c is the frequency count of spending category c – the number of transactions user i made in category c in period t – and F is the total number of transactions made by that user in that period. Then, using our definitions of unsmoothed and smoothed probabilities, we can write unsmoothed entropy as

$$H = -\sum_{c \in \mathcal{C}^+} \left(\frac{f_c}{F}\right) log \left(\frac{f_c}{F}\right), \tag{7}$$

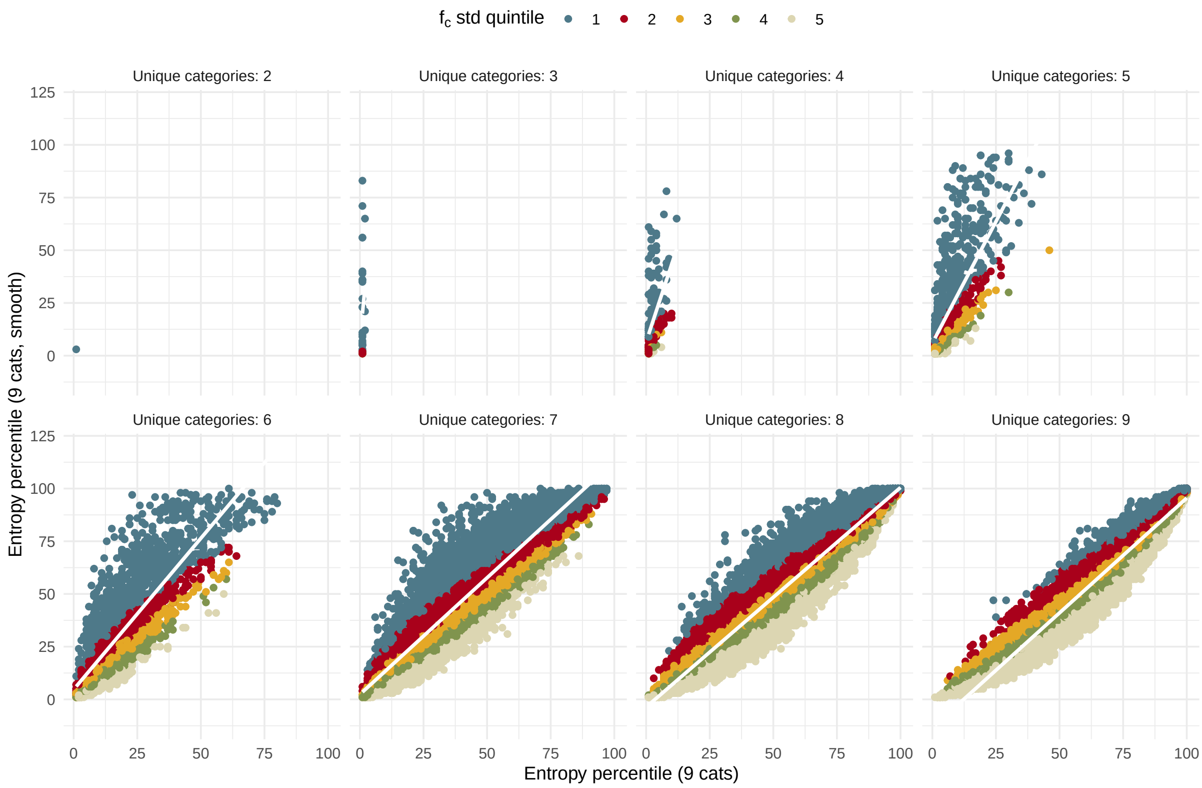
and smoothed entropy as:

$$H^s = -\sum_{c \in \mathcal{C}^+} \left(\frac{f_c + 1}{F + |\mathcal{C}|}\right) log \left(\frac{f_c + 1}{F + |\mathcal{C}|}\right) - |\mathcal{C}^0| \left(\frac{1}{F + |\mathcal{C}|}\right) log \left(\frac{1}{F + |\mathcal{C}|}\right), \tag{8}$$

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 a certain 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 the additively smoothed probabilities of categories with zero counts, which reduces to a constant term that is multiplied by the number of categories with a zero count.

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 7, and partitions the categories into either contributing to the sum on the left hand side of Equation 8 or to the constant term on the right hand side. The latter is the case since, for an exogenously fixed $|\mathcal{C}|$, $|\mathcal{C}^0| = |\mathcal{C} \backslash \mathcal{C}^+|$ – for a given number of total categories, the number of categories with a zero frequency count is the difference between the total number of categories and the number of categories with a positive frequency count. The second component is the variation of the frequency counts, f_c, which will determine the variation in the probabilities of a spend accurring in a given category. The third component is the total number of transactions (F). 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.

Figure 2: 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 frequency count standard deviation quintiles.

One starting point for thinking about why unsmoothed and smoothed entropy differ is thus to think about how these three components effect the two scores differently. From equations 7 and 8 we can see that the first part of smoothed entropy that sums over all spending categories with positive frequency 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.

Next, 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 counts frequency counts (f_c) are close to zero and (ii) if there is little variation in the counts.

Figure 2 visualises this intuition for our 9-category based entropy variable: it shows scatterplots 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.[10] 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 frequency variation that experience little or even a negative rank difference. Hence, while counts variation is some help in identifying cases with high entropy rank differences, it does not do so perfectly.

Table 4: Entorpy on components

	Entropy (9 cats) (1)	Entropy (9 cats, smooth) (2)	Entropy (48 cats) (3)	Entropy (48 cats, smooth) (4)
Unique categories ($\lvert\mathcal{C}^+\rvert$)	0.029	0.011	0.073	0.009
	[0.028; 0.030]	[0.010; 0.012]	[0.073; 0.074]	[0.008; 0.010]
Category counts std.	-0.251	-0.494	-0.276	-0.431
	[-0.257; -0.245]	[-0.505; -0.483]	[-0.283; -0.270]	[-0.441; -0.421]
Number of spend txns (F)	0.013	0.021	0.013	0.010
	[0.012; 0.013]	[0.021; 0.022]	[0.013; 0.013]	[0.009; 0.010]
Observations	389,790	389,790	389,790	389,790
R^2	0.73531	0.75081	0.93219	0.93063
Within R^2	0.37140	0.45346	0.80362	0.82674
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.

A different way to investigate whether the sign flip can be explained by the three components of the entropy formula is to test whether entropy captures anything about users' spending distribution that is predictive of savings behaviour beyond the information already contained in its three components. If this is were not the case, then we should be able to explain the sign flip solely based on the components.

[10]Figure 4 in Appendix B a similar plot with colouring based on the quintile of the total number of transactions. The result is very similar, which is why we do not show it here.

Table 5: 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.

As a first step, Table 4 shows results from regressing the unsmoothed and smoothed versions of both the 9-category and 48-category based entropy variables on the three components. The components explain about 75% of the variation in the 9-category-based and 95% of the 48-category-based variable. The directions of the effects are consistent across the two entropy variables as well as accross smoothed and unsmoothed versions and are as expected: an increase in the number of categories with a positive frequency count increases entropy (more elements are summed over in Equation 7 and more elements are part of the first part of the expression in Equation 8), a higher variation in spend category frequency counts reduces entropy (the uniform distribution is the maximum entropy distribution), and the total number of spending transactions increases entropy.[11]

Next, we look at 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 3 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 entorpy – while the width of the confidence intervals about double in both cases, reflecting the strong collinearity amont 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 Discussion

There 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.

[11] The direction of the relationship between the number of total spend transaction and entropy is less obvious than that of the other two components. Taking the derivative of Equation 7 with respect to F shows that the relationship is guaranteed to be positive if the frequency count of each category is less than $F/2$.

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, 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.

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 Interpreting entropy

B Additional results

B.1 Endogeneity

As discussed in Section 2.6, one concern one might have about our results in Section 3 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.

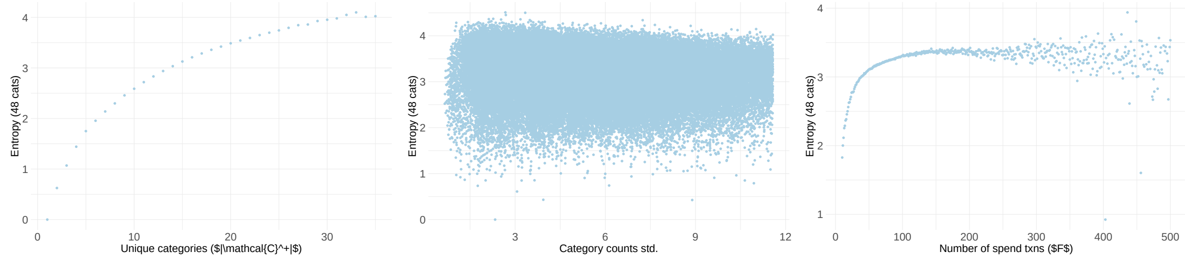
However, because transaction labelling is imperfect, it is possible that some transfers are misclassified as spends and included in the calculation of entropy scores. One way to deal with this is to lag entropy scores by one period. Table 6 presents results similar to the main results in the main text, but using entropy lagged by one year-month period as the independent variable of interest. We can see that the results are very similar to thos presented above.

Table 6: Effect of lagged entropy on P(savings transactions)

	(1)	(2)	(3)	(4)	(5)	(6)
Entropy lag (9 cats)	0.013 [0.007; 0.019]					
Entropy lag (48 cats)		0.020 [0.013; 0.026]				
Entropy lag (merchant)			0.016 [0.010; 0.022]			
Entropy lag (9 cats, smooth)				0.001 [-0.003; 0.005]		
Entropy lag (48 cats, smooth)					-0.009 [-0.013; -0.005]	
Entropy lag (merchant, smooth)						-0.012 [-0.015; -0.008]
Month spend	0.010 [0.009; 0.011]	0.010 [0.008; 0.011]	0.010 [0.008; 0.011]	0.010 [0.009; 0.011]	0.009 [0.008; 0.011]	0.009 [0.008; 0.010]
Month income	0.012 [0.010; 0.014]	0.012 [0.010; 0.014]	0.012 [0.010; 0.014]	0.012 [0.011; 0.014]	0.012 [0.010; 0.014]	0.012 [0.010; 0.013]
Has income in month	0.083 [0.066; 0.099]	0.082 [0.066; 0.099]	0.083 [0.066; 0.099]	0.083 [0.067; 0.100]	0.083 [0.067; 0.100]	0.084 [0.067; 0.100]
Income variability	0.001 [0.000; 0.002]	0.001 [0.000; 0.002]	0.001 [0.000; 0.002]	0.001 [0.000; 0.002]	0.001 [0.000; 0.002]	0.001 [0.000; 0.002]
Observations	375,251	375,251	375,251	375,251	375,251	375,251
R^2	0.47281	0.47289	0.47285	0.47275	0.47282	0.47300
Within R^2	0.00595	0.00610	0.00603	0.00583	0.00597	0.00631
User fixed effects	✓	✓	✓	✓	✓	✓
Year-month fixed effects	✓	✓	✓	✓	✓	✓

Notes: Results from estimating Equation 6. 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. Entropy variables are lagged by one period. Terms in brackets denote the upper and lower ends of a 95% confidence interval. Standard errors are clustered at the user-level.

Figure 3: 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).

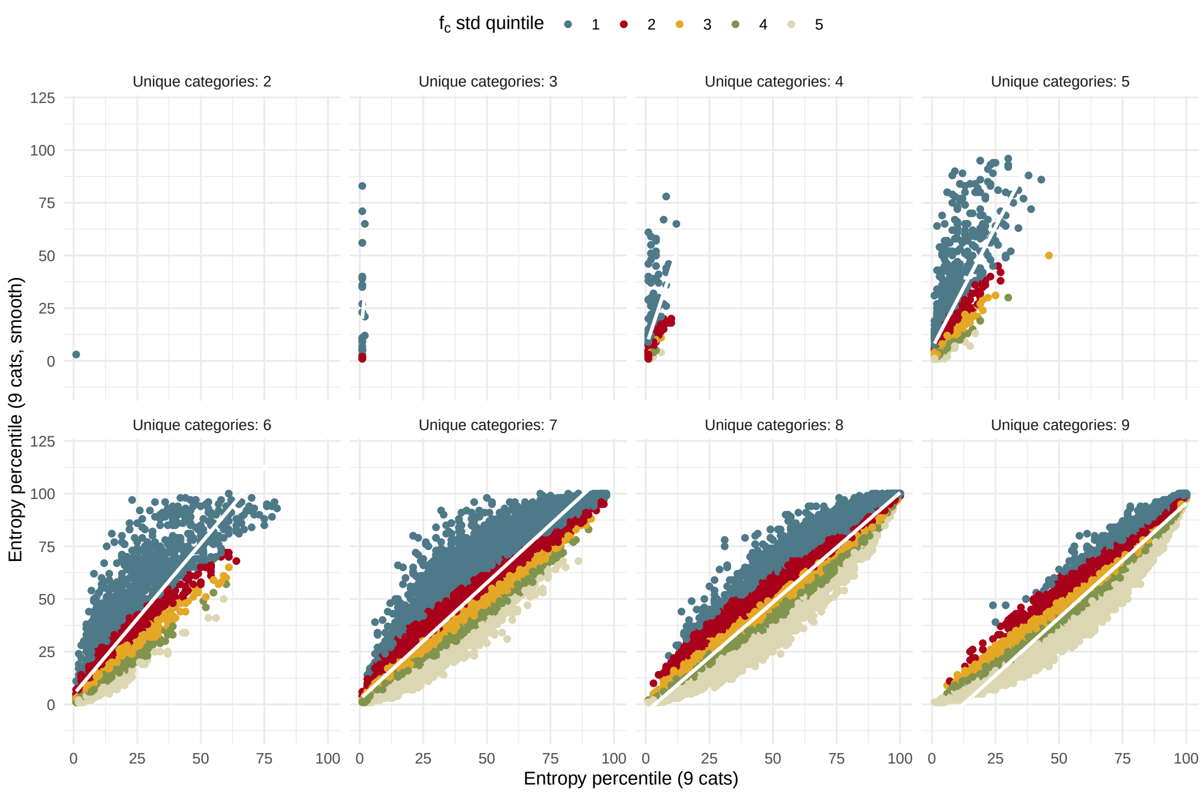
B.2 Entropy components

Figure 3 shows the empirical relationship with our 48-categories-based unsmoothed entropy variable and these three components.[12] We can see that for the values 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.

B.3 Effect of smothing on entropy (by number of spend quintile)

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

Figure 4: 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.