

# Entropy\*

Fabian Gunzinger  
Warwick Business School

Neil Stewart  
Warwick Business School

October 1, 2021

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Data</b>	<b>2</b>
2.1	Preprocessing . . . . .	2
2.2	Sample selection . . . . .	2

## 1 Introduction

Literature:

Muggleton et al. (2020) find that consumption entropy over categories correlates with financial distress.

Davenport et al. (2020) study the impact of COVID-19 on the spending and savings behaviour of MDB users.

Baker and Kueng (2021) summarises literature that uses mass financial transaction data to study household financial behaviour.

Becker (2017) finds that access to a fintech money management app increases first-time savings and savings account balances among 65,000 customers of a large European bank but that uptake is negatively correlated with financial sophistication.

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.

Paper:

Independent variable: entropy over categories and others

Outcome variables: first-time saving, average monthly savings

---

\*This research was supported by Economic and Social Research Council grant number ES/V004867/1. WBS ethics code: E-414-01-20.

## 2 Data

### 2.1 Preprocessing

Duplicate transactions

### 2.2 Sample selection

Table 1: Sample selection

	Users	Accounts	Transactions	Value (£M)
Raw sample.	283	1,281	681,755	125.1
At least 6 months of data.	246	1,201	674,665	123.7
At least one current account.	234	1,156	658,074	121.5
At least 5 monthly debits totalling GBP200.	161	707	457,545	82.4
Income payments in 2/3 of all observed months.	119	559	350,773	65.5
Yearly incomes between 5k and 100k.	51	194	133,350	16.1
No more than 10 active accounts in any year.	50	179	129,173	15.2
Debits of no more than 100k in any month.	50	179	129,173	15.2
Working-age.	44	157	112,417	12.6
Final sample.	44	157	112,417	12.6

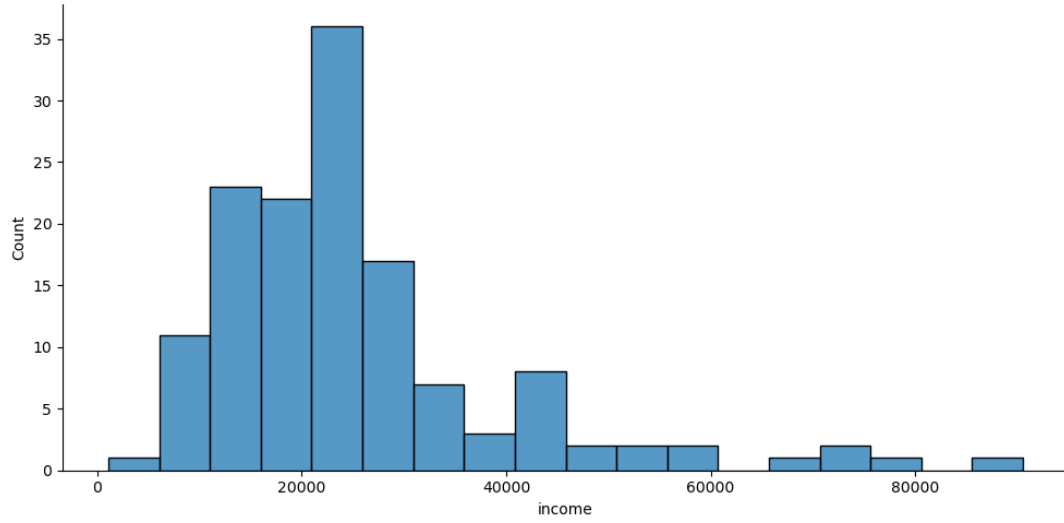
Table 2: Sample selection

	Users	Accounts	Transactions	Value (£M)
Raw sample.	2,730	13,272	6,950,912	1,229.6
At least 6 months of data.	2,432	12,589	6,896,903	1,219.5
At least one current account.	2,304	12,234	6,695,164	1,196.8
At least 5 monthly debits totalling GBP200.	1,518	7,915	4,906,252	897.3
Income payments in 2/3 of all observed months.	1,107	5,937	3,592,551	676.1
Yearly incomes between 5k and 100k.	561	2,936	1,786,631	280.0
No more than 10 active accounts in any year.	521	2,227	1,539,684	207.7
Debits of no more than 100k in any month.	497	2,096	1,445,382	177.5
Working-age.	430	1,811	1,301,806	156.8
Final sample.	430	1,811	1,301,806	156.8

Types of balances, from Becker (2017), which treats balance at end of each month as observations:

- Current account balance
- Debit balance (savings and current account balance)
- Pure savings (savings account balance only)
- Credit balance (loans and negative current account)
- Pure credit (loans only)
- Wealth held (debit - credit balance)

Figure 1: Distribution of user incomes



## References

- Baker, Scott R and Lorenz Kueng (2021). “Household Financial Transaction Data”. Tech. rep. National Bureau of Economic Research.
- Becker, G (2017). “Does fintech affect household saving behavior? findings from a natural field experiment”. Tech. rep. mimeo.
- Colby, Helen and Gretchen B Chapman (2013). “Savings, subgoals, and reference points”. In: Davenport, Alex, Robert Joyce, Imran Rasul, and Tom Waters (2020). “Spending and saving during the COVID-19 crisis: evidence from bank account data”. In: *Institute for Fiscal Studies, Briefing Note* 308.
- 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](https://psyarxiv.com/qabgm).