

Quality of Expenditure Data Collected with a Receipt Scanning App in a Probability Household Panel

Alexander Wenz University of Essex

Bureau of Labor Statistics October 3, 2018

An initiative by the Economic and Social Research Council, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Kantar Public

Acknowledgements

Funding



- Economic and Social Research Council (ESRC)
- National Centre for Research Methods (NCRM)
- ► In collaboration with Kantar Worldpanel

Team members

Annette Jäckle (Essex) Thomas Crossley (Essex) Jonathan Burton (Essex) Paul Fisher (Essex) Mike Brewer (Essex) Alexander Wenz (Essex) Brendan Read (Essex) Mick Couper (Michigan) Carli Lessof (Southampton) Joachim Winter (Munich) Cormac O'Dea (Yale) Alessandra Gaia (Milano)

Study aims



Collect high-quality data on monthly expenditure

 in the context of a general-purpose household panel survey: The UK Household Longitudinal Study (UKHLS)

Challenges

- Limited questionnaire space
- Maximize data quality
- Minimize respondent burden

Study aims



Approaches to measure expenditure

- Recall questions
- Expenditure diaries (paper/online)
- ► Data linkage (credit cards, loyalty cards, account aggregation)
- Receipt scanning app

Advantages of a receipt scanning app

- Collect objective data
- Capture more detailed data
- Lower respondent burden (?)

Background

Receipt data collection



- ► Market research: studies on shopping behavior
- ► Academic research: small-scale studies on food expenditure, paper receipts returned by post (French et al. 2008; Ransley et al. 2001)
- Development of smartphone-based receipt scanning
 - ► Food Acquisition and Purchase Survey: FoodAPS-2 (Yan et al. 2017)

Diary data collection

- Development of web-based expenditure diaries
 - Consumer Expenditure Survey: Gemini Project (Kopp 2016, Erhard & McBride 2017)

Study design



- Understanding Society Innovation Panel
- ► Stratified, clustered random sample of HHs in Great Britain
- ► Adults aged 16+ invited to download receipt scanning app
 - After annual household interview
 - ► Invitation by letter and email + reminders
 - Use app for one month
 - Own smartphone or tablet (iOS or Android)
 - Everyone invited, not only those with known access to mobile devices

Study design



- Incentives
 - £2 vs. £6 for app download
 - $\pounds 0.50$ per day app used
 - $\pounds 10$ bonus for using app for one month
 - ▶ £3 for debrief survey
 - ▶ Max. total: £30.50 (\approx \$40) / £34.50 (\approx \$45)
- ► Fieldwork: October-December 2016

App design – tasks

=	
	Submit Purchase or Nothing Bought Today
FAQs ?	Help and FAQs Heb with taking part
Understanding	Welcome to the Spending Study



- Scan shopping receipt
- Report spending in app
- Report no purchases today

Based on app developed by Kantar Worldpanel

App design - scan shopping receipt





App design - direct entry



Theorem and a contract oppay
Food and groceries
Clothes and footwear
Transport costs, e.g. petrol, car maintenance, public transport costs
Child costs, e.g. childcare, school equipment and fees
Home improvements and household goods, e.g. DIY, gardening, furniture, white goods or electrical goods
Health expenses, e.g. glasses, dental care, prescriptions, social care
Socialising and hobbies, e.g. going out (restaurants, pub, cinema, theatre, concert), gym or club membership, arts and crafts, children's activities
Other goods and services, e.g. books, magazines, DVDs, CDs, games, toys, beauty products, haircuts, manicures, massages
Holidays
Giving money or gifts to other people, e.g. money for children, gifts or money for relatives, donations to charity
Please select 1 or more options

Which of the following categories best describe what the

purchase was?

Participation

Low participation rate among general population

- 12.8% used app at least once (n = 270)
- Similar to other app-based studies

Participation bias: some respondents more likely to participate

- Women
- Frequent mobile device users
- Users of technology to manage finances
- Users of store loyalty cards
- Those generally cooparative with the survey

No evidence of bias on income- and spending-related indicators

Jäckle et al. (2017): Participation in a mobile app survey to collect expenditure data as part of a large-scale probability household panel: response rates and response biases, Understanding Society Working Paper, 2017-09

Participation

Little dropout over the month

 \blacktriangleright 81.5% of app users remain in the study for at least 29 days



Jäckle et al. (2017): Participation in a mobile app survey to collect expenditure data as part of a large-scale probability household panel: response rates and response biases, Understanding Society Working Paper, 2017-09

Other papers from this project

Willingness to do tasks on mobile device



- Willingness differs markedly between different types of tasks:
 e.g. more willing to do active than passive tasks
- Wenz et al. (*in press*): Willingness to use mobile technologies for data collection in a probability household panel, *Survey Research Methods* (Understanding Society Working Paper, 2017-10)

Respondent burden

- App users reported low subjective burden
- ► Objective burden (time per app use) does not predict dropout
- Read (2018): Respondent burden in a mobile app: evidence from a shopping receipt scanning study, Understanding Society Working Paper, 2018-04

Other papers from this project



Process quality

- How well do respondents comply with the scanning task?
- Measures of compliance
 - ► Used app every day: scan rather than direct entry
 - Number of purchases reported
 - Time between purchase and scan
- Lessof et al. (2018): Can we use an app on a smartphone or tablet to collect detailed expenditure data? Paper presented at the NCRM Research Methods Festival, Bath, UK



Outcome quality

How good are expenditure data collected with a receipt scanning app compared to benchmark data?

Research questions



- 1. Total expenditure in app compared to benchmark?
- 2. Category expenditure in app compared to benchmark?
- 3. Direct entry option important for which categories?
- 4. Which types of participants responded well to app task?
- 5. Different **substantive conclusions** when using app data compared to benchmark?



UK Living Costs and Food Survey (LCF)

- ► Main survey on household expenditure in UK
- ► Stratified, clustered random sample of HHs in UK
- Adults aged 16+ invited to complete paper spending diary for two weeks
- ► Fieldwork: April 2016-March 2017
- ► HH response rate: 45%

UKHLS App Study

- Receipt scanning app
- General-purpose survey
- Collect data on total expenditure
- Diary period: four weeks
- Expenditures collected in app
- Excluding Northern Ireland
- ► Oct-Dec 2016
- ► N = 259

- Paper spending diary
- Expenditure survey
- Collect data on item-level expenditure
- ► Diary period: two weeks
- Expenditures collected in
 diary + HH survey
- ► Including Northern Ireland
- Apr 2016-Mar 2017
- ► N = 2,213

UKHLS App Study

- Receipt scanning app
- General-purpose survey
- Collect data on total expenditure
- ► Diary period: two weeks
- Expenditures collected in app
- Excluding Northern Ireland
- ► Oct-Dec 2016
- ► N = 259

- Paper spending diary
- Expenditure survey
- Collect data on item-level expenditure
- ► Diary period: two weeks
- Expenditures collected in
 diary + HH survey
- ► Including Northern Ireland
- Apr 2016-Mar 2017
- ► N = 2,213

UKHLS App Study

- Receipt scanning app
- General-purpose survey
- Collect data on total expenditure
- ► Diary period: two weeks
- Expenditures collected in app
- Excluding Northern Ireland
- ► Oct-Dec 2016
- ► N = 259

- Paper spending diary
- Expenditure survey
- Collect data on item-level expenditure
- ► Diary period: two weeks
- Expenditures collected in
 diary + HH survey
- Excluding Northern Ireland
- Apr 2016-Mar 2017
- ▶ N = 2,213

UKHLS App Study

- Receipt scanning app
- General-purpose survey
- Collect data on total expenditure
- ► Diary period: two weeks
- Expenditures collected in app
- Excluding Northern Ireland
- ► Oct-Dec 2016
- ► N = 259

- Paper spending diary
- Expenditure survey
- Collect data on item-level expenditure
- ► Diary period: two weeks
- Expenditures collected in diary + HH survey
- Excluding Northern Ireland
- ► Oct-Dec 2016
- ► N = 2,213

Method

Inverse probability weighting to match sample composition of app study to benchmark data

Sump	ie compositio	in uncer	
		LCF	App-LCF
Mean Age		50.0	-0.4
% Female		51.7	-3.3
% Employe	d	58.7	-5.6
% Income	Q1	26.0	-5.5
	Q2	23.8	-3.4
	Q3	24.4	+5.7
	Q4	25.9	+3.2
% House ov	vned	72.4	-0.5
Mean HH s	ize	2.6	-0.1
Mean $\#$ ch	ildren in HH	0.6	-0.1
% Urban		75.5	+2.7
% Compute	er in HH	92.1	+3.5

Sample composition after IPW

Total expenditure



	Mean	SD	Q1	Median	Q3	Ν	KS-test
LCF	148.2	131.4	56.7	115.5	196.6	2,213	
Scan+DE	165.6	187.1	57.2	115.8	209.2	259	n.s.
Scan only	116.7	149.4	41.5	76.1	127.7	236	***

Category expenditure



Coding of scanned receipts

- 1. Data from images manually entered into database by Kantar Worldpanel
- 2. Automated category assignment for each item
 - Item description parsed
 - Compared against consumption classification COICOP
 - COICOP category assigned, then recoded into one of ten spending categories used in app
- 3. Category assignment manually checked

Category expenditure

MORRISONS Feel childs for you									
Nok1 Manager : Telephone : O Vat Number :	ng Lee King 1483 755 3434753	552 55							
Pick up a Car Saving for Chr DATE: 19/06/200 TILL: 0019 YOU were served	d and St istmas T 8 TIME: NO: by: JEN	art oday 17:54 01968232 I							
DESCRIPTION		2							
'M' FRESH SEA BREAM 'M'SIDE OF SALMON 'M' KIPPER FILLETS 'M' PORK LEG STEAK		2.88 D 3.08 D 0.56 D 2.93 D							
M BHDDDDL1 0.270kg @ £1.99/kg M BABYLEAF AND HER 'W' PCD CD/DES	2	0.54 D 1.39 D							
n ACO UNITED SUBJECT OF A CONTROL OF A CONTR		2.59 D 0.99 D 1.34 D 1.29 D 1.88 D 0.99 D 0.55 D 0.94 D 0.94 D 0.94 D 1.86 O 1.96 O 1.96 O 2.59 A							
Items Sold: 20	TOTAL	£35.54							
	CASH	£40.00							
Change	3	£4.46							
VAT A 17.5% (1 VAT B 5.0% (1 VAT D 0.0% (1 VAT Total	22.59) 5.98) 26.97)	£0.39 £0.28 £0.00 £0.67							
MULTI £0 SAVI AT MORE Thank you for sho Please of	SAVE .08 NGS ISON pping a call aga	IS IS t Morrisons In							



Receipt-level data

- Store
- Date and time of purchase
- Total amount
- Payment method
- Loyalty card used

Item-level data

- Product description
- Price
- Units bought + price per unit

Category expenditure

Spending categories

- 1. Food and groceries
- 2. Clothes and footwear
- 3. Transport costs
- 4. Child costs
- 5. Home improvements and household goods
- 6. Health expenses
- 7. Socialising and hobbies
- 8. Other goods and services
- 9. Holidays (= Vacation)
- 10. Giving money or gifts to other people
- 11. Other

Based on qualitative research by d'Ardenne & Blake (2012)



Category expenditure



Coding of direct entries

- ► If one spending category selected: no additional coding
- If multiple spending categories selected:
 - Amount divided by ratio of category expenditures from receipt scanning + direct entry



	Mean	SD	Q1	Median	Q3	N	KS-test
LCF	12.8	30.4	0.0	0.0	13.5	2,213	
Scan+DE	11.3	20.6	0.0	0.0	12.5	259	n.s.
Scan only	8.6	17.8	0.0	0.0	9.8	236	n.s.



	Mean	SD	Q1	Median	Q3	N	KS-test
LCF	28.5	40.1	6.0	15.8	35.0	999	
Scan+DE	23.1	24.5	5.9	13.3	32.0	134	n.s.
Scan only	18.7	22.3	2.5	10.7	24.5	115	***



	Mean	SD	Q1	Median	Q3	N	KS-test
LCF	32.3	32.3	5.1	24.4	49.4	2,213	
Scan+DE	40.9	30.2	14.6	38.3	60.8	259	***
Scan only	34.3	27.0	12.9	29.7	54.1	236	**



	Mean	SD	Q1	Median	Q3	N	KS-test
LCF	36.9	32.0	11.7	30.1	53.1	1,935	
Scan+DE	43.5	29.2	16.2	41.9	64.3	248	***
Scan only	36.3	26.4	14.6	32.4	55.6	224	**



	Mean	SD	Q1	Median	Q3	N	KS-test
LCF	37.8	51.4	7.5	22.8	49.4	2,213	
Scan+DE	16.6	21.2	0.8	8.3	22.0	259	***
Scan only	5.8	15.1	0.4	0.4	4.9	236	***



	Mean	SD	Q1	Median	Q3	N	KS-test
LCF	42.0	52.6	11.5	26.8	54.0	1,988	
Scan+DE	21.6	21.8	5.1	13.5	33.9	194	***
Scan only	10.7	19.2	1.3	4.0	10.1	129	***

Total expenditure by subgroup



		Female			Male	
	Median	N	KS-test	Median	N	KS-test
LCF	133.3	1,144		98.4	1,069	
Scan+DE	92.6	160	**	129.6	99	*
Scan only	77.3	149	***	71.9	87	n.s.

34 / 40

Total expenditure by subgroup



	Income: Below median			Income	Income: Above median		
	Median	Ν	KS-test	Median	Ν	KS-test	
LCF	88.1	1,102		143.8	1,111		
Scan+DE	74.2	102	*	143.3	157	n.s.	
Scan only	68.1	91	***	77.3	145	***	

Conclusions



- Promising method for collecting high-level expenditure data in general population
- Total expenditure (scan + direct entry) comparable to benchmark (LCF)
- Category expenditure more comparable for some categories than for others
- Direct entry option important for most categories
- Data quality varies by respondent characteristics

Discussion



Follow-up study

- Test methods to increase participation in mobile data collection
- ► Samples: Innovation Panel, Lightspeed Online Panel
- Adults invited to download expenditure diary app
 - ► Use app for 31 days
 - Own smartphone (iOS or Android)
 - ► No scanning. Direct entry only
- ► Fieldwork started May 2018

Discussion



How to increase participation?

- Browser-based alternative to app (sequential mixed-mode)
- ► Experiment #1: Invitation to use app
 - Within interview vs. letter plus email
- ► Experiment # 2: Feedback on reported spending
 - Additional section in app: cumulative summary of expenditure by category
 - Feedback promised vs. feedback not promised but given vs. no feedback
 - Effect on participation + reported spending?

More information



Project webpage https://www.iser.essex.ac.uk/research/projects/understandinghousehold-finance-through-better-measurement



Thank you.

Alexander Wenz awenz@essex.ac.uk