Paying to be poor:

Uncovering the scale and nature of the poverty premium

Methodological Appendix

Sara Davies, Andrea Finney and Yvette Hartfree

November 2016

University of BRISTOL
Personal Finance Research Centre

This Appendix was produced by

Sara Davies, Andrea Finney and Yvette Hartfree

Personal Finance Research Centre

University of Bristol

Contents

1 Defining low-income households

In this study low-income households were defined as those with a household income of 70 per cent of median or below (after housing costs). This definition is higher than the standard 'poverty line' measure of 60 per cent of median income, used in the UK and other European Union countries to classify people as poor.

This higher threshold was used for both theoretical and pragmatic reasons. From a theoretical perspective, households with an income of 70 per cent of median are still likely to find it hard to make ends meet and so may share some of the same risk factors for incurring a poverty premium as households living in poverty. Comparing poverty thresholds to the UK Minimum Income Standard, it shows that for non-pensioner households the Minimum Income Standard is higher than the threshold of 70 per cent of median income (Davis et al., 2016).[1] This indicates that households who have an income of 70 per cent of median income are still likely to have difficulty affording goods and services that are considered necessary to achieve a socially acceptable minimum standard of living. In other words, they nonetheless have ***low incomes***.

From a practical perspective, a higher threshold accounts for the practical difficulties in screening households into research by income, where income is necessarily only captured approximately. In other words, it allows for measurement error and ensures coverage of those who are not poor, but who may be on the fringes of poverty. Its effect is also to increase the pool of people from which to sample. This was considered to be a particular issue for recruiting pensioner households where only 14 per cent of pensioner households are below the poverty line. By using a threshold of 70 per cent of median income the pool of pensioners increases to 24 per cent.

For recruiting focus group participants, proxies were used to identify working age households with low incomes, these were receipt of Working Tax Credit or Housing Benefit. For pensioner households an

[1] Davis, A., Hill, K., Hirsch, D., and Padley, M. (2016) *A Minimum Income Standard for the UK in 2016*, York: Joseph Rowntree Foundation.

income threshold equivalent to slightly above receipt of State Pension plus Pension Credit was used. For recruiting low-income households to participate in the survey equivalised income thresholds of 70 per cent of median income (after housing costs) were used, whereby households with an income above this level were screened out.

2 Focus groups

The aim of the focus groups was to understand how people living with low incomes experience the poverty premium on a day-to-day basis. The focus groups asked people about how they shopped and paid for goods and services, the extent to which they felt they could access the best deals, and the areas of household spending where they struggled most to get a good deal.

Seven focus groups were held with low-income households in February 2016. They were designed to cover the following dimensions: region, urban and more rural areas, and age / life stage (as shown in Table 1). The areas were chosen on the basis of their having higher levels of poverty. A recruitment questionnaire was used to identify households on a low-income. Receipt of Working Tax Credit or Housing Benefit was used as a proxy to identify low-income working age households and among pensioner households an income threshold slightly higher than receipt of State Pension plus Pension Credit was used.

In addition the groups aimed to include a mix of:

- males and females;
- working and non-working households (among those aged 18-55);
- households with and without a car in the town and rural town locations (Telford and Bideford);
- pensioner households who did and did not use the internet for shopping and paying bills;
- and to include some households who paid for their gas or electric using a pre-payment meter.

The groups lasted around 90 minutes with eight people attending each group.

Table 1 Focus group design

Group	Age group	Location	Area Type
1	65+	West Midlands - Telford	town
2	26-55	West Midlands - Telford	town
3	18-25	Glasgow - Govan	city
4	26-55	Glasgow - Govan	city
5	65+	South West - Bideford	rural town
6	26-55	South West - Bideford	rural town
7	26-55	London - Hackney	city

The focus groups were recorded and transcribed verbatim. The transcripts were analysed thematically by area of household spending (identified in previous research) e.g. fuel, telecommunications, to identify the key issues and understand the experiences of participants.

3 Survey

A module of survey questions about the key elements of the poverty premium was included on an Ipsos-MORI face-to-face omnibus survey in Spring 2016. The survey covers a nationally and regionally representative sample of 2,000 adults in Great Britain per wave. All interviews were carried out by Ipsos MORI interviewers in-home, using CAPI (Computer Assisted Personal Interviewing). The data were weighted to the general population.

The survey module included screening questions to identify low-income households and to ensure that the survey was answered by either the Chief Income Earner in the household or their spouse / partner. The survey module was included on two omnibus survey waves in April and May 2016 to achieve a sufficient sample size of low-income households, defined as those with a household income of 70 per cent of median income or below (after housing costs).

Respondents were screened into the poverty premium module if they said their income was less than or around the same as the given threshold for 70 per cent median income, equivalised to take account of their household's composition (the number of adults and children making up the household). This resulted in an initial sample of 1,129 of potentially low-income households. We additionally took the precaution to screen out from the analysis those whose income band, given in response to a later question in the omnibus survey, was sufficiently higher than the upper income threshold for their household's composition for them to be realistically considered as living at or around the 70 per cent median income threshold. This resulted in a sample available for analysis of 947 respondents. As such, we can be confident that, based on the information available, the households represented by the resulting sample were living at, around or below 70 per cent median income for the country as a whole. We refer to these households as low-income households.

In addition to the survey questions, the omnibus also collected data on respondents' socio-demographic characteristics. The background characteristics of survey participants are shown in Table 2. All analyses were undertaken using IBM SPSS based on weighted data.

Table 2 Demographics of survey respondents

		Weighted data		Unweighted data	
		Frequency	Percent	Frequency	Percent
Household composition	Single adult	177	20.1	204	21.5
	Two adults	72	8.1	79	8.3
	Single adult with children	230	26.0	250	26.4
	Two adults with children	161	18.2	161	17.0
	Three or more adults without children	156	17.7	168	17.7
	Three or more adults with children	87	9.8	85	9.0
	Total	882	100.0	947	100.0
Whether CIE was of working age	No	207	23.4	231	24.4
	Yes	656	74.4	693	73.2
	Refused	19	2.2	23	2.4
	Total	882	100.0	947	100.0
Tenure	Being bought on a mortgage	142	16.1	123	13.0
	Owned outright by household	200	22.6	219	23.1
	Rented from local authority	172	19.5	215	22.7
	Rented from a private landlord	262	29.7	255	26.9
	Belongs to housing association	102	11.6	131	13.8
	Other	3	0.4	3	0.3
	Refused	2	0.2	1	0.1
	Total	882	100.0	947	100.0
Work status of Chief Income earner	CIE in full time work or self-employment	342	38.8	325	34.3
	CIE in part time work	102	11.5	102	10.8
	CIE retired	213	24.1	246	26.0
	CIE not working for some other reason	224	25.4	272	28.7
	Refused	2	0.2	2	0.2
	Total	882	100.0	947	100.0
Work status of respondent	Respondent in full time work	263	29.8	241	25.4
	Respondent in part time work	110	12.5	108	11.4
	Respondent retired	221	25.0	255	26.9
	Respondent not working for some other reason	276	31.3	330	34.8
	Refused	13	1.5	13	1.4
	Total	882	100.0	947	100.0
Social Grade	AB	95	10.7	79	8.3
	C1	208	23.6	256	27.0
	C2	217	24.6	217	22.9
	DE	363	41.1	395	41.7
	Total	882	100.0	947	100.0

Ethnicity	White	755	85.5	811	85.6
	Non-White	124	14.0	131	13.8
	Refused	4	0.4	5	0.5
	Total	882	100.0	947	100.0
Financial inclusion	Does not have a debit card	17	2.0	22	2.3
	Has a debit card	821	93.1	876	92.5
	Total	838	95.0	898	94.8
	Refused	44	5.0	49	5.2
	total	882	100.0	947	100.0
Household access to credit card	No	546	61.9	594	62.7
	Yes	292	33.1	304	32.1
	Refused	44	5.0	49	5.2
	Total	882	100.0	947	100.0
Level of digital inclusion	Has used the internet in last 3 months to buy products / services online	478	54.2	503	53.1
	Has used the internet in last 3 months to seach for products/services want to but NOT to buy products / services online	95	10.8	97	10.2
	Has NOT used the internet to search for products / services to buy or to buy them	309	35.0	347	36.6
	Total	882	100.0	947	100.0
Presence of car in household	No access to a car	352	39.9	405	42.8
	Access to a car	531	60.1	542	57.2
	Total	882	100.0	947	100.0
Income Band	system missing/not asked	2	0.3	3	0.3
	UP TO £4,499	50	5.7	52	5.5
	£4,500 - £6,499	52	5.9	63	6.7
	£6,500 - £7,499	58	6.6	68	7.2
	£7,500 - £9,499	72	8.2	83	8.8
	£9,500 - £11,499	60	6.8	65	6.9
	£11,500 - £13,499	49	5.5	55	5.8
	£13,500 - £15,499	43	4.9	47	5.0
	£15,500 - £17,499	35	4.0	37	3.9
	£17,500 - £24,999	97	11.0	95	10.0
	£25,000 - £29,999	32	3.6	33	3.5
	£30,000 - £39,999	15	1.8	15	1.6
	£40,000 - £49,999	2	0.3	2	0.2
	DON'T KNOW	129	14.7	135	14.3
	REFUSED	183	20.7	194	20.5
	Total	882	100.0	947	100.0
Area	Rural	207	23.4	23.4	23.4
	Suburban	209	23.7	23.7	47.1
	Urban	280	31.8	31.8	78.9
	Metropolitan	186	21.1	21.1	100.0
	Total	882	100.0	100.0	
Government office region	East Midlands	55	6.2	6.2	6.2
	Eastern	66	7.5	7.5	13.7

	London	99	11.2	11.2	24.9
	North East	49	5.5	5.5	30.5
	North West	92	10.4	10.4	40.9
	Scotland	93	10.5	10.5	51.4
	South East	101	11.5	11.5	62.9
	South West	67	7.6	7.6	70.5
	Wales	56	6.4	6.4	76.9
	West Midlands	98	11.1	11.1	88.1
	Yorks and Humberside	105	11.9	11.9	100.0
	Total	882	100.0	100.0	

4 Cluster analysis

The function of cluster analysis is to identify 'natural' structures within a data set based on multiple variables. It allocates cases (in this case households) into groups in such a way that maximises similarity (homogeneity) within each group while simultaneously maximising differences (heterogeneity) between the groups. Cluster analysis can therefore be seen as a statistical approach to segmentation or typology construction. The analysis uses a set of pre-defined measures of interest, in this case individual poverty premiums variables. Therefore, the objective here is to create homogenous and distinct groups of households based on the level and nature of their exposure to premiums across these areas. However, in order for cluster analysis to be considered successful, it not only has to identify distinct and interpretable groups, but those groups also need to be identifiable based on other characteristics, such as demographic characteristics.

For the purposes of this analysis, we have used 26 poverty premium variables. This is lower than the number of premiums we describe elsewhere, and reported in the resulting breakdowns of the clusters (including for the purposes of calculating the costs). This is because we removed the any home insurance and any car insurance premiums (used to calculate premiums associated with living in a deprived area) as these would tend to drive the cluster analysis too strongly by financial inclusion and exclusion (and car ownership), rather than by the experience of premiums per se. This reduced the available premiums to 25, however we additionally split out paper billing for broadband and landlines in order to better reflect the underlying, differential use of these services (even though use of paper billing on either is counted as a single premium when it comes to calculating the cost), resulting in 26 premium variables for use as the 'cluster variate'. These are shown in the list below. Note, that it treats the fuel switching premium as one variable, only being split out again for the calculation of the cost each resulting cluster incurs on average.

The cluster variate:

Household fuel-use

- Payment on receipt of bill, electricity
- Payment on receipt of bill, gas
- Prepayment meter, electricity
- Prepayment meter, gas
- Not switched to best fuel tariff

Paper billing

- Paper billing, electricity
- Paper billing, gas
- Paper billing, landline
- Paper billing, broadband
- Paper billing, mobile

Insurance

- Insurance for specific items
- Insurance for mobiles
- Direct debit payment, home insurance
- Direct debit payment, car insurance

Difficulty accessing good value stores for food and groceries

- Shopping premium

Access to money

- Fee-charging ATM
- Fee-charging cheque-cashing
- Prepaid card

Use of higher-cost credit

- Rent-to-own
- Payday loan
- Home collected loan
- Pawnbroking loan
- Subprime personal loan
- Subprime credit card
- Mail order catalogues
- Hamper schemes

Cluster analysis is a really a collection of multivariate techniques. The main distinction is between hierarchical and non-hierarchical methods. Hierarchical methods are powerful and sophisticated methods which can simultaneously evaluate a large number of solutions. Non-hierarchical methods are computationally less heavy than hierarchical methods but enable one or more preferred solutions to be refined iteratively, allowing the re-assignment of cases to alternative clusters to improve the initial solution. Here, we use both methods in combination: hierarchical clustering followed by a common non-hierarchical method, k-means clustering (also known as centroid clustering).

This, two-stage, approach is widely regarded as an optimal approach to producing the best, most parsimonious (simplest most distinguishing) solution. However, k-means clustering is not recommended for binary (no/yes) measures, which all of our poverty premiums take the form of. The two-stage clustering process was therefore undertaken not on the 26 raw, binary variables but on a subset of 23 composite variables, constructed using principal components analysis (PCA), which represented those 26 variables.

The use of PCA afforded two advantages. First, it returned continuous, or scale (measured scored on a scale with mean 0 and standard deviation of 1), variables which are amenable to both forms of cluster analysis. Second, it resolved inherent relationships which naturally existed between the binary variables and which might risk artificially driving the cluster analysis (for example, a household with a prepayment meter for electricity would be highly expected to also have a prepayment meter for gas, and this correlation alone might determine a cluster). Despite reducing the number of premium variables for use in the cluster analysis, this approach nonetheless retained the full richness of the data: all 26 binary variables remained represented by the 23 resulting continuous variables.

From the initial hierarchical analysis, we identified solutions ranging from three to eight clusters as favourable solutions (based on a diagnostic chart, the dendrogram, and Analysis of Variance, which examines the ratio of homogeneity to heterogeneity mentioned above). Based on these results, we requested solutions with three to eight clusters inclusive in the second stage of cluster analysis, k-means clustering, encompassing our two indicated solutions and one either side of these. The k-means clustering also used the cluster centres (or centroids) for each cluster, produced in the hierarchical

clustering, as the starting point (initial centre) for the clusters and refined the solutions based on these.

The k-means clustering found that five- and seven-cluster solutions were optimal (based on Analysis of Variance). The additional granularity offered by the seven-cluster solution, combined with acceptable cluster sizes, led us to select the seven-cluster solution as the preferred solution. Our interpretation of the resulting clusters was made based on an analysis of the clusters by the original binary variables.

We undertook a separate, parallel run of the original 26 binary variables in hierarchical clustering in order to compare and validate the results of the solution above. This confirmed that a seven-cluster solution was optimal and returned similar results in the composition of the clusters, including in respect of the largest group (of low exposure to the poverty premium), another large group defined by prepayment meter use and characterised by marginalisation along a number of dimensions, and one defined by exposure to poverty-related insurance premiums. While similar in composition, this cluster solution suffered from small sample sizes for two of the groups and was therefore not considered for interpretation as a solution in its own right.

We used **pen portraits** as a means of illustrating some of the clusters. The portraits provided have been chosen to represent the likely characteristics and potential premium exposure of households in a particular cluster. For each pen portrait, we have selected one survey respondent who represents the typical socio-demographic characteristics of the group. We have described them by these characteristics, albeit without disclosing their true identity (any stated names, gender or precise ages or incomes are fictitious). Then we have described the combination of actual premiums they reported that their household was exposed to. In this way, the portraits are not intended to be representative of all households in this cluster, but are instead illustrative of what members of a cluster may be exposed to.

5 Survey questionnaire

1. Which, if any, of these is the main method that you, or someone else in your household, pay for electricity for this home?

SINGLE CODE

1. Monthly direct debit or standing order
2. Monthly, after you have received the bill or statement (not by direct debit or standing order)
3. Quarterly (once every three calendar months) direct debit or standing order
4. Quarterly (once every three calendar months), after you have received the bill or statement (not by direct debit or standing order)
5. Using a pre-payment meter (keycard or token)
6. In some other way (please specify)
7. Not applicable/no electricity
8. Don't know

2. Which, if any, of these is the main method that you, or someone else in your household, pay for mains gas for this home?

SINGLE CODE

1. Monthly direct debit or standing order
2. Monthly, after you have received the bill or statement (not by direct debit or standing order)
3. Quarterly (once every three calendar months) direct debit or standing order
4. Quarterly (once every three calendar months), after you have received the bill or statement (not by direct debit or standing order)
5. Using a pre-payment meter (keycard or token)
6. In some other way (please specify)
7. Not applicable/no mains gas
8. Don't know

3. When, if at all, did you or someone else in your household last switch gas or electricity supplier? If you have switched both, please think about the one you switched most recently.

SINGLE CODE, ALLOW DK

1. In the last 12 months
2. More than 12 months ago, but less than two years
3. More than two years ago but less than five years ago
4. Five or more years ago
5. Have never switched

4. For which, if any, of the following types of services do you, or does someone else in your household, receive a paper bill in the post? CODE ALL THAT APPLY.

MULTI CODE, RANDOMISE, ALLOW NULL AND DK

1. Electricity
2. Gas
3. Landline
4. Broadband
5. Mobile phone

5. Now thinking about the following types of insurance, as far as you are aware, which, if any, do you or someone else in your household currently have? CODE ALL THAT APPLY

1. Home contents insurance (Contents insurance is insurance on the furniture or other household contents or any personal possessions)
2. Buildings insurance
3. Car, motorbike or any other vehicle insurance
4. Mobile phone insurance
5. Insurance for specific household items such as kitchen appliances, TV etc

6. And how do you or someone else in your household usually pay for your home contents insurance?

SINGLE CODE

1. Annually, upfront
2. Monthly, by direct debit
3. Other
4. Don’t know

7. And how do you or someone else in your household usually pay for your car or motorbike or other vehicle insurance?

If they have more than one vehicle, please code all that apply.

MULTI CODE 1-2, SINGLE CODE 3

1. Annually, upfront
2. Monthly, by direct debit
3. Other
4. Don’t know

8. How easy or difficult is it for you or someone else in your household to get to, what you consider to be, good-value stores for your food and grocery shopping?

SINGLE CODE, ALLOW DK, FORWARD AND REVERSE LIST

1. Very easy
2. Fairly easy
3. Fairly difficult
4. Very difficult

9. As you may know, some cash machines charge a fee for using them. As far as you are aware, in the last 12 months have you or someone else in your household used a cash machine that charged for using it?

MULTI CODE 1-2

1. Yes- I have
2. Yes- someone else in the household has
3. No- I have not but I don't know if someone else in the household has
4. No- no-one in the household has
5. Don't know

10. As far as you are aware, in the last 12 months have you or someone else in your household been charged a fee to cash a cheque?

For example this could have been done using Shopacheck, Cash Generator, The Money Shop etc.

MULTI CODE 1-2

1. Yes- I have
2. Yes- someone else in the household has
3. No- I have not but I don't know if someone else in the household has
4. No- no-one in the household has
5. Don't know

For the next question we will be asking you to fill out the survey by yourself.

11. As far as you are aware, which, if any, of these types of credit have you, or has someone else in your household, used in the last 12 months? CODE ALL THAT APPLY

MULTI CODE, RANDOMISE

1. 'Rent-to-own' store, where items are generally paid for weekly or monthly but are not owned by the purchaser until the cost has been paid in full, such as BrightHouse, Perfect Home, Buy as You View
2. Payday loan
3. Personal loan from company that collects payments from your home, sometimes called a 'doorstep lender' or 'home collected credit' e.g. Provident, Greenwoods
4. Pawnbroking loan

5. A personal loan from a company that does not require a good credit rating e.g. Satsuma Loans, Pounds to Pocket, Cash Converters
6. A credit card from a company that does not require a good credit rating e.g. Ocean, Luma, Vanquis
7. A prepaid card – this is a card that you load money onto and use like a debit or credit card
8. Mail order catalogues (where you pay in instalments) from a company that does not require a good credit rating e.g. Marisota, Jacamo, Fashion World, Park
9. Christmas food hamper scheme, where you pay in instalments, e.g. Park
10. None of these
11. Don't know
12. Refused

12. And what did you use this credit to pay for? <insert credit used> (repeat for each type of credit used)

MULTI CODE, RANDOMISE, ALLOW DK

1. Day to day spending such as food or bills
2. A holiday
3. Washing machine
4. Cooker
5. Fridge/freezer/ fridge freezer
6. TV
7. Bedframe or mattress
8. Sofa / armchair / 3-piece suite
9. Small electrical items e.g. microwave, hairdryer, kettle
10. Christmas
11. Other (please specify)

Personal Finance
Research Centre

Visit www.bris.ac.uk/geography/research/pfrc