Preferences for Cash vs. Card Payments: An Analysis using German Household Scanner Data

CALOGERO BRANCATELLI[1]

November 10, 2019

ABSTRACT

Using data on 19 million transactions from a German panel of up to 18,000 households from 2009 to 2015 I analyze household preferences for cash versus card payments. I exploit a unique institutional setting in which virtually all households have access free-of-charges to the same debit card, the *girocard*, which has the same liquidity as cash, and which allows to make payments of any, even small, denomination at all considered retailers without any additional benefits or costs. Still, I find that households exhibit great differences in their use of cash vs. card payments, and that these differences are also stable over the course of the covered seven years. The institutional setting allows to rule out common explanations for households' different payment choices.

JEL Classification: L5, D12, G20

Keywords: payment behavior, payment preferences, payment instruments, cash usage, consumer heterogeneity, homescan data

---

[1] Goethe University Frankfurt, Chair of Finance and Economics, Theodor-W-Adorno-Platz 3, 60323 Frankfurt, Germany; Email: brancatelli@econ.uni-frankfurt.de.

I thank my supervisor Roman Inderst for continuous support and guidance. I also thank Krzysztof Wozniak, Livio Stracca, Adrian Fritzsche, Sergey Turlo, Sebastian Schuler, Marco Kotschedoff and seminar participants at Goethe University Frankfurt, Deutsche Bundesbank and The Nielsen Company (Germany) GmbH for helpful comments. I want to emphasize that only the author is responsible for the contents of this paper, which do not necessarily represent the opinion of The Nielsen Company (Germany) GmbH.

I. Introduction

While over the last decades payments have shifted from cash to cards, at least in some countries cash is still prevalent (or even the most prominent means of payment). Using representative Nielsen Homescan data from around 18,000 German households over the period 2009 till 2015, accounting for around 19 million shopping trips for fast-moving consumer goods (FMCGs), I show how consumer preferences alone account for the still widespread use of cash, and that consumers have strongly heterogeneous preferences that are, in addition, very stable over time.

To draw these conclusions I rely on specific features of the German payment infrastructure. Virtually all German households have a debit card, the so-called girocard, as this is provided free of additional charges with their current account. There are also no (issuer-specific) benefits ("reward points") linked to card usage. German banks have co-operated since the 1990s to achieve such universal coverage, and at least for the covered time period all retailers covered in the Homescan data have installed the respective terminals and accept these cards free of charge and for any, i.e., even small, denominations.[2] In terms of liquidity, girocard payments and cash payments are equivalent.[3] Finally, almost all banks are part of three ATM-networks that allow cash withdrawals free of charge, and I also present analyses restricted to households living in urban environments with thus a reasonably comparable access to ATMs.

In sum, for all households included in the panel and the entire period from 2009 to 2015, I have the unique situation that virtually all households had the same type of debit card (girocard), which has the same liquidity as cash, and were able to make payments of any, even small, denomination at all considered retailers without any additional benefits or costs. Consequently, in contrast to the previous literature reviewed below, in this institutional environment such potential differences can be excluded as explanations for households' different payment choices. Controlling for transaction value, as well as other (socio-demographic) co-variates that are frequently used in the literature, I still find large and persistent differences in households' choices.

To isolate and quantify such differences in household payment choices, which are thus not explained by the availability or (monetary) costs and benefits of card vs. cash payments, I can make use of the following key features of my data set: First, the use of Homescan data, e.g., in contrast

[2] I can also rule out changes that would have made one means of payment faster than the other, or more attractive than the other, both for consumers or retailers. Furthermore, I conduct additional checks (not reported here) to rule out that retailers did not follow a policy of refusing acceptance of girocard payments for small denominations.

[3] The girocard is equipped with a chip that records the most recent payments, so that even when an immediate (online) verification for the required funds was not possible, an authorization is still possible. Even when consumers have an overdraft facility for their current account, it makes no difference whether they use this to withdraw cash or to make the respective payment with their card.

to point-of-sale (POS) data[4], allows me to attribute payments to individual households; second, in contrast to, for instance, household diaries[5], I obtain a large number of observations for each household, notably also as scanned purchases are not restricted to food items; third, by observing a large fraction of households for various years, i.e., for up to seven years, provided that households' payment choices are relatively stable over time, these can be measured relatively precisely; fourth, as most households are covered for (almost) the entire period of 7 years from 2009 to 2015, I can analyze how stable households' preferences are.

As virtually all of the analyzed FMCG transactions are made by either cash or debit card (that is, the discussed girocard), I employ a linear probability model and consider the payment decision as a binary choice problem.[6] My first analysis explores how households' choice of payment methods has changed over the analyzed seven years. When I do not employ household fixed effects, albeit using both transaction value as well as standard sociodemographic controls, I find that overall the likelihood of cash decreased by around 3 to 3.6 percentage points, i.e., by less than 0.5 percentage points per year. Notably, even this small change is largely due to a change in sample composition over time, as once I introduce household fixed effects, the cumulative change drops by two third to only around 1.3 to 1.6 percentage points from 2009 to 2015.[7] My first result is, thus, that in the aggregate households seem to have almost not changed their preferences between paying by card or with cash over seven years. The introduction of household fixed effects greatly improves the regression model's explanatory power, lifting the (adjusted) R^2 from around 0.2 to around 0.45. While this already speaks to the role of unexplained differences in household preferences, I explore this further as follows.

The higher propensity to use cash for smaller denominations combined with the large number of such transactions mask potential differences in household preferences. For this reason, I explore differences in household preferences mainly with a focus on transaction sizes between 40 Euro and 80 Euro. To obtain a sufficiently large number of shopping trips (observations) per household and to also be able to analyze the stability of preferences, for this analysis I further restrict attention

---

[4] Data on the store/POS level is used, for instance, in Klee (2008).

[5] Diaries, covering often only few days of transactions, are used, for instance, in Rysman (2007), Arango, Huynh and Sabetti (2015) or Wakamori and Welte (2017).

[6] Being able to restrict the choice set to only two options and applying a linear probability model is a great advantage of my setting as this makes the use and interpretation of household fixed effects standard (in contrast to non-linear models).

[7] Albeit I am mainly interested in the yearly changes as well as the role of household fixed effects, I also conduct the analysis by instrumenting transaction size with weekdays, as from Monday to Saturday there is largely a monotonically increasing relationship between transaction volume per shopping trip and weekday. This procedure, which is meant to account for a potential "ticket uplift" when customers pay by card follows Cohen, Rysman and Wozniak (2018).

to households that are in the sample for at least six (out of the possible seven) years. This is still the majority of households (ca. 60 percent). Controlling for transaction size and sociodemographic variables, I show that (thus unexplained) differences are large (and precisely measured) and stable over time. Notably, households which have a controlled propensity to use cash (as captured by the fixed effect) below the median in the first half of their duration, have a 87 percent probability of being still below the median in the second half of their duration, and a similar figure applies to the transition rate of households above the median.

Thus, this paper documents large and (over time) stable differences in preferences over cash and card payments between households, which are not explained by those differences that are commonly analyzed in the literature, such as differences in access to different means of payment or in acceptance between merchants. As the absence of other sources of potential variation restricts the scope of analyzing potential determinants of such observed differences,[8] the detected large and stable differences in household behavior may not tell us much about the strength of individual preferences. However, they suggest that at least in Germany the use of cash vs. card payments is unlikely to change unless there is a fundamental change in the payment infrastructure. As the current roll-out of contactless card payment technologies primarily affects payments for small denominations, it is at least not obvious that this will change behavior for higher denominations, which is the focus of my analysis of household differences.

The results of this paper contribute to a growing literature that documents and explains consumers' and merchants' choices between different payment methods, such as cash, cards, or cheques. As I noted above, the particular institutional setting allows me to abstract from differences in merchants' acceptance of different payment methods (see, e.g., Wakamori & Welte, 2017) as well as from potential differences in consumers' costs (see, e.g., Koulayev, Rysman, Schuh & Stavins, 2016). As I can rule out such differences, my focus is on consumers' (unexplained) heterogeneity. Such heterogeneity has also been studied elsewhere, e.g., by Kalkreuth, Schmidt and Stix (2014) for Germany or Shuh and Stavins (2010) for the US. Using diary information and a survey, Kalkreuth et al. (2014) stressed "behavioral" motives as an explanation for such heterogeneity, such as consumers' preference for keeping immediate control over their liquidity.[9] While I

[8] For instance, it seems unlikely that regional differences in crime rates are sufficiently important as an explanation for different cash holdings and thus cash usage (Alvarez & Lippi, 2009). Generally, girocard holders are also virtually insured by their issuers against the abuse of their card. Finally, though authorization of girocard payments frequently rely on online verification, this not a requirement (and offline transactions account for a large fraction of all transactions, where either information on last usage is taken from the embedded chip or retailers may also circumvent such verification altogether ("Lastschriftverfahren")).

[9] As already noted, I neither need to model card-adoption or, more generally, card holding, as a girocard is, without additional costs, part of a current account, while almost every German household has a current account (and thus a girocard). Such adoption, both for debit and credit cards, potentially triggered by reward

document large and stable differences between consumers, for lack of variation in the monetary benefits and costs or in merchants' acceptance, I cannot model their choice from primitives as, for example, in Fiebig, Keane, Louviere and Wasi (2010) or Wakamori and Welte (2017).[10]

Most related to my analysis is a recent paper by Cohen et al. (2018), which also makes use of Homescan data to analyze household choice between cash, cards, and cheques in the US. I share with their analysis the documentation of substantial (otherwise unexplained) differences in household preferences. The particular institutional environment of my analysis allows me, however, to focus only on two means of payment, which, using a linear probability model, allows for a much simpler consideration of household fixed effects. The institutional environment also allows me to safely rule out various other explanations for differences in household choices, as explained above.

The remainder of this paper is structured as follows. In section II, I describe my data and the institutional setting of the analysis. I present my empirical strategy and estimation results for the analysis of aggregate payment preferences over time in section III. Section IV is dedicated to my analysis of individual household preferences. I conclude in section V. Full estimation results are presented in Appendix A and additional descriptive statistics in Appendix B.

II. Data Description and Institutional Background

i. Description of Data

I employ Nielsen Homescan data provided for Germany covering the period starting from January 2009 till December 2015. The data records for up to 18 thousand households all shopping trips in the Fast-Moving Consumer Goods (FMCG) segment, largely food retailing and drugstores.[11] For some of the following analysis, I identify 15 major retailers (12 grocery retailers and 3 drugstores).

programs, is instead modelled, for instance, in Simon, Smith and West (2010), Rysman (2007) or Zinman (2009).

[10] In this literature, a survey documents various stated reasons for the preference of one payment method over the other, such as speed, avoidance of fraud, fees or reward systems, delayed payment, or the possibility to get back cash. As already noted, I can rule out at least objective differences in fraud, also over time, card payments never generated costs or benefits for consumers, and, to what I know, cash-back was only introduced later (and is still not very common across all retailers). As for speed, various communications with one of the largest issuer of payment terminals suggest that, over the considered time period, I can exclude substantial advances in the speed of girocard payments, and I can also rule out differences at least across the larger retailers.

[11] The original sample size was 14 thousand households per year, which was increased to around 18 thousand in 2010.

The initial data comprises over 19 million such observations in total. Upon preliminary cleaning[12] of the dataset, I impose the following two main sample restrictions:

I drop all transactions that were not paid by either cash or debit card (which is almost always the aforementioned girocard; see also below). I thereby drop 3.2 percent of observations.[13] I finally keep only so-called "well-reporting" households. These can be easily identified in the dataset as only these households have been provided with a representative weighting factor by Nielsen. I thereby drop 16.77 percent of all observations. In total, I am left with over 14 million shopping trips.

Overall, cash is used for 84.07 percent of all transactions, with card payments accounting for the remaining 15.93 percent. In terms of value, the card share is predictably substantially larger and equal to 32.51 percent. The average transaction value of a card payment is 37.63 Euro and thus more than double the average cash bill (of 18.60 Euro). Table 1 provides an overview over the usage of cash and card payments, where for ease of exposition I have focused on the initial and final year of the entire time period, while calculating the respective cash share for various transaction sizes.[14]

Finally, I note that over half of the households in my sample are covered for almost the entire period. As noted below, some of the subsequent analysis is conducted only for these households (precisely, for all households that are covered for at least six years). Figure 4 in Appendix B shows the distribution of households over years in the sample. I observe various sociodemographic covariates, which are reported for the household head[15], of which I employ income, age, gender,

---

[12] Initially, in around 6.8 percent of observations the payment method is missing, so I drop these observations. In few cases, households visited the same store more than once in a given day, using different means of payment. This affects 0.5 percent of all observations, which are dropped from the analysis due the absence of enough detail in the time stamp of the purchase to differentiate intraday purchases within one store. Furthermore, while a scan should typically take place when the shopping bags are unpacked, I still employ the following restriction. As restricted shopping hours for conventional retailers should make Sunday shopping exceptional, I drop all observations with a Sunday timestamp (2.4 percent). I also truncate at the top and bottom 0.1 percent with regards to transaction value for outlier correction. Notably, all number refer to the initial number of observations, unless otherwise stated

[13] Of these, 1.33 percent were made up by credit card payments. The remaining 1.86 percent comprise "other payment methods", which are reportedly made up mostly of vouchers. Cheques are no longer in use in Germany (as the introduction of the girocard system by German banks served also the purpose of replacing cheques).

[14] The average transaction value of a card payment, over all years, is equal to 36.14 Euro over all grocery stores and equal to 26.70 Euro over all drugstores. The respective values for cash payment are 16.92 Euro and 10.86 Euro.

[15] The household head is defined as the person conducting the purchases and running the household most of the time, i.e. it is not necessarily the main earner. Household income, however, refers to the main income of the household.

family status, education, next to the number of household members. Descriptive statistics are provided in Table 10 in Appendix B.

Table 1: Mean and standard deviation of cash share

Transaction value	Mean 2009	SD 2009	Mean 2015	SD 2015
[0€,20€]	94.41	12.05	92.77	14.40
]20€,40€]	86.60	20.42	85.80	21.81
]40€,60€]	82.73	24.28	82.54	24.77
]60€,80€]	82.04	25.78	82.38	25.80
>80€	81.17	27.32	80.67	27.14

Notes: Average computed over household cash shares in the respective year.

ii. Institutional Background

The obtained Homescan data confirms evidence from previous research that for German households and for FMCG, the relevant payment choice is restricted to that between cash and the girocard, which is a debit card that is provided free of charge as part of a current account. Virtually all households in my final sample have evidently access to a girocard as they use this at least once as a means of payment.[16]

The girocard system is a domestic payment system to which all German and foreign banks have access.[17] The girocard is also a multi-purpose card that is used to withdraw cash from the issuer's or another bank's ATM.[18] When a transaction is made, typically this needs to be authorized by the issuing bank, which is either done online or offline by comparison with data stored on a chip. As such authorization is only provided when sufficient funds are available on the respective account, card payments do also not extend any form of credit and do not provide additional liquidity to the household. I am also not aware of any other benefits, such as reward programs, that were at any time linked to girocard payments. Merchants pay a small fee to the issuing bank (typically, between 0.2 percent and 0.3 percent). I confirmed that all major (15) retailers in the sample, which account for 70 percent of all transactions, accepted these girocards at any point in time in the sample and for all denominations. It is prohibited that merchants charge customers for the use of girocards (differently from that of cash).

[16] Precisely, only 0.1 percent of households in the sample never use such a debit card. Furthermore, survey evidence suggests that even including under-age respondents, currently 98 percent of respondents hold such a card (Bundesbank, 2017).

[17] Various European countries had or still have national payment systems, including the Dutch "PIN", the Finnish "Pankkikortti" or the Luxembourgish "Bancomat" network.

[18] This is free of charge both at the issuer's bank as well as at all banks that belong to the same network of banks. Almost all banks belong to one of three such networks, i.e. that of private, public or cooperative banks. Withdrawals at other banks is possible with the girocard, but comes at an additional charge.

I can thus draw the following conclusion. For the entire period of seven years, households had the same access to girocard payments, and there were also no differences in terms of acceptance at individual merchants. For households such card payments did also not involve any additional monetary benefits or costs. As described in the introduction, I can thus abstract from many of the differences by which the extant literature has explained variations in households' preferences between different means of payment.

III. Analysis of Preferences for Cash vs. Card Payments over Time

i. Empirical strategy

In this section, I am interested in the question whether, *in the aggregate*, household preferences for cash vs. card payments have shifted over time. As stated above, since virtually all of the analyzed FMCG transactions are made by either cash or debit card (that is, the discussed girocard), the decision of the household can be modelled as a binary choice problem. For this purpose, I employ a linear probability model (LPM) with year fixed-effects, next to other covariates, which I estimate by means of Ordinary Least Squares (OLS). In a second step, I also extend my model with household fixed effects to capture the role of (unexplained) household preferences and to control, for instance, for changes in sample composition over time.

Despite the well-known limitations of the LPM, being able to restrict the choice set to only two options and applying OLS is a great advantage of my setting over previous literature as this makes the use, and later also the interpretation, of household fixed effects straightforward in contrast to non-linear models.[19] In fact, in the subsequent section, this allows me to then shift the focus of my analysis and ask whether *individual* households change preferences over time.

[19] The main limitation of the LPM, when used for prediction, is that the model yields probabilities outside the interval [0,1]. However, not being interested in predicting payment choices as such, the use of LPM with fixed effects not only makes the interpretation of fixed effects parameters standard and their estimation computationally easy, it also allows me to mitigate a common econometric problem encountered when estimating fixed effects, namely the *incidental parameters problem*. This econometric problem occurs in the absence of a sufficiently high number of observations per households (e.g. for households that are not in the panel over a long time or have documented only few shopping trips). As noted in Baltagi (2008), while for the linear panel data regression model with fixed T, other coefficients of interest included in the model can be estimated consistently, this is not the case in non-linear models such as Probit or Logit. For a detailed discussion on this in a similar context as in my paper, see Cohen et al. (2018). Notwithstanding, to address doubts on the model choice, I re-estimate my baseline model without fixed effects also using Probit and document average marginal effects in Table 9 of appendix A.III.

In the following, I implement the described empirical strategy using two empirical baseline models. As described, I first specify a linear model to estimate the conditional likelihood of paying cash without controlling for fixed effects. Formally, I estimate the following model:

$$cash_i = \beta \ln(V_i) + \gamma \boldsymbol{X}_k + \phi \boldsymbol{\Psi}_t + \epsilon_i \quad (1)$$

The dependent variable $cash_i$ is a binary indicator if transaction i is paid with cash; V_i represents the transaction value of transaction i; $\boldsymbol{\Psi}_t$ is a vector of year dummies; and the vector $\boldsymbol{X}_k$ contains a set of socio-demographic covariates, such as the number of household members or the age, gender, education and marital status pertaining to the household head of household k. The model is estimated with a constant. In the second baseline model, I add household fixed effects, represented by a vector of household dummies $\boldsymbol{\alpha}_k$, and drop time-invariant socio-demographic covariates. Thus, the second baseline model can be denoted as:

$$cash_i = \beta \ln(V_i) + \gamma \boldsymbol{X}_k + \phi \boldsymbol{\Psi}_t + \eta \boldsymbol{\alpha}_k + \epsilon_i \quad (2)$$

In both models, the coefficients of interest are collected in the vector ϕ and measure the aggregate *ceteris paribus* change in the likelihood of cash payments over time.

The baseline results are reported in column 1 and 3 of Table 2 respectively. While in the baseline regressions transaction value enters the model as logarithm, Table 2 also shows the following two variations for each model: Columns 2 and 5 report results using a set of dummy variables for the quintile of the respective transaction value. Columns 3 and 6 show analogous results using transaction value intervals of 20 Euro.

Moreover, to address the possibility that the transaction value may depend on the means of payment, I follow Cohen et al. (2018) and use the day of week of the transaction as an instrument. Table 2 shows the results only for the case of the baseline regression with household fixed-effects in column 7. Details on the instrumental variable strategy are contained in Appendix A.I.

Finally, I note that all estimates are reported with standard errors clustered at the household level.

ii. Estimation results

In this subsection, I discuss the main regression results presented in Table 2 with a focus on the change in (aggregate) behavior over time. Detailed estimates are reported in Appendix A.II.

Compared to the base year of 2009, over all transaction values the regression results without household fixed effects suggest a cumulative decrease in the likelihood of cash payments between 3 and 3.7 percentage points. While this is already small, as it amounts to a yearly change of around 0.5 percentage points only, this figure is still considerably smaller when one accounts for household fixed effects, as the cumulative change over seven years amounts then to under 1.7 percentage points. Note also that the result is slightly lower with the instrumental variable method (column

7). These findings suggest that, over time, there has been hardly any change in households' choice of payment methods, at least for FMCG transactions.

Table 2: Summary of OLS, FE and IV regression estimates on the determinants of cash payment

	Cash usage						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(value)	-0.134***			-0.121***			-0.133***
	(0.00)			(0.00)			(0.01)
Value quintiles	No	Yes	No	No	Yes	No	No
Value bins	No	No	Yes	No	No	Yes	No
Year 2010	-0.006***	-0.007***	-0.008***	-0.001	-0.001*	-0.002***	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2011	-0.013***	-0.013***	-0.015***	-0.002*	-0.002**	-0.004***	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2012	-0.019***	-0.020***	-0.023***	-0.004***	-0.005***	-0.007***	-0.002*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2013	-0.025***	-0.026***	-0.028***	-0.007***	-0.008***	-0.010***	-0.005***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2014	-0.030***	-0.031***	-0.032***	-0.010***	-0.011***	-0.013***	-0.008***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2015	-0.034***	-0.035***	-0.037***	-0.013***	-0.014***	-0.016***	-0.011***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Socio-Demographic Controls	Yes	Yes	Yes	Yes	Yes	No	No
HH Fixed-Effects	No	No	No	Yes	Yes	Yes	Yes
Instrumental Variable	No	No	No	No	No	No	Yes
Constant	Yes	Yes	Yes	No	No	No	No
Observations	12,931,206	12,931,206	12,931,206	14,116,209	14,116,209	14,116,209	14,116,209
R^2	0.185	0.203	0.198	0.447	0.458	0.452	
Adjusted R^2	0.185	0.203	0.198	0.446	0.457	0.452	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports estimates from the two baseline regressions, excluding fixed effects (columns 1-3) and including fixed effects (columns 4-6). The table also reports extended estimates from column 4 using the day of week as instrumental variable for the transaction value (column 7). In all specified columns, the dependent variable is a dummy for cash payment, which is 1 if the transaction was payed with cash and 0 if it was paid with card. Other means of payment are excluded from the analysis. Full regression results are reported in Appendix A. Standard errors are reported in parentheses and clustered at the household level. The base year for time fixed effects is 2009.

In Appendix A.III I report additional regression results showing that this finding is robust to restricting the sample to i) households being in the sample for at least six years and ii) "city dwellers", i.e. households living in urban areas, which further reduces any existing differences in access to cash.[20] In line with my estimates above, preferences exhibit almost no change over time in both cases. Admittedly, this stark finding is partially due to the overwhelming importance of cash payments for smaller denominations, which make up for the majority of FMCG transactions. In fact, when I restrict attention to shopping trips with a bill value exceeding 40 Euro, changes are

[20] To be able to identify urban households, I enrich my dataset with information on population densities on the zip code level from the German Statistical Office, Destatis. A household is considered *urban* if the population density in the respective zip code of residence exceeds 1,000 inhabitants/Km².

slightly larger. Accounting for fixed effects, the cumulative change over seven years amounts then to still under 1.8 percentage points as shown in Table 5 in the appendix.

In what follows, I demonstrate that even on the household level preferences (as captured through household fixed effects) are relatively stable. While for low transaction values this may not be surprising given the documented dominance of cash payments, for larger transaction values I do find pronounced differences in households' likelihood to use cash or card payments. In Table 2 the importance of accounting for such (persistent) differences in unexplained household preferences can already be seen by comparing the explanatory power of the different specifications as shown by the sharp increase in the adjusted R^2 from columns 1-3 to columns 4-6.

Finally, I follow recent work by Chen, Huynh and Shy (2019) and consider the possibility that cash usage may be driven not only by the transaction value as such but also by the perceived burden of the amount of coins needed to conduct a cash payment. I refer to transactions that can be paid with either no or low amount of coins as "easy-cash" transactions and to transactions involving high amount of coins as "non-easy cash" transactions respectively. I thus include a set of dummies to my baseline models and present results in Table 8.[21] Notice that the coefficients on the included dummies are highly significant. However, overall they add very little explanatory power compared to my baseline results and leave my main coefficients of interest nearly unaffected.

IV. Analysis of Individual Preferences

i. Setup of the analysis

In the previous section, I documented persistence in cash preferences in two ways. First, I demonstrated a substantial gain in explanatory power when adding household fixed effects in my baseline model. Second, accounting for household fixed effects, I observed only a very small change in card and cash preferences, *in the aggregate,* over time. In this section, I depart from this aggregate perspective and zoom in on the level of individual households to document heterogeneity and stability of individual payment behavior.

For this purpose, I impose additional sample restrictions before conducting the analysis. As already noted, the extreme propensity to use cash for smaller denominations combined with the large number of such transactions in the sample mask potential differences in household preferences. For this reason, I explore differences in household preferences mainly with a focus on transaction sizes in the interval between 40 Euro and 80 Euro. To obtain a sufficiently large number of

[21] I consider various definitions for „easy cash" transactions, namely integer transaction values, transaction values with only one decimal, linear combinations of 5 Euro and transaction values of 5, 10, 20 or 50 Euro as these are the most common bills in circulation.

shopping trips, i.e. observations, per household and to also be able to analyze the stability of preferences, for this analysis, I further restrict attention to households that are in the sample for at least six (out of the possible seven) years. This is the majority of households and obviously the vast majority of transactions.

The analysis evolves in two steps. First, to explore heterogeneity in payment preferences, I re-run regression 6, i.e. my baseline model with fixed effects replacing the logarithm of transaction value by a vector of dummies for transaction value intervals, denoted by the vector $\boldsymbol{Z_i}$. More formally, I thus estimate the following model:

$$cash_i = \beta \boldsymbol{Z_i} + \gamma \boldsymbol{X_k} + \boldsymbol{\phi \Psi_t} + \eta \boldsymbol{\alpha_k} + \epsilon_i \qquad (3)$$

For reasons of computational feasibility, I conduct the analysis on a random subsample of 5000 households. The imposed sample restriction allows me to estimate these household fixed effects with high precision and high statistical power since it ensures that I can rely on a substantial and stable amount of transactions for each household over time.[22]

Next, to focus on the intertemporal stability of preferences, I construct a dummy variable, m_k, splitting the duration in sample of each household at each individual midpoint and interact household fixed effects, as well as all other explanatory variables, with it.[23] Including all previous terms[24], I thus estimate the following extension of model 3 for the second part of my analysis:

$$cash_i = \beta_1 \boldsymbol{Z_i} + \beta_2 \boldsymbol{Z_i} \times m_k + \gamma_1 \boldsymbol{X_k} + \gamma_2 \boldsymbol{X_k} \times m_k + \boldsymbol{\phi}_1 \boldsymbol{\Psi_t} + \boldsymbol{\phi}_2 \boldsymbol{\Psi_t} \times m_k + \eta_1 \boldsymbol{\alpha_k} + \eta_2 \boldsymbol{\alpha_k} \times m_k + \epsilon_i \qquad (4)$$

Notably, adding these interaction terms to the previous model allows me to measure the change of each household's individual marginal propensity to pay with cash over time and, as I show below, to also quantify these changes in form of transition probabilities.

ii. Heterogeneity of preferences

Figure 1 documents the distribution of estimated household fixed effects estimated with model 3. The x-axis shows the magnitude of the estimated coefficients. Numerically, all values can be interpreted as differential marginal propensity of each household to pay with cash vis-à-vis a base

---

[22] While the imposed sample restrictions already substantially mitigate the potential harm of estimating parameters inconsistently, this problem of incidental parameters may still arise when including also households with only few shopping trips. To address this potential caveat, I run additional estimations of the baseline models not shown here, where I restrict the sample to households with more than 100 transactions over time, and verify that this leaves the results of my analysis virtually unchanged.

[23] More precisely, the dummy variable is one if the transaction was conducted after the midpoint and is zero otherwise. Notably, by allowing the midpoint to vary across households and by requiring households to be in sample at least six years, the individual midpoints may differ up to one year.

[24] For ease of exposition, I denote coefficients with subscripts to distinguish interacted from non – interacted terms below.

household that is dropped for obvious reasons of multicollinearity when running the estimation. Note that for my analysis of heterogeneity of payment behaviour I am interested in the shape of the distribution and the relative distance of households to each other. The distance between two households can be interpreted as the differential marginal probability of two households to pay with cash, and thus as indicator for differences of otherwise unexplained household cash preferences. For instance, the marginal probability to pay with cash of a household with a value of minus 0.8 is 60 percentage points larger than the marginal probability of a household with a value of minus 0.2.[25]

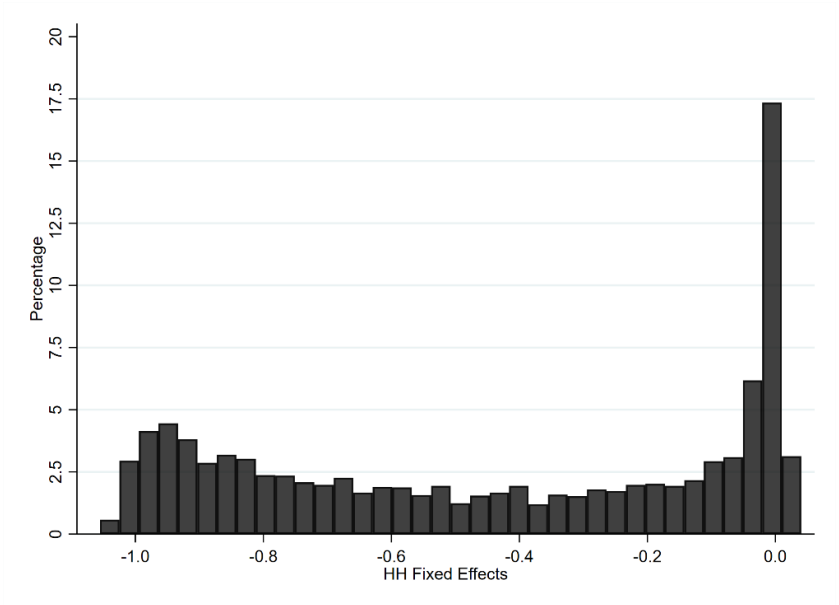

Figure 1: Distribution of household fixed effects

Notes: Sample restricted to 5000 randomly drawn households that were in the sample at least 6 years and conducted transactions in the interval [40€, 80€]. The figure shows the distribution of household fixed effects resulting from running regression 6 by means of OLS with a set of household dummies.

Notably, the distribution of fixed effects has two particular characteristics, namely it exhibits i) concentration around two peaks but also ii) a wide variation. Intuitively, the two peaks illustrate the fact that in the considered range of relatively high transaction values, some households stick to one payment method, either cash or card payment, and hardly randomize between the two.

It is noteworthy that the unique institutional setting of my analysis allows me to confidently relate such observed behavioural differences to differences in otherwise (unexplained) household preferences as I can rule out that any observed difference in payment behaviour stems from

[25] Admittedly, due to the well-known limitations of the LPM mentioned above, the approach also yields values outside the interval [0,1], for which this interpretation is not valid.

differences in access to the different payment forms or other confounding factors mentioned in previous literature.

However, despite the design of the analysis, the observed strength of variation of the fixed effects estimates still raises the question how precisely measured these coefficients are. To address potential concerns about the precision of the estimates, I propose a procedure of pairwise testing for equality of the estimated coefficients.

For this, I draw a random subsample of 500 household fixed effects (i.e. 10 percent of the regression sample) for which I run Wald tests on equality of each fixed effect vis-à-vis the other 499 fixed effects, extracting the two-sided p-value of every single test. Notably, all tests are conducted using clustered standard errors from the regression above. The analysis shows that 98 percent of the obtained p-values are below 1 percent, thus indicating highly statistical significant differences of the coefficients from each other. This finding suggests that the observed heterogeneity in Figure 1 is not a consequence of imprecise measurement of household preferences, but that this difference is rooted in deeply heterogeneous preferences across households, which are precisely identified through the setup of my analysis.

iii. Stability of preferences

So far, I showed that over the covered seven years German households, *overall,* did not change their preferences regarding the use of cash or card payments (or that they did so only to a very small amount). However, I also documented, at least for larger transaction amounts, pronounced differences in household preferences. In this section, I bring these two findings together and ask how stable such preference differences were over the seven years, *within household.*

For this purpose, as noted above, I estimate model 4, where I fully interact the previous regression with a household-specific dummy splitting transactions before and after the midpoint of the household's duration in sample. In analogy to the analysis of heterogeneity, these resulting interacted fixed effects can be interpreted as the household-specific change of the controlled marginal propensity to pay with cash over time.

Figure 2 shows the distribution of interacted household fixed effects. I find a pronounced peak around zero highlighting that, on the level of individual households, most preferences are fairly stable over time and therefore do not change.

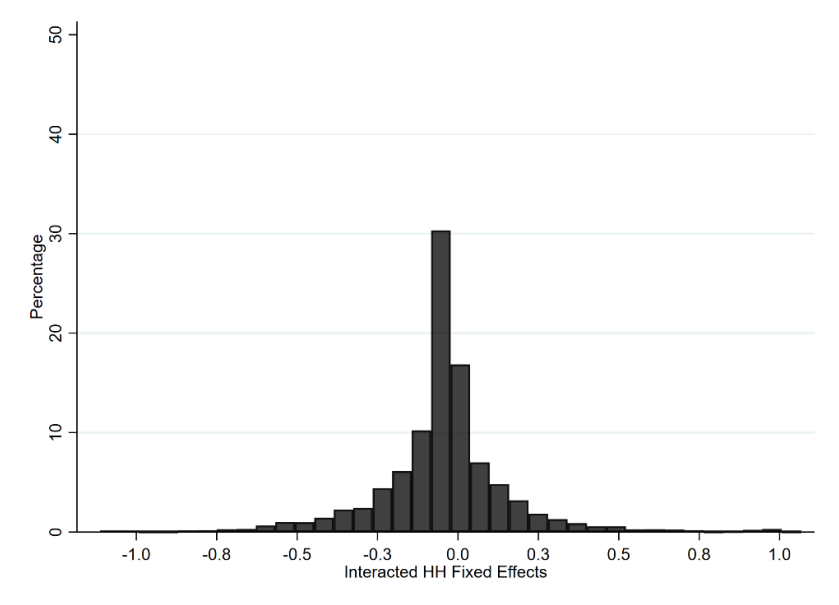

Figure 2: Distribution of household fixed effects interacted with individual midpoint of duration in sample

Notes: Sample restricted to 5000 randomly drawn households that were in the sample at least 6 years and conducted transactions in the interval [40€, 80€]. The figure shows the distribution of interacted household fixed effects obtained through regression 6 fully interacted with a dummy splitting transactions before and after the midpoint of the household's duration in sample.

Furthermore, to quantify this observed persistence on the household level, I estimate probabilities for the transition from any given position in the distribution of fixed effects over time. To accomplish this, I make use of the obtained fixed effects and interacted fixed effects parameters and proceed as follows:

First, in order to define a reference point for each household's preferences before the midpoint of the duration in sample, I group households into quartiles by ordering them according to the magnitude of their fixed effect parameter collected in vector η_I of model 4. Second, to define a point of comparison after the midpoint for each household, I compute the respective household quartile after the midpoint by adding up the fixed effect parameter with the interacted fixed effect parameter and ordering households accordingly. Based on these household positions in the distribution, I define the transition probability as the conditional relative frequency of households switching from a given quartile before the midpoint to any of the four quartiles after the midpoint. I report these probabilities in Table 3.

I find that households which have a controlled propensity to use cash (as captured by the fixed effect) below the median in the first half of their duration in sample, have a 87 percent probability of being still below the median in the second half of their duration, and a similar figure applies to the transition rate of households above the median. In terms of quartiles, thus the likelihood to stay in the same quartile lies between 54 percent and 73 percent across the four quartiles.

Table 3: Transition distribution of household fixed effects

	To Q1	To Q2	To Q3	To Q4
From Q1	71.91	22.90	3.71	1.48
From Q2	23.74	55.48	18.38	2.39
From Q3	3.38	17.97	54.74	23.91
From Q4	0.99	3.63	23.17	72.22

Notes: Sample restricted to 5000 randomly drawn households that were in the sample at least 6 years and conducted transactions in the interval [40€, 80€]. The table shows probabilities that a household transitions from a given quartile in the distribution of household fixed effects before the household individual midpoint to any quartile after the midpoint.

I have thus documented that household preferences are not only strongly heterogeneous but also stable and hence persistent over time. Put differently, the results presented in this section suggest that households who have a high initial propensity to pay with cash or card, tend to stick to their decision, despite the fact that both means of payment are equally available and equally employable at all points of sales of consideration over the entire time period.

V. Conclusion

Using detailed transaction data over seven years for a large panel of households, I analyze whether and how households differ in their payment choice and whether and how this has changed over time. The universal availability of the girocard for virtually all households at all covered retailers, for all denominations, and over the whole covered period, from 2009 to 2015, as well as the fact that the particular debit card has the same liquidity as cash, both allow me to confidently rule out various alternative explanations for the observed stark differences in households' preferences for cash vs. card payments even for given transaction sizes. My analysis is further simplified by the fact that over the considered period and for fast-moving consumer goods I can virtually rule out any other payment method but cash and the girocard.

Despite the dominance of cash for smaller transaction values, I document how (unexplained) consumer preferences alone account for the still widespread use of cash, and that consumers have strongly heterogeneous preferences that are, in addition, very stable over time.

A drawback of the particular institutional environment studied in this paper is, however, that the absence of any (exogenous) variations in the availability or the costs of different means of payment does not allow me to estimate from observed behavior alone the strength of households' preferences, e.g., in terms of compensating differentials. I thus do not know, for instance, how strongly households would react to a retailer's potential attempt to ban cash payments. While such a ban seems not to be imminent, card payments have recently become more attractive for retailers

due to the fact that the European Commission imposed a cap on the respective commissions.[26] Given households' documented persistent preferences to use one payment method over the other, it is however unlikely that, from retailers' perspective, such cost savings will materialize.

While this paper is being written, new payment technologies, such as girocards with Near Field Communication (NFC) chips, advance and affect the relative speed of cash and card payments at the point of sale at least for small denominations where no longer a pin is needed. However, whether possible small time savings for larger denominations are sufficient to tilt the observed stable behavior of those households with seemingly strong or at least stable preferences for cash over time, remains to be seen.

[26] The legal background for this is the adoption of the interchange fee regulation in the European Union (Regulation (EU) No 2015/751 of 29 April 2015 on interchange fees for card-based payment transactions), which from December 2015 onward capped interchange fees for cards issued and used in Europe (maximum of 0.2 percent for debit cards and 0.3 percent for credit cards). In Germany, the financial regulator applied these rules more strictly to the girocard system, basically imposing on card issuers the obligation to ensure that for each individual transaction the final ad valorem fee paid by the merchant must not exceed 0.2 percent.

References

Arango, C., Huynh, K. P., & Sabetti, L. (2015). Consumer payment choice: Merchant card acceptance versus pricing incentives. *Journal of Banking & Finance*, *55*, 130–141.

Baltagi, B. (2008). *Econometric analysis of panel data*. John Wiley & Sons.

Bundesbank, D. (2017). Zahlungsverhalten in Deutschland 2017-Vierte Studie über die Verwendung von Bargeld und unbaren Zahlungsinstrumenten. *Frankfurt: Deutsche Bundesbank*.

Chen, H., Huynh, K. P., & Shy, O. (2019). Cash versus card: Payment discontinuities and the burden of holding coins. *Journal of Banking & Finance*, *99*, 192-201.

Cohen, M., Rysman, M., & Wozniak, K. (2018). Payment choice with consumer panel data.

Fernando Alvarez, F. L. (2009). Financial Innovation and the Transactions Demand for Cash. *Econometrica*, *77*(2), 363–402.

Fiebig, D. G., Keane, M. P., Louviere, J., & Wasi, N. (2010). The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity. *Marketing Science*, *29*(3), 393–421.

Kalckreuth, U. L.F. von, Schmidt, T., & Stix, H. (2014). Using Cash to Monitor Liquidity: Implications for Payments, Currency Demand, and Withdrawal Behavior. *Journal of Money, Credit and Banking*, *46*(8), 1753–1786.

Klee, E. (2008). How people pay: Evidence from grocery store data. *Journal of Monetary Economics*, *55*(3), 526–541.

Koulayev, S., Rysman, M., Schuh, S., & Stavins, J. (2016). Explaining adoption and use of payment instruments by me consumers. *The RAND Journal of Economics*, *47*(2), 293–325.

Rysman, M. (2007). An Empirical Analysis of Payment Card Usage. *Journal of Industrial Economics*, *55*(1), 1–36.

Schuh, S., & Stavins, J. (2010). Why are (some) consumers (finally) writing fewer checks? The role of payment characteristics. *Journal of Banking & Finance*, *34*(8), 1745–1758.

Simon, J., Smith, K., & West, T. (2010). Price incentives and consumer payment behaviour. *Journal of Banking & Finance*, *34*(8), 1759–1772.

Wakamori, N., & Welte, A. (2017). Why Do Shoppers Use Cash? Evidence from Shopping Diary Data. *Journal of Money, Credit and Banking*, *49*(1), 115–169.

Zinman, J. (2009). Debit or credit? *Journal of Banking & Finance*, *33*(2), 358–366.

Appendix A Extended results

A.I Instrumental variable discussion

Following Cohen et al. (2018), I employ the *day of week* of a transaction as an instrumental variable in order to address the potential problem of simultaneous causality between payment choice and transaction value. Simultaneous causality may arise if the transaction value of a purchase is influenced by the means of payment that is used for a transaction. It is noteworthy that the institutional background of my analysis, through which I can rely on the availability of girocard and cash at virtually any given time, mitigates considerably the risk of simultaneous causality that could stem, for instance, from behaviour induced by impediments in the access to any of both means of payment. Overall, this makes an instrumental variable approach less relevant. Nevertheless, to rule out possible influences from such rare cases, I follow Cohen et al. (2018) and show that the variable day of week satisfies both conditions required of instruments: it is relevant, in that it has a direct impact on transaction value, and valid, in that it does not have a direct impact on the payment choice.

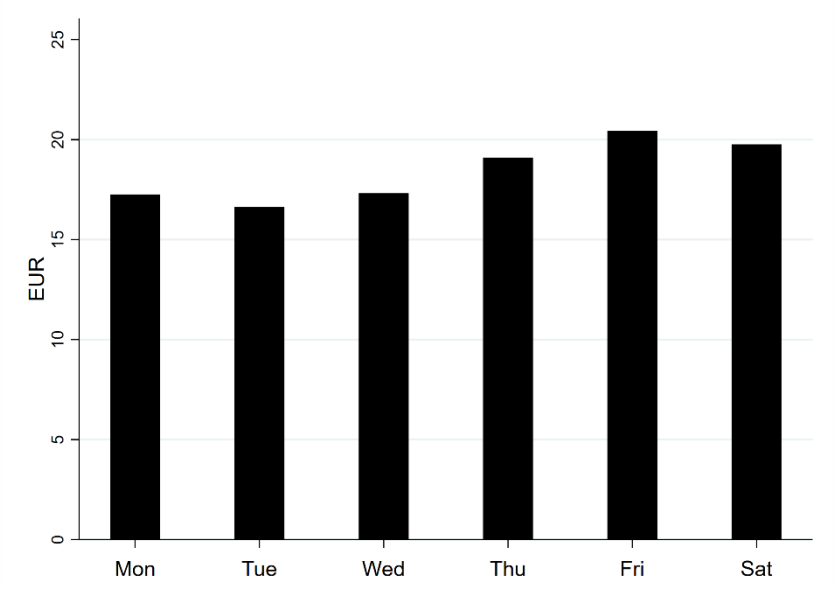

Figure 3: Average transaction value per week day

Notes: Sunday transactions excluded from the sample. The figures shows the average transaction value by day of week across households and years.

The relevance of day of week as instrument for transaction value becomes immediately visible in Figure 3, where I illustrate that consumers conduct smaller purchases on weekdays and larger shopping trips on Fridays and Saturdays, in analogy to findings in Cohen et al. (2018). Moreover, the institutional setting of my analysis ensures validity of day of week, in line with Cohen et al. (2018), as it rules out a potential relationship between the choice of cash or card and the day of week, as I can be confident that availability of either payment method is not subject to particularities of given days of the week, such as particular day-specific cash reliance, for the vast majority of my sample.

Furthermore, when using day of week as instrumental variable I rely on my assumption regarding the coincidence of scan date and transaction date stated above and prior exclusion of confounding Sunday transactions.

A.II Full Estimation results

Table 4: OLS, FE and IV regression estimates on the determinants of cash payment

	Cash usage						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(value)	-0.134***			-0.121***			-0.133***
	(0.00)			(0.00)			(0.01)
Value in]p20,p40]		-0.027***			-0.029***		
		(0.00)			(0.00)		
Value in]p40,p60]		-0.092***			-0.092***		
		(0.00)			(0.00)		
Value in]p60,p80]		-0.207***			-0.197***		
		(0.00)			(0.00)		
Value in >p80		-0.415***			-0.374***		
		(0.00)			(0.00)		
Value in]20€,40€]			-0.222***			-0.197***	
			(0.00)			(0.00)	
Value in]40€,60€]			-0.374***			-0.329***	
			(0.00)			(0.00)	
Value in]60€,80€]			-0.473***			-0.411***	
			(0.00)			(0.00)	
Value in >80€			-0.534***			-0.470***	
			(0.01)			(0.00)	
Year 2010	-0.006***	-0.007***	-0.008***	-0.001	-0.001*	-0.002***	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2011	-0.013***	-0.013***	-0.015***	-0.002*	-0.002**	-0.004***	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2012	-0.019***	-0.020***	-0.023***	-0.004***	-0.005***	-0.007***	-0.002*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2013	-0.025***	-0.026***	-0.028***	-0.007***	-0.008***	-0.010***	-0.005***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2014	-0.030***	-0.031***	-0.032***	-0.010***	-0.011***	-0.013***	-0.008***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2015	-0.034***	-0.035***	-0.037***	-0.013***	-0.014***	-0.016***	-0.011***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Low age (10-39)	-0.120***	-0.118***	-0.119***	0.011***	0.011***	0.010***	0.011***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
High income (>2250 EUR)	-0.035***	-0.033***	-0.032***	-0.004***	-0.004**	-0.003**	-0.004**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Female	0.019***	0.018***	0.016***				
	(0.00)	(0.00)	(0.00)				
Married/living together	0.029***	0.030***	0.029***				
	(0.00)	(0.00)	(0.00)				
Separated	0.012**	0.011**	0.010**				
	(0.00)	(0.00)	(0.00)				
Widowed	0.047***	0.045***	0.044***				
	(0.01)	(0.01)	(0.01)				
Apprenticeship or studied	-0.035***	-0.033***	-0.032***				
	(0.00)	(0.00)	(0.00)				
No. of persons in HH	-0.011***	-0.009***	-0.009***	-0.004***	-0.004***	-0.004***	-0.003**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	1.233***	1.052***	0.999***				
	(0.01)	(0.01)	(0.01)				
Observations	12,931,206	12,931,206	12,931,206	14,116,209	14,116,209	14,116,209	14,116,209
R^2	0.185	0.203	0.198	0.447	0.458	0.452	
Adjusted R^2	0.185	0.203	0.198	0.446	0.457	0.452	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports estimates from the two baseline regressions, excluding fixed effects (columns 1-3) and including fixed effects (columns 4-6). The table also reports extended estimates from column 4 using the day of week as instrumental variable for the transaction value (Column 7). In all specified columns, the dependent variable is a dummy for cash payment, which is 1 if the transaction was payed with cash and 0 if it was paid with card. Other means of payment are excluded from the analysis. Standard errors are reported in parentheses and clustered at the household level. The base year for time fixed effects is 2009.

Table 5: OLS, FE and IV regression estimates on the determinants of cash payment – Restricted on transaction interval [40€, 80€]

	Cash usage						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(value)	-0.264***			-0.182***			0.499***
	(0.01)			(0.00)			(0.11)
Value in]p20,p40]		-0.025***			-0.017***		
		(0.00)			(0.00)		
Value in]p40,p60]		-0.049***			-0.034***		
		(0.00)			(0.00)		
Value in]p60,p80]		-0.094***			-0.063***		
		(0.00)			(0.00)		
Value in >p80		-0.136***			-0.094***		
		(0.00)			(0.00)		
Value in]40€,60€]			-0.164***			-0.099***	
			(0.02)			(0.01)	
Value in]60€,80€]			-0.258***			-0.161***	
			(0.02)			(0.01)	
Year 2010	-0.010***	-0.010***	-0.010***	-0.003	-0.003	-0.003	-0.004
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2011	-0.021***	-0.021***	-0.021***	-0.005*	-0.005*	-0.005*	-0.008***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2012	-0.033***	-0.033***	-0.033***	-0.009***	-0.009***	-0.009***	-0.014***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2013	-0.035***	-0.035***	-0.036***	-0.008***	-0.009***	-0.009***	-0.018***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2014	-0.037***	-0.037***	-0.038***	-0.010***	-0.010***	-0.011***	-0.023***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2015	-0.042***	-0.042***	-0.043***	-0.015***	-0.015***	-0.017***	-0.032***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Low age (10-39)	-0.137***	-0.137***	-0.138***	0.018***	0.018***	0.018***	0.019***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
High income (>2250 EUR)	-0.094***	-0.094***	-0.095***	-0.006*	-0.006*	-0.006*	-0.009***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Female	0.007	0.007	0.006				
	(0.01)	(0.01)	(0.01)				
Married/living together	0.024**	0.024**	0.024**				
	(0.01)	(0.01)	(0.01)				
Separated	0.021	0.021	0.021				
	(0.01)	(0.01)	(0.01)				
Widowed	0.100***	0.100***	0.100***				
	(0.02)	(0.02)	(0.02)				
Apprenticeship or studied	-0.105***	-0.105***	-0.105***				
	(0.01)	(0.01)	(0.01)				
No. of persons in HH	-0.009**	-0.009**	-0.010**	-0.002	-0.002	-0.003	-0.009***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	1.765***	0.777***	0.909***				
	(0.03)	(0.02)	(0.02)				
Observations	1,255,427	1,255,427	1,255,427	1,359,619	1,359,619	1,359,619	1,359,619
R^2	0.043	0.043	0.040	0.555	0.555	0.554	
Adjusted R^2	0.043	0.043	0.040	0.548	0.548	0.547	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports estimates from the two baseline regressions, excluding fixed effects (columns 1-3) and including fixed effects (columns 4-6). The table also reports extended estimates from column 4 using the day of week as instrumental variable for the transaction value (Column 7). The regression sample is restricted to transactions between 40 and 80 Euro. In all specified columns, the dependent variable is a dummy for cash payment, which is 1 if the transaction was payed with cash and 0 if it was paid with card. Other means of payment are excluded from the analysis. Standard errors are reported in parentheses and clustered at the household level. The base year for time fixed effects is 2009.

A.III Robustness checks

Table 6: OLS, FE and IV regression estimates on the determinants of cash payment – Restricted on households in sample at least 6 years

	Cash usage						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(value)	-0.131***			-0.117***			-0.132***
	(0.00)			(0.00)			(0.01)
Value in]p20,p40]		-0.022***			-0.023***		
		(0.00)			(0.00)		
Value in]p40,p60]		-0.082***			-0.081***		
		(0.00)			(0.00)		
Value in]p60,p80]		-0.196***			-0.184***		
		(0.00)			(0.00)		
Value in >p80		-0.411***			-0.366***		
		(0.00)			(0.00)		
Value in]20€,40€]			-0.218***			-0.193***	
			(0.00)			(0.00)	
Value in]40€,60€]			-0.378***			-0.330***	
			(0.00)			(0.00)	
Value in]60€,80€]			-0.485***			-0.419***	
			(0.01)			(0.00)	
Value in >80€			-0.552***			-0.480***	
			(0.01)			(0.01)	
Year 2010	-0.007***	-0.007***	-0.008***	-0.001	-0.002*	-0.002***	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2011	-0.009***	-0.010***	-0.010***	-0.003***	-0.004***	-0.005***	-0.002*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2012	-0.013***	-0.014***	-0.016***	-0.005***	-0.006***	-0.008***	-0.003**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2013	-0.018***	-0.018***	-0.019***	-0.008***	-0.009***	-0.010***	-0.006***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2014	-0.022***	-0.022***	-0.022***	-0.011***	-0.011***	-0.012***	-0.008***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2015	-0.026***	-0.026***	-0.027***	-0.012***	-0.013***	-0.014***	-0.010***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Low age (10-39)	-0.117***	-0.115***	-0.115***	0.014***	0.014***	0.012***	0.014***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
High income (>2250 EUR)	-0.032***	-0.030***	-0.030***	-0.004**	-0.003	-0.003	-0.003*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Female	0.022***	0.020***	0.018***				
	(0.00)	(0.00)	(0.00)				
Married/living together	0.032***	0.031***	0.031***				
	(0.01)	(0.01)	(0.01)				
Separated	0.011*	0.010	0.010				
	(0.01)	(0.01)	(0.01)				
Widowed	0.052***	0.049***	0.048***				
	(0.01)	(0.01)	(0.01)				
Apprenticeship or studied	-0.031***	-0.029***	-0.028***				
	(0.01)	(0.01)	(0.01)				
No. of persons in HH	-0.015***	-0.013***	-0.013***	-0.004**	-0.004**	-0.003**	-0.003
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	1.224***	1.045***	0.999***				
	(0.01)	(0.01)	(0.01)				
Observations	9,078,027	9,078,027	9,078,027	9,926,133	9,926,133	9,926,133	9,926,133
R^2	0.187	0.210	0.208	0.444	0.459	0.455	
Adjusted R^2	0.187	0.210	0.208	0.444	0.458	0.455	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports estimates from the two baseline regressions, excluding fixed effects (columns 1-3) and including fixed effects (columns 4-6). The table also reports extended estimates from column 4 using the day of week as instrumental variable for the transaction value (Column 7). The regression sample is restricted to households being in the sample at least 6 years. In all specified columns, the dependent variable is a dummy for cash payment, which is 1 if the transaction was payed with cash and 0 if it was paid with card. Other means of payment are excluded from the analysis. Standard errors are reported in parentheses and clustered at the household level. The base year for time fixed effects is 2009.

Table 7: OLS, FE and IV regression estimates on the determinants of cash payment – Restricted on households living in urban areas (>1000 inhabitants/Km²)

	Cash usage						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(value)	-0.133***			-0.121***			-0.149***
	(0.00)			(0.00)			(0.01)
Value in]p20,p40]		-0.035***			-0.036***		
		(0.00)			(0.00)		
Value in]p40,p60]		-0.109***			-0.106***		
		(0.00)			(0.00)		
Value in]p60,p80]		-0.226***			-0.213***		
		(0.00)			(0.00)		
Value in >p80		-0.422***			-0.382***		
		(0.01)			(0.01)		
Value in]20€,40€]			-0.230***			-0.202***	
			(0.00)			(0.00)	
Value in]40€,60€]			-0.371***			-0.326***	
			(0.01)			(0.00)	
Value in]60€,80€]			-0.471***			-0.409***	
			(0.01)			(0.01)	
Value in >80€			-0.530***			-0.468***	
			(0.01)			(0.01)	
Year 2010	-0.005**	-0.005**	-0.007***	-0.002	-0.002*	-0.003**	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2011	-0.015***	-0.016***	-0.018***	-0.003*	-0.003**	-0.005***	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2012	-0.020***	-0.022***	-0.025***	-0.005***	-0.006***	-0.009***	-0.002
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2013	-0.025***	-0.027***	-0.030***	-0.009***	-0.009***	-0.012***	-0.004*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2014	-0.032***	-0.033***	-0.036***	-0.012***	-0.012***	-0.015***	-0.007***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2015	-0.036***	-0.036***	-0.040***	-0.015***	-0.016***	-0.019***	-0.010***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Low age (10-39)	-0.137***	-0.136***	-0.139***	0.017***	0.017***	0.018***	0.016***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
High income (>2250 EUR)	-0.029***	-0.026***	-0.025***	-0.005**	-0.005*	-0.005*	-0.004*
	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Female	0.020***	0.018***	0.016***				
	(0.01)	(0.01)	(0.01)				
Married/living together	0.038***	0.037***	0.036***				
	(0.01)	(0.01)	(0.01)				
Separated	0.002	0.001	0.000				
	(0.01)	(0.01)	(0.01)				
Widowed	0.043***	0.041***	0.040***				
	(0.01)	(0.01)	(0.01)				
Apprenticeship or studied	-0.040***	-0.038***	-0.037***				
	(0.01)	(0.01)	(0.01)				
No. of persons in HH	-0.018***	-0.016***	-0.017***	-0.001	-0.000	-0.001	0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	1.237***	1.067***	1.009***				
	(0.01)	(0.01)	(0.01)				
Observations	4,837,877	4,837,877	4,837,877	5,300,826	5,300,826	5,300,826	5,300,826
R^2	0.183	0.200	0.189	0.446	0.456	0.446	
Adjusted R^2	0.183	0.200	0.189	0.445	0.455	0.445	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports estimates from the two baseline regressions, excluding fixed effects (columns 1-3) and including fixed effects (columns 4-6). The table also reports extended estimates from column 4 using the day of week as instrumental variable for the transaction value (Column 7). The regression sample is restricted to households living in zip codes with at least 1000 inhabitants/Km². In all specified columns, the dependent variable is a dummy for cash payment, which is 1 if the transaction was payed with cash and 0 if it was paid with card. Other means of payment are excluded from the analysis. Standard errors are reported in parentheses and clustered at the household level. The base year for time fixed effects is 2009.

Table 8: OLS, FE and IV regression estimates on the determinants of cash payment – Full regression including non-easy cash indicators

	Cash usage						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(value)	-0.134***			-0.121***			-0.134***
	(0.00)			(0.00)			(0.01)
Value in]p20,p40]		-0.026***			-0.029***		
		(0.00)			(0.00)		
Value in]p40,p60]		-0.091***			-0.091***		
		(0.00)			(0.00)		
Value in]p60,p80]		-0.205***			-0.196***		
		(0.00)			(0.00)		
Value in >p80		-0.413***			-0.372***		
		(0.00)			(0.00)		
Value in]20€,40€]			-0.219***			-0.195***	
			(0.00)			(0.00)	
Value in]40€,60€]			-0.373***			-0.327***	
			(0.00)			(0.00)	
Value in]60€,80€]			-0.471***			-0.409***	
			(0.00)			(0.00)	
Value in >80€			-0.532***			-0.468***	
			(0.01)			(0.00)	
Year 2010	-0.006***	-0.007***	-0.008***	-0.001	-0.001*	-0.002***	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2011	-0.013***	-0.013***	-0.014***	-0.002**	-0.002**	-0.004***	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2012	-0.019***	-0.020***	-0.023***	-0.004***	-0.005***	-0.007***	-0.002*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2013	-0.025***	-0.026***	-0.028***	-0.007***	-0.008***	-0.010***	-0.005***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2014	-0.030***	-0.031***	-0.032***	-0.011***	-0.011***	-0.013***	-0.008***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2015	-0.034***	-0.035***	-0.037***	-0.014***	-0.015***	-0.017***	-0.011***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Low age (10-39)	-0.120***	-0.117***	-0.119***	0.011***	0.011***	0.010***	0.011***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
High income (>2250 EUR)	-0.035***	-0.033***	-0.032***	-0.004***	-0.004**	-0.003**	-0.004**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Female	0.019***	0.018***	0.016***				
	(0.00)	(0.00)	(0.00)				
Married/living together	0.029***	0.029***	0.029***				
	(0.00)	(0.00)	(0.00)				
Separated	0.012**	0.011**	0.010**				
	(0.00)	(0.00)	(0.00)				
Widowed	0.047***	0.045***	0.044***				
	(0.01)	(0.01)	(0.01)				
Apprenticeship or studied	-0.035***	-0.033***	-0.032***				
	(0.00)	(0.00)	(0.00)				
No. of persons in HH	-0.011***	-0.009***	-0.009***	-0.004***	-0.004***	-0.004***	-0.003**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Integer values	-0.014***	-0.017***	-0.022***	-0.012***	-0.013***	-0.019***	-0.011***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Integer values with no two-digit decimal	-0.004***	-0.015***	-0.027***	-0.000	-0.009***	-0.019***	0.004*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Lin. combination of 5 EUR values	-0.071***	-0.056***	-0.015***	-0.057***	-0.045***	-0.008**	-0.071***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)
5, 10, 20, and 50 EUR values	0.005	0.017***	0.006	0.005	0.012**	0.003	0.016**
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)
Constant	1.314***	1.118***	1.051***				
	(0.01)	(0.01)	(0.01)				
Observations	12,931,206	12,931,206	12,931,206	14,116,209	14,116,209	14,116,209	14,116,209
R^2	0.185	0.204	0.199	0.447	0.459	0.453	
Adjusted R^2	0.185	0.204	0.199	0.446	0.458	0.452	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports estimates from the two baseline regressions, excluding fixed effects (columns 1-3) and including fixed effects (columns 4-6). The table also reports extended estimates from column 4 using the day of week as instrumental variable for the transaction value (Column 7. All regressions contain a set of dummies controlling for non-easy cash transactions. Each dummy is 1 if the transaction value is defined as non-easy cash and 0 else. In all specified columns, the dependent variable is a dummy for cash payment, which is 1 if the transaction was payed with cash and 0 if it was paid with card. Other means of payment are excluded from the analysis. Standard errors are reported in parentheses and clustered at the household level. The base year for time fixed effects is 2009.

Table 9: Baseline OLS estimates and average marginal effects from Probit regression on the determinants of cash payment

	Cash usage					
	OLS			Probit		
	(1)	(2)	(3)	(4)	(5)	(6)
ln(value)	-0.134***			-0.151***		
	(0.00)			(0.00)		
Value in]p20,p40]		-0.027***			-0.029***	
		(0.00)			(0.00)	
Value in]p40,p60]		-0.092***			-0.097***	
		(0.00)			(0.00)	
Value in]p60,p80]		-0.207***			-0.212***	
		(0.00)			(0.00)	
Value in >p80		-0.415***			-0.413***	
		(0.00)			(0.00)	
Value in]20€,40€]			-0.222***			-0.221***
			(0.00)			(0.00)
Value in]40€,60€]			-0.374***			-0.368***
			(0.00)			(0.00)
Value in]60€,80€]			-0.473***			-0.462***
			(0.00)			(0.00)
Value in >80€			-0.534***			-0.521***
			(0.01)			(0.01)
Year 2010	-0.006***	-0.007***	-0.008***	-0.007***	-0.007***	-0.008***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2011	-0.013***	-0.013***	-0.015***	-0.014***	-0.014***	-0.016***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2012	-0.019***	-0.020***	-0.023***	-0.021***	-0.022***	-0.025***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2013	-0.025***	-0.026***	-0.028***	-0.025***	-0.027***	-0.029***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2014	-0.030***	-0.031***	-0.032***	-0.028***	-0.032***	-0.034***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year 2015	-0.034***	-0.035***	-0.037***	-0.032***	-0.036***	-0.038***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Low age (10-39)	-0.120***	-0.118***	-0.119***	-0.089***	-0.092***	-0.095***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
High income (>2250 EUR)	-0.035***	-0.033***	-0.032***	-0.029***	-0.031***	-0.030***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Female	0.019***	0.018***	0.016***	0.020***	0.020***	0.017***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Married/living together	0.029***	0.030***	0.029***	0.035***	0.035***	0.033***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Separated	0.012**	0.011**	0.010**	0.016***	0.017***	0.013**
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
Widowed	0.047***	0.045***	0.044***	0.063***	0.064***	0.061***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Apprenticeship or studied	-0.035***	-0.033***	-0.032***	-0.032***	-0.032***	-0.032***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
No. of persons in HH	-0.011***	-0.009***	-0.009***	-0.005***	-0.008***	-0.009***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	1.233***	1.052***	0.999***			
	(0.01)	(0.01)	(0.01)			
Observations	12,931,206	12,931,206	12,931,206	12,931,206	12,931,206	12,931,206
R^2	0.185	0.203	0.198			
Adjusted R^2	0.185	0.203	0.198			

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table reports estimates from the baseline regressions excluding fixed effects and using OLS (columns 1-3) as well as excluding fixed effects and using Probit (columns 4-6). For Probit estimates in columns 4-6 I report average marginal effects. In all specified columns, the dependent variable is a dummy for cash payment, which is 1 if the transaction was payed with cash and 0 if it was paid with card. Other means of payment are excluded from the analysis. Standard errors are reported in parentheses and clustered at the household level. The base year for time fixed effects is 2009.

Appendix B Descriptive statistics

Table 10: Summary statistics of categorical sociodemographic control variables

Variable	Value	Share of transaction value (%)
Income	<=1500 Euro	27.50
Income	1500-1750 Euro	8.61
Income	1750-2250 Euro	19.53
Income	2250-3000 Euro	20.60
Income	>3000	23.76
Income	Missing	0.00
Age	10-19	0.00
Age	20-39	14.44
Age	40-49	21.14
Age	50-59	23.63
Age	60-69	23.79
Age	>=70	17.00
Gender	Male	27.18
Gender	Female	72.82
Fam. Status	Single	17.03
Fam. Status	Married/Living together	63.40
Fam. Status	Separated	12.99
Fam. Status	Widowed	6.05
Fam. status	Missing	0.53
Education	Apprenticeship or teriary education	64.64
Education	Tertiary education	19.17
Education	No degree	2.37
Education	Not yet in higher education	5.54
Education	Missing	8.28

Notes: Sociodemographic variables refer to the person running the household, i.e. the household head. Income refers to the main income of the household, i.e. not all income sources are included.

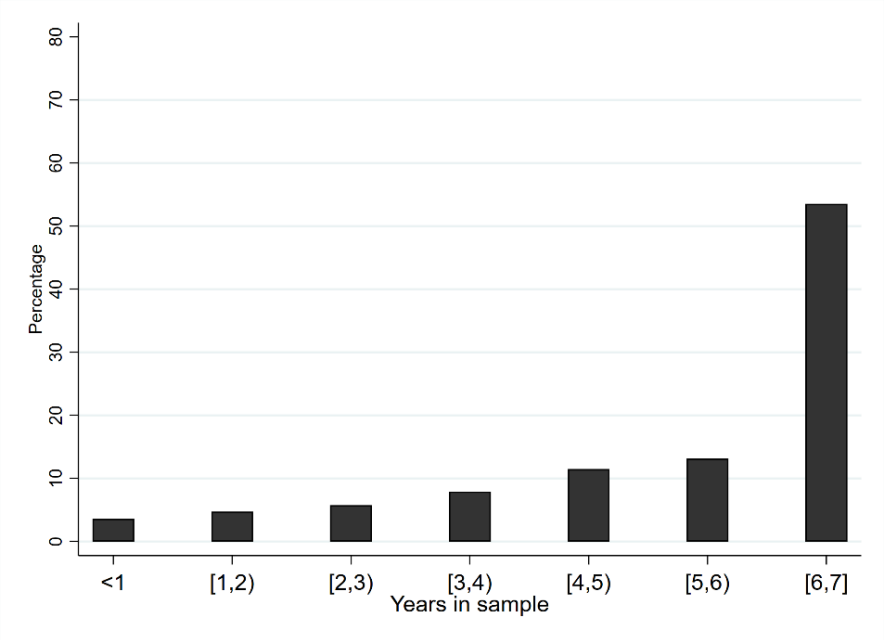

Figure 4: Distribution of years in sample

Notes: The figure shows the distribution of the household duration in sample. The figure highlights that more than 50 percent of the households in sample are observed over the almost entire time span of more than 6 years.