Finance and Economics Discussion Series

Federal Reserve Board, Washington, D.C.
ISSN 1936-2854 (Print)
ISSN 2767-3898 (Online)

Understanding Preferences for Payment Cards using Household Scanner Data

Marc Rysman, Shuang Wang, Krzysztof Wozniak

2025-096

Please cite this paper as:
Rysman, Marc, Shuang Wang, and Krzysztof Wozniak (2025). "Understanding Preferences for Payment Cards using Household Scanner Data," Finance and Economics Discussion Series 2025-096. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2025.096.

Understanding Preferences for Payment Cards using Household Scanner Data *

Marc Rysman
Boston University

Shuang Wang†
Charles River Associates

Krzysztof Wozniak
Federal Reserve Board

September 15, 2025

Abstract

We use consumer panel scanner data to examine households' payment choices, a new application of such data. In particular, we study the long-term shift towards payment cards, as well as the role of transaction size in determining choices. We find that idiosyncratic household preferences are a key driver of payment choice. Our estimates suggest that transaction size, while important, may have a smaller effect on payment choice than previously thought, and that the effect varies substantially across households. Our results further suggest that idiosyncratic household preferences evolve slowly over time, explaining only a third of the increase in card use over the seven-year period in our data. Taken together, our findings have potential policy implications not just for the adoption of new methods such as instant payments, but also around potential costs to households from sun-setting older payment methods such as checks.

1 Introduction

Over the past several decades, the US payments system has shifted from paper payment instruments, namely cash and check, to digital instruments, such as debit cards and credit cards. This shift is important because digital payments are typically regarded as superior in many dimensions: they are faster and cheaper to process, easier for customers to keep track of, and in many ways offer superior protection from crime and fraud. Despite this change, however, cash and check continue to

*Researchers own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. Thanks to Ian Meeker for excellent research assistance.

†Shuang Wang worked on this paper while a PhD student at Boston University.

play a large role in the United States. Anecdotal evidence of young people adopting digital payment while older households persist with cash and check suggests that demographics and heterogeneity between households could be key to explaining the enduring popularity of paper payment instruments. As alternative payment methods multiply and traditional payment methods come under scrutiny as being inefficient and fraud-prone, understanding the determinants of payment method choice is of substantial policy interest.[1]

This paper studies the determinants of payment method choice in both the short and long term. In the short term, across shopping trips, we focus on the transaction size as an important determinant. Transaction size has been central to the discussion of payment choice, with households more likely to pay with non-cash instruments for larger transactions. Previous papers, such as Klee (2008) and Wang and Wolman (2016), have studied the effect of transaction size on payment choice by using scanner data drawn from retailers. However, because these datasets did not allow the authors to track individuals over time, the resulting estimates were not able to separate the within and between effects. In particular, while the previous literature documents that the choice of card is correlated with the size of the transaction, it is possible that this results from households that pay with cards more often having higher transaction sizes on average. In this paper, we use a comprehensive consumer panel dataset to study payment method choice for the first time, which allows us to fully separate the within and between effects.

We also study the long-term evolution of payment method choice. While it is natural to ascribe changes in card use to changes in household preferences, alternative explanations are that there are shifts in the composition of transaction volumes or transaction sizes. For instance, if older households prefer cash and check while younger households prefer cards, gradual growth over time in the number of transactions made by younger households would result in an aggregate increase in card usage even if household preferences for payment methods did not actually change. Naturally, household expiration by older households and household formation by younger households would have a similar effect. Our paper separates out these compositional changes from within-household preference changes in preferences over a seven-year period. We further study the extent to which demographics explain payment choice, and whether those relationships have changed over time. Finally, our dataset exhibits a substantial increase over time in the reporting of payment choice, which we address as well.

This paper leverages a consumer scanner dataset to obtain transaction-level data on payment choice. NielsenIQ maintains a panel of households that tracks in great detail their purchases of food and non-food items for home use across all retail outlets in all U.S. markets (except Alaska and Hawaii). These types of data are common for marketing studies. In addition, NielsenIQ tracks the payment method choice of each trip. We access the data through the Kilts Center of the University

[1] The Federal Reserve lists as one of its five functions: "fosters payment and settlement system safety and efficiency through services to the banking industry and the U.S. government that facilitate U.S.-dollar transactions and payments." See

of Chicago, which recently made the payment choice data available.[2] To our knowledge, no previous academic work has used such data to study payment choice. A recent paper that makes use of that payment information in the NielsenIQ data is Wang (2025), which integrates it into a larger study of the pricing in the payments market.

In order to fully capture the many factors driving payment choice, we estimate a multinomial discrete choice model with household-quarter-choice fixed effects. With over 110,000 households, 28 quarters of data, and 3 payment choices, our richest specification translates into more than 2.4 million fixed effects. Such a setting presents challenges to maximum likelihood estimation. We rely on a new method by Chen, Meeker, Rysman, and Wang (2025) that utilizes the MM algorithm for efficient parameter estimation followed by the jackknife of Dhaene and Jochmans (2015) to address the incidental parameters problem.

While our analysis confirms that transaction size is an important determinant of short-term payment choice, accounting for household heterogeneity suggests the effect is not just smaller in magnitude, it also varies considerably across households. In particular, we find that going from the 1st quartile of the empirical distribution of transaction size, \$11.46, to the 3rd quartile, \$55.40, leads to, on average, a 21.6 percentage point increase in the probability of the payment being made using a card. Notably, we find that our model specification with a full set of household-quarter-choice fixed effects results in a lower effect on average than the model with only choice fixed effects. This finding suggests the impact of transaction size on payment choice is smaller than had been estimated by papers not able to directly account for heterogeneity in unobserved household payment preferences, although the difference is only moderate in magnitude. Moreover, models with household-specific fixed effects unveil substantial heterogeneity in transaction size effects across households. For instance, while the average effect across all households is 21.6 percentage points, the effect varies from as low as 3 percentage points in the 20th percentile of the distribution to more than 30 percentage points in the 80th percentile of the distribution.

The long-term analysis focuses on the increase in card usage share of more than 13 percentage points over the seven-year period in our data. We use our model to decompose the factors driving this change into (a) changes in individual household preferences, (b) changes in the number and value of transactions, and (c) entry and exit of households from the sample. Our results show that only about 40% of the growth in popularity of card payments is due to changes in individual households' preferences. This finding suggests that individual household preferences change relatively slowly. An implication is that potential public policy efforts to shift households to digital payments may take time to yield substantial results.

Overall, our paper makes several contributions. We demonstrate that consumer scanner data can be a powerful tool for studying payment choice. We present new results on the importance of transaction size in determining short-term payment choice, and show that accounting for persistent

[2]The Kilts Center requested payment choice data from NielsenIQ in part based on our request.

unobserved household heterogeneity reduces the magnitude of that effect. We decompose long-term trends in payment choice and find a relatively limited role for changes in household preferences in driving these trends. Finally, we utilize a new method for addressing large numbers of fixed effects in a multinomial logit model.

2 Literature Review

There are many studies whose aim is to identify the determinants of payment choice, with the majority focusing on the decision in the short term. However, few studies are able to track the payments of individual households, especially for payments made with cash. One method for tracking payment choice is to survey consumers retrospectively, as used in Schuh and Stavins (2010) and Koulayev, Rysman, Schuh, and Stavins (2016). These papers rely on a survey that asks consumers about payment use over the previous month. However, because shopping trip details are not captured alongside payment choice, data from such surveys make it difficult to study the determinants of each individual choice, or why choice varies across shopping trips. Another method is to ask survey participants to fill out a diary of payment behavior, as used in Rysman (2007), Arango, Huynh, and Sabetti (2015), and Wakamori and Welte (2017). A well-known example is the Diary of Consumer Payment Choice (Bayeh, Cubides, and O'Brien, 2024). While such diaries are an important data source, Jonker and Kosse (2009) raises questions about their accuracy. In particular, the authors show that the daily number of transactions in seven-day surveys is significantly less than in one-day surveys, suggesting data from payment diaries may suffer from "diary fatigue." Thus, these surveys tend to cover short periods, such as a few days at a time, which is often not enough to allow meaningful analysis of individual payment preferences. A third widely-used method is to obtain data directly from consumer bank accounts, as do White (1975), Dutkowsky and Fusaro (2011), and Stango and Zinman (2014). While data thus obtained do not suffer from diary fatigue, they typically provide no information on cash usage. Moreover, individual consumers may have multiple transaction accounts, some of which may not show up in the available transaction record.

Consumer scanner data have important advantages over these data sources. In particular, in our dataset, we observe payment choice decisions for individual households continuously over a period of seven years, something that no existing diary dataset can match. At the same time, our data have certain limitations. First, the NielsenIQ data do not capture every transaction a household makes. Nonetheless, the dataset is probably most complete with regard to grocery trips, a significant touch-point for payment choice, and an important focus of the payments industry. Second, the method that NielsenIQ uses to track payments does not allow us to distinguish between debit and credit card payments, a common issue in payment literature. Importantly, though, we are able to distinguish between the three most common retail payment instruments: cash, check, and payment card.

A paper closely related to ours is Klee (2008), which also uses scanner data from grocery purchases to study payment choice. However, because the data are drawn from the cash register of a grocery chain, Klee (2008) cannot track consumers over time. Moreover, the data do not contain consumer demographics, so the author accounts for them by using census data for store locations. This contrasts with our paper, where we observe household demographics directly, and importantly, can use household identifiers to account for unobserved heterogeneity using panel techniques such as fixed effects. In addition, our study covers packaged food shopping from a wide array of retail channels, not just a single store. Like us, Klee cannot distinguish between debit and credit, although she distinguishes between signature and PIN-based card transactions. Wang and Wolman (2016) follows a similar approach. Ultimately, most of the papers we discuss here rely on datasets that cover relatively short time periods. We are not aware of another paper that attempts to decompose long-term changes in payment instrument use the way we do. Like us, Wang (2025) utilizes the payments information in the NielsenIQ data. However, the focus of his paper is quite different, as he provides a model of interchange fees, including merchant and consumer update decisions. He does not analyze changes over time in payment usage.

3 Data

In our empirical investigation, we use the NielsenIQ Consumer Panel Dataset, available through the Kilts Center for Marketing at the Chicago Booth School of Business. The dataset provides detailed coverage of purchase choices at the household level, including the quantity bought and price paid for each product. The data include detailed information on the characteristics of products purchased based on UPC codes scanned by the panelists. Crucially for our study, households indicate how they paid for each shopping trip. NielsenIQ verifies the information using receipts submitted by panelists.

3.1 Descriptive statistics

We focus our study on three payment choices: cash, check, and card. Following the approach adopted in previous literature (for example, Klee, 2008), the *card* category pools purchases categorized as either debit or credit card. This approach reflects concerns about panelists not distinguishing accurately between the two types of payment card.[3] Treating credit and debit card purchases as a single category is consistent with how payment cards offer consumers substantially greater

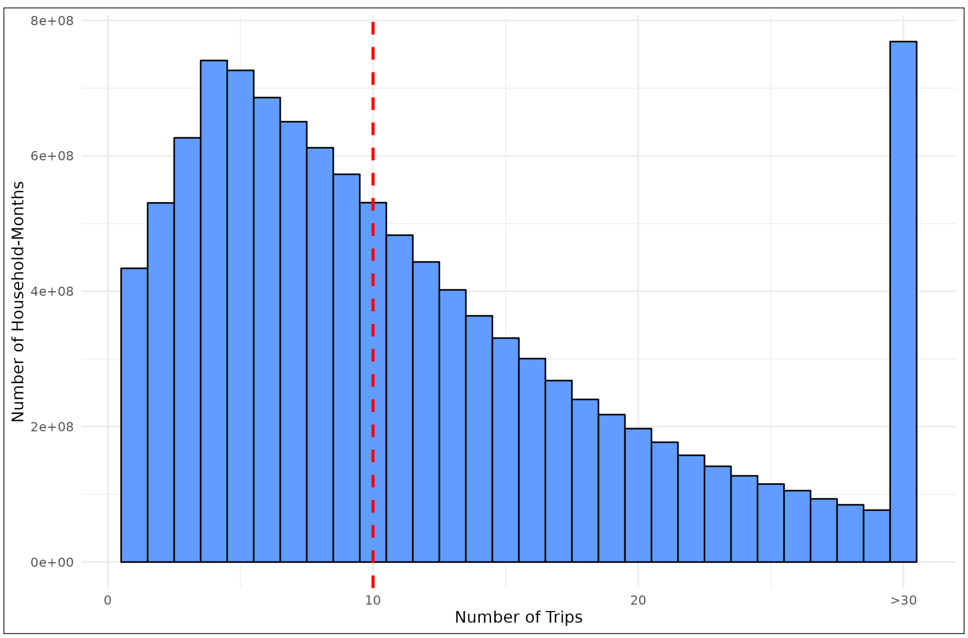
[3]Debit card purchases have typically been authenticated either by signature or PIN (Personal Identification Number, consisting of 4 to 6 digits), while credit card purchases have typically been authenticated by signature. Industry studies and previous literature suggest that many U.S. consumers do not fully appreciate the difference between credit card purchases and those debit card purchases authenticated via signature. Previous versions of the NielsenIQ Consumer Panel Dataset appeared to give households contradictory instructions on this issue, for instance, instructing households to indicate "credit" if they used a signature, and we were not able to fully verify the instructions for the current dataset.

convenience and generate greater payment efficiencies than paper payment instruments, such as check or cash.

Our dataset spans the seven-year time period between January 1, 2013 and December 31, 2019, and captures over 72 million shopping trips made by 117,975 households.[4] Reflecting turnover in the dataset, we observe the average household for just over half the sample period. Nonetheless, there is a considerable number of households that remain in the sample for much longer than the average, with almost a quarter of them remaining in the dataset for the full seven years.

For household-years in which the household is present for the full year, the average number of trips per year is 169.94. The number of shopping trips per month varies substantially across the dataset. Figure 1 shows that the median number of shopping trips per month is 10, equivalent to a shopping trip every three days on average.

Figure 1: Distribution of shopping trips in a month



Note: The red line shows the median. The final bar pools all instances when households reported 30 or more shopping trips in a single month.

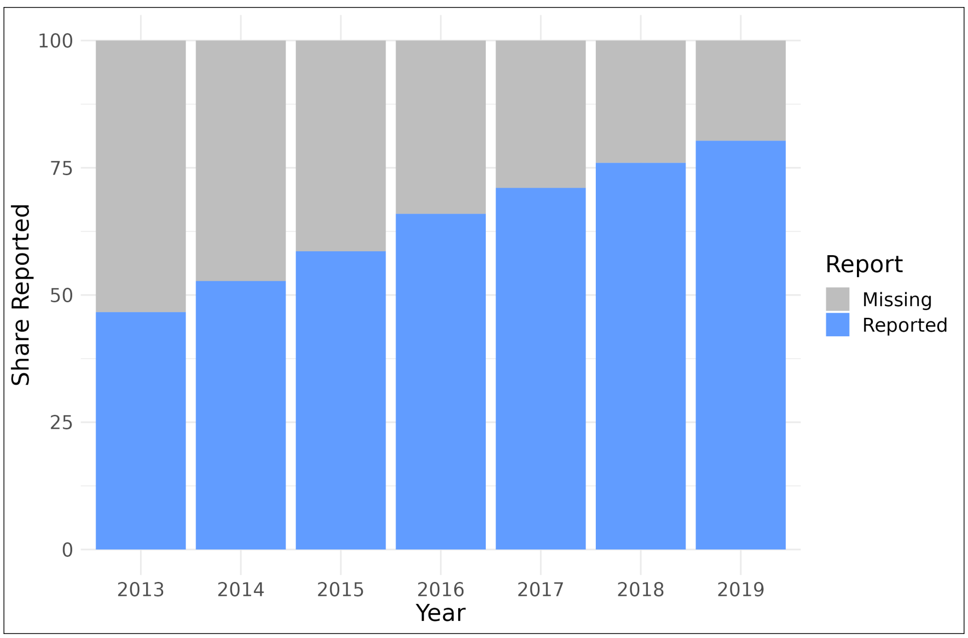
[4]Although the NielsenIQ Consumer Panel Dataset runs over a decade, the version of the data available through Kilts Center for Marketing includes payment choice information beginning only in 2013. We stop at the end of 2019 to avoid the effects of COVID on payments, which we view as outside the scope of our paper. Because Kilts has so far made little post-pandemic data available, we do not engage with that here.

3.2 Missing payments information

For our estimation, we exclude two types of observations from the raw data. They are “Scanner does not collect Method of Payment” and “Other Payment”, which account for 34.9% and 1.68%, respectively, of the trips in the raw data. In this subsection, we consider whether excluding these observations is likely to create issues with our analysis.

Looking at observable shopping trip characteristics, we find little difference in shopping trips based on whether or not the scanner captured the payment method used. For instance, the average transaction value for trips with a payment method reported is \$46.71, compared to \$45.35 for those without. Of potentially more import, reporting of payment method changes substantially over time. Figure 2 shows that the share of payments reported grows from 53% in 2013 to over 80% in 2019. Thus, naive calculations of the growth of card use could fall prey to reporting bias.

Figure 2: Share of shopping trips for which payment is reported over time



Looking deeper, we see that reporting of payment method differs substantially across households. In particular, we find that most households either always report the payment method (62.2% in our sample) or never report (19.2% in our sample). Among the 18.6% of households that report payment choice for some trips but not others, we find that once a household reports, it tends to do so for the remainder of their stay in the dataset. To see this, consider household reporting within a quarter. For household-quarters, only 1.2% exhibit both reporting and non-reporting. In this environment, there appears to be a very limited role for selection from trip to trip in determining whether we observe payment method information. Thus, we do not expect trips to differ substantially based on whether or not they contain payment method information in a way that might affect our analysis.

Nonetheless, it is important to address reporting issues in computing summary statistics such as decompositions of the growth of card use. We address this issue with an inverse probability weighting scheme. In particular, we estimate a probit model predicting whether a shopping trip contains payment choice information or not. The independent variables in the probit regression are household income, household size, the presence of children, race, Hispanic origin, and head of household education. These are variables that NielsenIQ lists as important in generating its own sampling weights. We also include dummy variables for the year-quarter of the shopping trip. Note that the dependent variable is at the level of the shopping trip, whereas the explanatory variables vary only at the level of the household or quarter. This is based on the previous analysis showing the importance of the household and time in determining reporting status.

Let $\hat{p}_{it}$ be the predicted probability of reporting for household i in period t from the probit regression. Let w^s_{it} be the sampling weight provided by the NielsenIQ survey, which varies by year, where each weight w^s_{it} represents the number of similar individuals in the population. Our new weight is:

$$w_{it} = \frac{w^s_{it}}{\hat{p}_{it}}$$

Note that the sampling weights from the NielsenIQ data, w^s_{it}, vary across households and years, whereas our weight w_{it} varies across households and year-quarters.

To see the effect of weighting, we graph the number of transactions by payment type over time. In Figure 3, the upper panel is weighted by NielsenIQ weights, and the lower panel uses our weights w_{it} that are further adjusted for the probability of reporting. The upper panel shows a steady increase in the total number of transactions over time, but this is due to the change in the probability of reporting payment choice over time. Using our weights (the lower panel) shows no trend in payments, although still reveals the increase in the share of card payments over time.

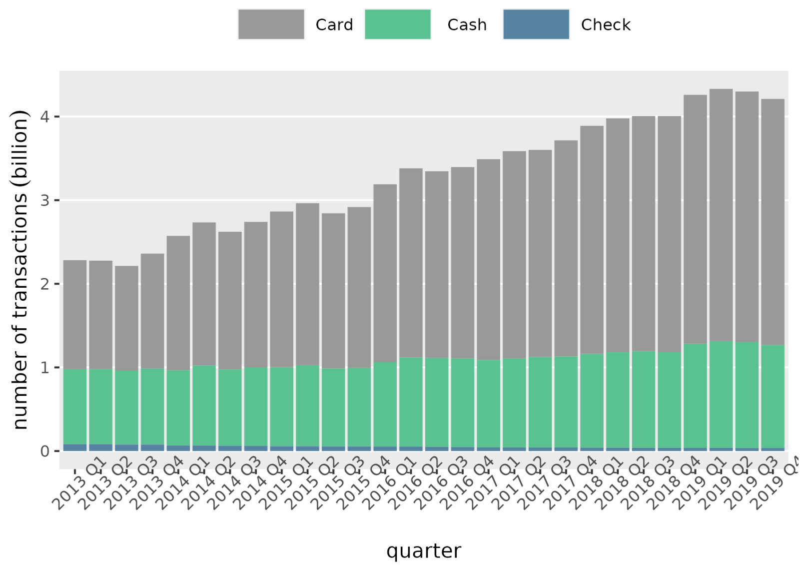
We calculate summary statistics using only observations for which we observe the payment method (46,753,560 observations total), and adjust using our weights w_{it}. The weighting adjustment makes only tiny differences in most statistics, such as market shares, but does have an impact on levels of payments over time, such as in Figure 3. Using our weighting scheme, we find that card accounts for 65% of transactions over our entire sample, whereas cash accounts for 33% and check accounts for 2%. We do not use weights in our regressions.

3.3 Demographics

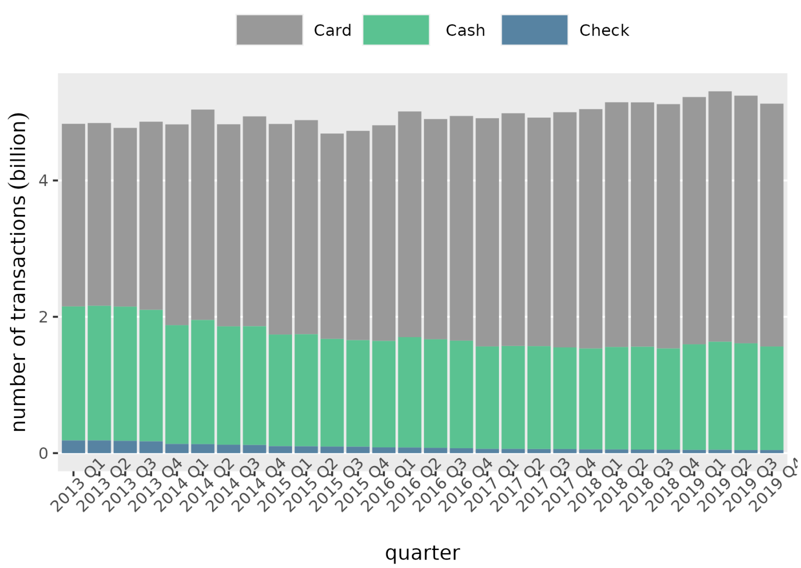
We focus on several demographic variables. We confirm the patterns found in previous research in our dataset. Higher income households tend to pay with card, as shown in Figure 4. Households

Figure 3: Transactions over time

(a) Weighted to address population frequency (w_{it}^s)



(b) Weighted to address population frequency and payment reporting (w_{it})



making \$20K to \$25K report well over 40% of their transactions as cash, whereas those with more than \$100K in income have less than 25% of transactions in cash. Similarly, there is a large change across education levels. Figure 5 shows that households with *some high school* report about 50%

of transactions in cash while those with *post-college degrees* report less than 25%. Check use is somewhat higher among low-education and low-income households, although on a small overall level.

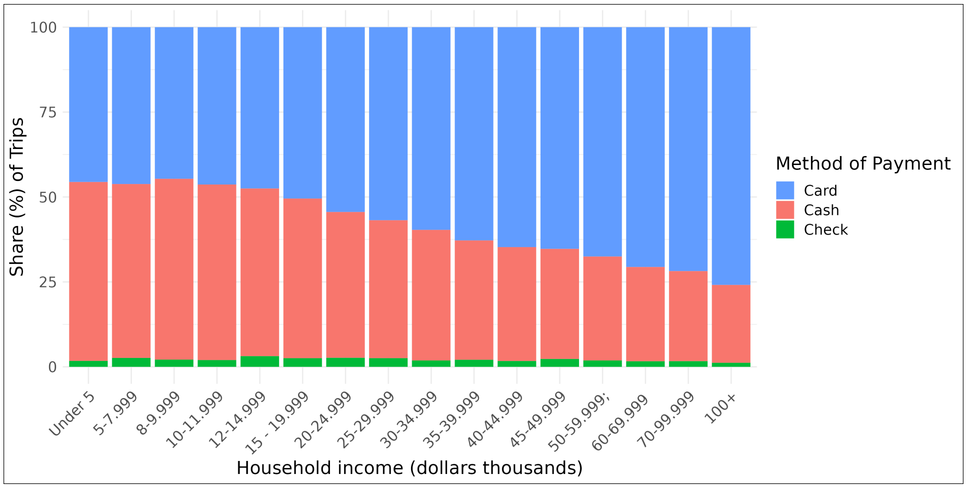


Figure 4: Payment method choice by household income

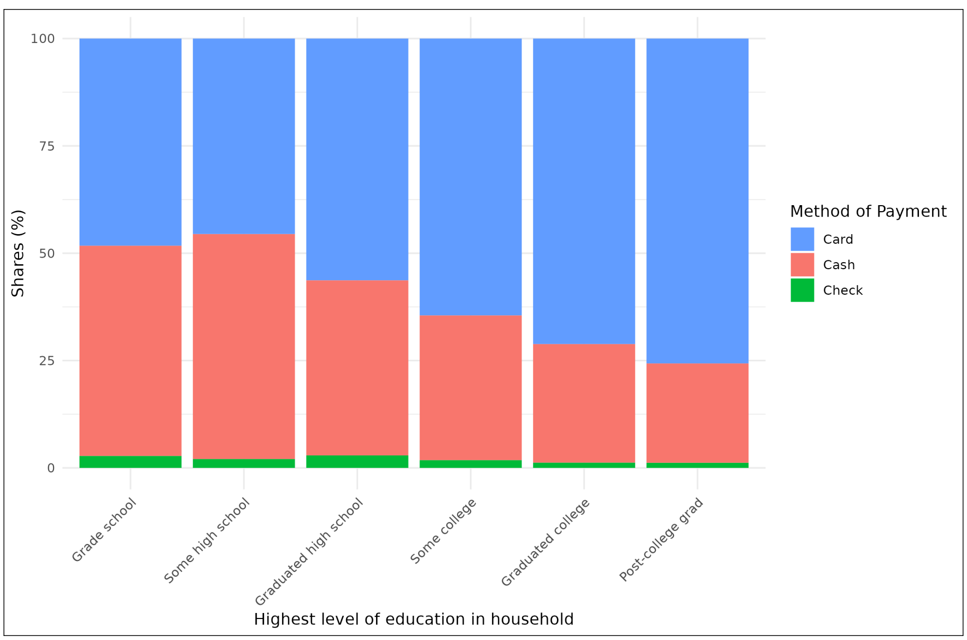


Figure 5: Payment method choice by highest household education level

Common perception is that younger households pay with cards more than older households.

That appears in our dataset, although the effect is not enormous. Figure 6 shows a stacked bar graph of the shares of transactions for each payment choice by 5-year age bins based on the oldest member of the household. The shares of both cash and check grow with age (although under-25s are a slight anomaly). Households of age 25-29 pay with cash and check combined for about 25% of their transactions whereas 65+ households are about 40%.

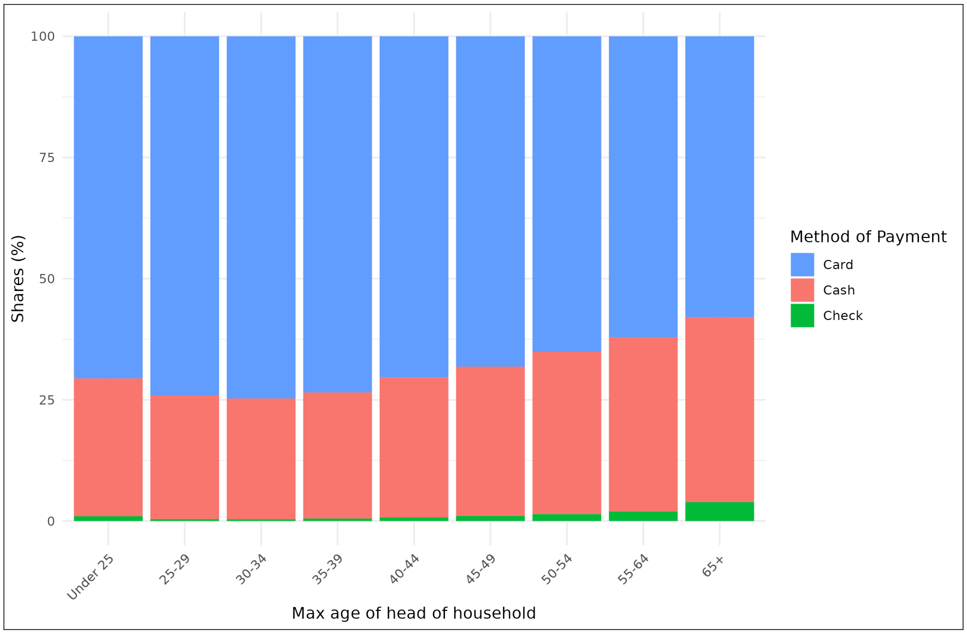


Figure 6: Payment method choice by age of oldest household member

3.4 Single-homing

Single-homing describes households that use the same payment choice for every transaction. We find that less than a third of households do this. More than 30% households put less than 80% of their transactions on their most preferred payment choice. While the market shares for check overall are quite low, we find 30.1% of households pay with check at least once. Thus, there is significant within-household variation in payment choice.

3.5 Transaction size

Following the previous literature, we examine *transaction size* as a key driver of payment choice. Table 1 illustrates the distribution of *transaction size* in our dataset. While the average *transaction size* is $46.08, the variation in *transaction size* is large. In particular, the 10^{th} percentile in the distribution is just \$5.09, the interquartile range goes from \$11.46 to \$55.40, and the 90^{th} percentile is \$105.90.

Table 1: *Transaction size* distribution (\$)

Mean	Std. Err.	10%	[25%, 75%]	90%
46.08	65.56	5.09	[11.46, 55.40]	105.90

Figure 7 illustrates how important *transaction size* is in determining payment choice. In particular, the figure shows that the market share of cash falls from above 60% to below 20% as *transaction size* moves from \$5 to \$150, with most of the remaining share absorbed by card. Similar to card, the market share for check also increases with *transaction size*, although it rises to only around 4% for the largest transactions. It is important to recognize, however, that while check has a low market share overall, this is not because its use is limited to only a small minority of households - in our dataset, 36.3% of the households pay with check at least once.

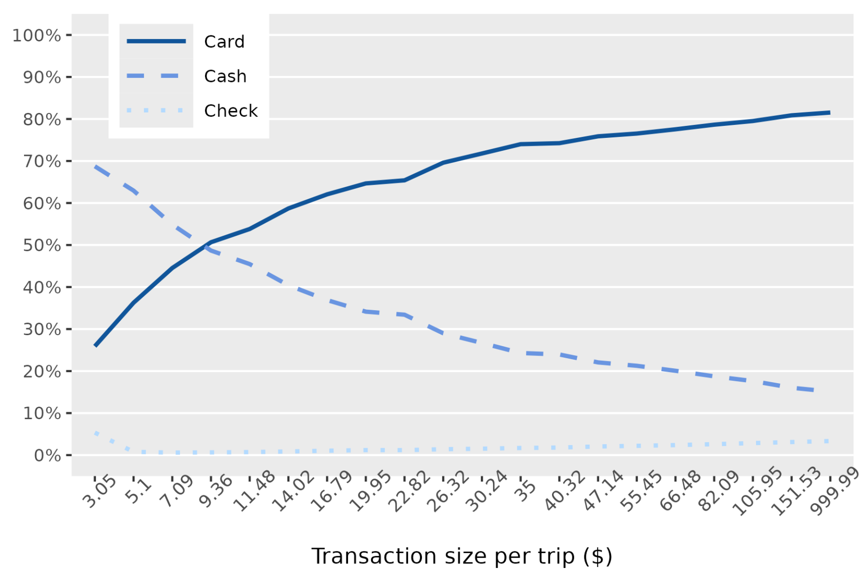


Figure 7: Market share in transactions, by *transaction size*

4 Regression Results

Before turning to our multinomial logit model, we establish some preliminary findings using a linear probability model. While the linear probability model handles fixed effects easily, it can

Table 2: Results from linear probability model

	(1)	(2)	(3)
ln(*transaction size*)	0.128	0.119	0.119
	(0.0001)	(0.0001)	(0.00001)
HH FEs		✓	
HH-year-quarter FEs			✓
R^2	0.109	0.427	0.483

Note: Dependent variable is an indicator for card use for a household-transaction. 46,753,560 observations.

be used to analyze only binary outcomes. Our preliminary analysis focuses on households' choice of making card payments instead of paying using either of the two other options captured in our dataset, cash and check.

4.1 Linear probability model

Let $Y_{it} = 1$ if household i uses a card on transaction t and 0 otherwise. We specify a linear model:

$$Y_{it} = \beta \ln(x_{it}) + \xi_{iq(i,t)} + \varepsilon_{it},$$

where x_{it} is the *transaction size* in dollars. The variables $\xi_{iq(i,t)}$ are household-quarter fixed effects, where $q(i,t)$ is the quarter of transaction t for household i. We also experiment with other specifications for fixed effects, such as fixed effects that vary by household i but are fixed across time. The econometric error term ε_{it} is mean independent from the explanatory variables.

Table 2 shows the results for three model specifications: (1) a base specification with only log *transaction size* as an explanatory variable, (2) including household fixed effects, and (3) including combined household-quarter fixed effects. Going from Column 1 to include household fixed effects (Column 2) increases from 0.11 to 0.43, which suggests the importance of accounting for household fixed effects. The relatively small further increase in R^2 to 0.48 from adding household-quarter fixed effects (column 3) suggests that changes in household preferences over time may play a smaller role in explaining payment choice than variation in payment preferences across households.

The coefficient on *transaction size* is positive and statistically significant in all specifications. Introducing household fixed effects causes the coefficient to decrease in magnitude, which indicates that the within effect is smaller than the between effect. That is, households that pay with cards more also have larger transactions. That means that previous research that relied on cash register data and could not account for household heterogeneity overstated the effect of *transaction size* on card use. The coefficient falls from 0.127 to 0.119 under household fixed effects, so a regression that ignored household fixed effects would overstate the effect by 6.7%.

The dataset includes store identifiers so we could include store fixed effects. That might be useful for addressing heterogeneity in payment acceptance. However, store fixed effects cause the number of observations to drop about in half because there are many stores visited only once. The large dropped sample presumably creates selection issues so we do not pursue this. Note that there are relatively few retailers with restrictions on payment choice among our choices.

4.2 Multinomial logit model

We now turn to specifying the multinomial logit model for use in our analysis of households' payment choice. Doing so allows us to analyze households' choice between all three payment methods simultaneously. Households face an exogenously determined set of shopping trips with predetermined *transaction sizes* for which they must choose a payment instrument. In particular, household i paying with instrument $j \in \{cash, check, card\}$ on shopping trip t receives utility:

$$u_{ijt}(x_{it}, \theta) = \bar{u}_{ijt}(x_{it}, \theta) + \varepsilon_{ijt} = \beta_j \ln(x_{it}) + \xi_{ijq(i,t)} + \varepsilon_{ijt}.$$

Reusing some notation from the previous subsection, $q(i,t)$ is again the quarter when the shopping trip takes place and x_{it} is a scalar representing the *transaction size* in dollars. Here, ε_{ijt} is distributed Extreme Value. As is standard, we normalize the mean utility of one choice to zero. In particular, we normalize the utility of $j = cash$ to zero, so $\beta_{cash} = 0$ and $\xi_{i,cash,q(i,t)} = 0$ for all i and q. We interpret the rest of the coefficients as the value relative to the value for cash. The parameters $\theta = \{\beta_{card}, \beta_{cash}, \{\xi_{ijq}\}_{(i,j,q)}\}$ are to be estimated.

Thus, the probability of choosing j is:

$$P_j(x_{it}, \theta) = \frac{\exp\left(\bar{u}_{ijt}(x_{it}, \theta)\right)}{1 + \exp\left(\bar{u}_{i,card,t}(x_{it}, \theta)\right) + \exp\left(\bar{u}_{i,check,t}(x_{it}, \theta)\right)},$$

where $\bar{u}_{i,cash,t}(x_{it}, \theta) = 0$.

Our specification with household-choice-quarter fixed effects has about 2.4 million fixed effects to estimate, which presents numerical challenges for standard implementations of the multinomial logit in terms of computer memory and time. We follow the approach in Chen et al. (2025), which provides a method to handle this in a computationally efficient way. They develop an implementation of the Minorization-Maximization (MM) Algorithm, a generalization of the Expectation-Maximization (EM) Algorithm, to iteratively linearize an approximation of the model at a given set of parameters, apply linear techniques to address fixed effects (i.e., demeaning) and obtain new parameter estimates, and then update the approximation of the linearization based on the new set of parameters. The algorithm converges to the parameters that maximize the likelihood function, that is, parameters that are numerically identical to parameters from standard gradient search (up to optimization error, which is present in any non-linear search routine). Because parameter estimation is handled with linear techniques, the MM algorithm is faster than methods based on non-linear search, and uses less memory as well.

Another issue with non-linear panel data estimation is the incidental parameters problem. The incidental parameters problem arises when the time dimension of a panel dataset is not large enough, typically understood to be around 30 time periods, to generate unbiased estimates of household fixed effects. In our setting, the "time dimension" is the number of transactions per household, and for most households, this number is relatively high. The average number of transactions per household-quarter is 39.6. However, recall that we wish to identify two fixed effects for each household, one for card relative to cash and one for check relative to cash. Because households pay with their third choice relatively few times, we can still face the incidental parameters problem. In order to address the incidental parameters problem, Chen et al. (2025) recommend bias reduction following Dhaene and Jochmans (2015), which we implement.

4.3 Parameter results

Results of our estimation are shown in Table 3. Standard errors in this table are conventional maximum likelihood standard errors derived from the inverse of the Hessian matrix. As the Hessian is very large, we exploit the sparsity of the matrix in order to invert it, as described in Chen et al. (2025).

Table 3: Results from multinomial logit

	(1)	(2)	(3)
ln(*transaction size*):			
check ($\beta_{\texttt{check}}$)	0.670	0.970	1.060
	(0.001)	(0.002)	(0.002)
card ($\beta_{\texttt{card}}$)	0.677	0.996	1.101
	(0.0003)	(0.001)	(0.001)
Fixed effects:			
check	-4.790	-7.436	-7.616
	(0.004)	[-14.620, 7.501]	[-15.416, 8.730]
card	-1.282	-1.828	-1.632
Household-choice FEs		✓	
Household-quarter-choice FEs			✓
Number of FEs	2	190,672	2,423,332

Notes: Multinomial logit model predicting the choice of cash, card, or check. Standard errors are in parenthesis. For Columns (2) and (3), rather than report standard errors in the fixed effect rows, the table reports the minimum and maximum fixed effects. The number of observations for each regression is 46,753,560.

As expected, we find in all specifications that the estimated coefficients on *transaction size* are positive for both check and card. This agrees with findings in previous papers, as well as the trends presented in Figure 7 – namely, that the likelihood of households paying with check or card increases significantly with *transaction size*. In the first column, which has no controls for household heterogeneity, we see that the fixed effects for check and card are both negative, indicating that at low transaction values, cash is most popular. However, the coefficient on *transaction size* is high enough that, on average, card is preferred to cash at a *transaction size* of \$7 or more. The parameters predict that check is not preferred to cash until *transaction size* is above \$1,200 on average, far outside the range of our data.

Column (2) adds household-choice fixed effects, and column (3) adds household-choice-quarter fixed effects, allowing each household to hold a different preference for each choice in each quarter. The table reports the mean fixed effect for card and check as well as the minimum and maximum values. We can see enormous variation in fixed effects for both specifications (2) and (3), suggesting that there is substantial heterogeneity in payment preferences across households. We further explore this heterogeneity below. The coefficient on *transaction size* grows with the number of fixed effects, which is different from the linear probability model. However, the parameters may not be comparable given the range of fixed effects. To get a better sense of the impact of accounting for unobserved household heterogeneity, we turn to computing average marginal effects.

4.4 Marginal effects and household heterogeneity

To calculate a marginal effect, we calculate the semielasticity of the probability of a choice to *transaction size*. That is, it is the percentage point change in the probability of a choice in response to a percentage change in *transaction size*. Given our multinomial logit assumption, the marginal effect (ME) is:

$$\mathrm{ME}_{ijt} = \frac{\partial p_{ijt}(\boldsymbol{\theta})}{\partial x_{it}} x_{it} = p_{ijt}(\boldsymbol{\theta}) \left(\beta_j - \sum_{k=1}^{3} \beta_k p_{ikt}(\boldsymbol{\theta}) \right). \tag{1}$$

Results appear in Table 4. To summarize, we report the average marginal effect (AME) of *transaction size* on each payment method by averaging Eq. (1) across all households i and trips t. In the first panel, the three columns report the results with no fixed effects, household-choice fixed effects, and household-choice-quarter fixed effects, corresponding to the three columns in Table 3. The second panel reports the average marginal effects after bias correcting with the split-panel jackknife. As expected, bias correction moves the results closer to zero, although the results are similar. We focus on the comparison of Columns 4 and 5 in the second panel to Column 1 in the first panel.

As expected, the estimated coefficient on *transaction size* is positive for check and card and

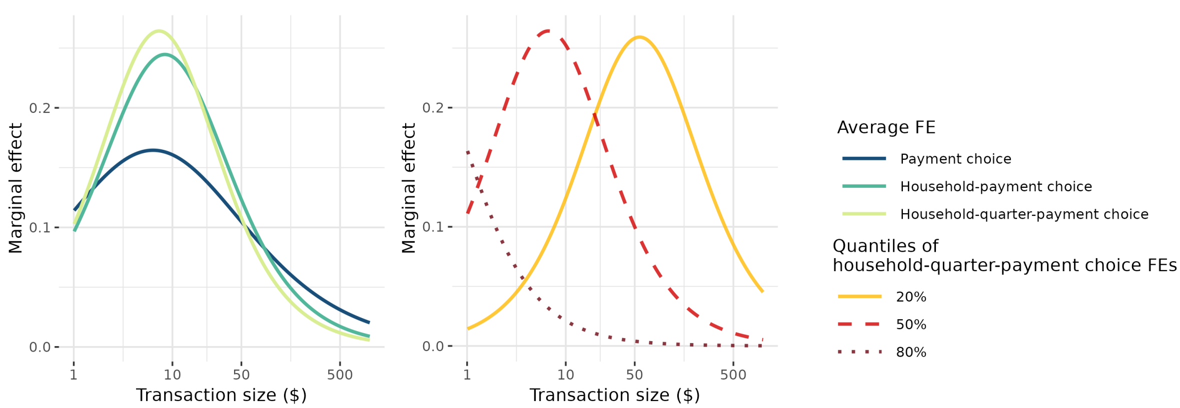
Table 4: Average marginal effects of *transaction size*

	Uncorrected			Bias-Corrected	
	(1)	(2)	(3)	(4)	(5)
Cash	-0.127	-0.119	-0.116	-0.118	-0.115
Check	0.004	0.005	0.005	0.005	0.005
Card	0.124	0.114	0.112	0.113	0.110
Household-choice FEs		✓		✓	
Household-quarter-choice FEs			✓		✓

Notes: The table reports the marginal effect of a change in *transaction size* the probability of the three outcomes: card, cash, and check. Columns 1-3 correspond to Columns 1-3 in Table 3. Columns (4) and (5) report bias-corrected estimates for Columns (2) and (3) to address the incidental parameters problem.

negative for cash. Also, as expected, accounting for household heterogeneity reduces the absolute value of the estimated coefficients. For example, the marginal effect of *transaction size* on the use of cards when there are no household fixed effects is 0.124 whereas the effect when using household-choice-quarter fixed effects is 0.110, about a 11.3% decline. Similarly, the effect on cash is about 9.6% closer to zero.

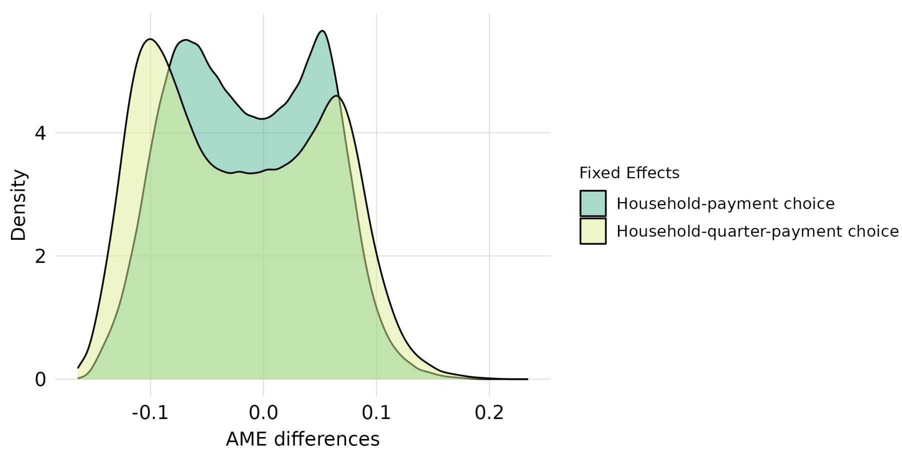
Figure 8: Marginal effect of log of *transaction size* on card usage



In addition to studying average marginal effects, our model also allows us to study household heterogeneity in the marginal effect both across *transaction sizes* and across households. To study how the marginal effect varies with *transaction size*, we plot the marginal effects of *transaction size* on card usage for the three different fixed effect specifications in the left panel of Figure 8. For this figure, we hold fixed effects at the average estimated values. In particular, we find that the two richer household fixed effect specifications generate substantially higher marginal effects for moderate *transaction sizes*. In this sense, accounting for unobserved heterogeneous household preferences suggests that for a broad range of "typical" purchases (for example, the interquartile range reported in Table 1 is \$11.46 to \$55.40), *transaction size* has a potentially much larger impact on households' payment choice than previously thought. Moreover, this result suggests that for small or large purchases, households' payment choice is even less likely to be affected by *transaction size* than previously thought.

The right panel in Figure 8 further examines how the impact of *transaction size* on payment choice varies across households. In particular, the graph illustrates the marginal effect of *transaction size* on card usage for the 20^{th}, 50^{th}, and 80^{th} percentiles of the distributions of the estimated household-quarter-card fixed effects.[5] The difference between households is clear when you consider a transaction in the \$50 region (recall from Table 1 that the mean is \$46.08), where additional spending increases the likelihood of the 20^{th} percentile household paying with card substantially more than it does for the 80^{th} percentile household. Overall, Figure 8 reveals significant heterogeneity in marginal effects under the specifications with expanded fixed effects.

Figure 9: Distribution of difference from the product FE model in marginal effect



Although Figure 8 is informative about how the overall distribution of marginal effects changes with richer models, it does not give a sense of how much it differs for individual households in our dataset, or how wrong we would be about individual households if we used the simpler model. To further explore this heterogeneity, we calculate the difference in AME for each household quarter. We compute the difference between the baseline specification (only choice fixed effects) and the two other specifications with richer fixed effects. Figure 9 presents the distribution of this difference.[6]

[5]The graph holds the check fixed effect constant at the average value; averaging these lines together over the realized fixed effects leads to the household-quarter-choice fixed effect line in the left panel.

[6]To ease the comparison between Figure 8 and Figure 9, the AMEs in Figure 9 are with respect to the log of

We see that differences in AME of 0.1 and -0.1 are common. Comparing this difference to the baseline AME, Figure 8 shows that the AME for the payment choice fixed effect model is always below 0.2, so the changes in AME are relatively large. Also, the dip in the middle of the distributions suggests that there are relatively few households for which the baseline specification is accurate. Thus, bias from leaving out household-quarter fixed effects has a substantial impact on measured individual marginal effects.

Figure 10: Probability of using card vs. *transaction size*

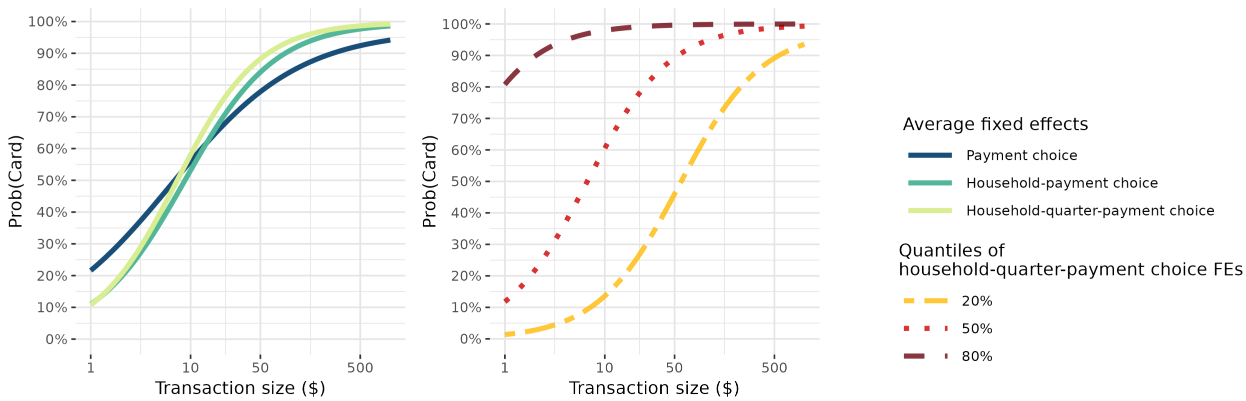


Figure 10 provides another illustration of how accounting for unobserved heterogeneous household payment preferences using a rich set of fixed effects can yield very different predictions. Consider an increase of *transaction size* from \$10 to \$50, which the baseline model predicts would increase the likelihood of a household paying with card by around 23 percentage points. By contrast, the model with household-quarter-payment choice fixed effects not only predicts more accurately that the likelihood will increase by almost 33 percentage points for the average household, it shows that the actual impact will differ greatly between households, ranging from below 3 percentage points (80^{th} percentile household) to over 30 percentage points (20^{th} percentile household).

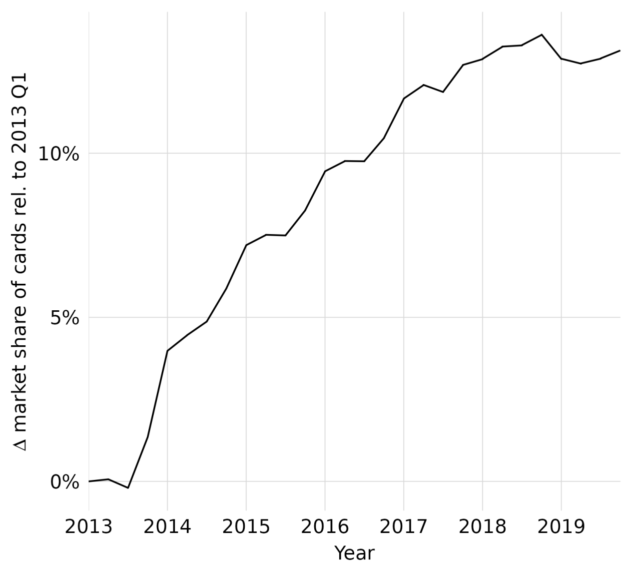
5 Long-term decomposition

The goal of this section is to use our model to calculate the importance of a variety of factors in explaining the growth of card usage over time. Card use has grown 13.1 percentage points in our sample, as described in Figure 11. One of the key factors that could have contributed to this growth is a gradual increase over time in household preferences for card payments. At the same time, changes in the composition of transactions or *transaction sizes* across households could

transaction size.

also have resulted in a shift of payments towards card. Intuitively, consider a young household that always pays using a card and an older household that always uses cash or check. If the young household has children, its average number of shopping trips and average *transaction size* will likely both increase. Similarly, once the older household reaches retirement age, its average number of shopping trips and average *transaction size* will likely both shrink. In this example, the market share of card would increase purely due to changes in the composition of transaction number and size, without any changes in preferences of individual households. Similarly, the older household leaving the sample and being replaced by another young household that favors card over cash/check would result in further growth in card's market share, this time due to entry and exit of households from the sample.

Figure 11: Change in market share of card usage



To facilitate discussion, we introduce new notation. First, we use $\mathcal{I}_q$ to denote the set of households in quarter q and $\mathcal{T}_{iq}$ to denote the set of trips household i took in quarter q. The transactions market share of payment choice j in the quarter Q is thus:

$$s_{jQ} = \frac{1}{\sum_{i \in \mathcal{I}_Q} |\mathcal{T}_{iQ}|} \sum_{i \in \mathcal{I}_Q} \sum_{t \in \mathcal{T}_{iQ}} \frac{\exp\left(\beta_j \ln(x_{it}) + \xi_{ijQ} + \alpha_{m(i,t)}\right)}{\sum_{k=1}^{J} \exp\left(\beta_k \ln(x_{it}) + \xi_{ikQ} + \alpha_{m(i,t)}\right)}.$$

The final market share s_{jQ} may differ from some earlier market share s_{jq} for several reasons: (a) the number of transactions $|\mathcal{T}_{iq}|$ can change, (b) the average size for those transactions x_{it} can change, (c) household preferences $\xi_{ijq(i,t)}$ can change, or (d) the set of households $\mathcal{I}_q$ can change, which can be further broken down into entry and exit. We proceed by sequentially fixing each of these values at their realization in the first quarter each household i is observed in the data, denoted as $\underline{q}(i)$, or in the case of exit the last quarter denoted as $\overline{q}(i)$, and then computing market shares for the last quarter.

Transaction size distribution within households: For each household present in the last quarter, we fix the number of trips and the *transaction size* on each trip at the level of their first quarter, but take their final period fixed effects. We calculate the household-level choice probabilities and then aggregate them to market shares with the number of trips in the current quarter as weights. So the counterfactual last quarter market share is:

$$s_{jQ}^1 = \frac{1}{\sum_{i \in \mathcal{I}_Q} |\mathcal{T}_{iQ}|} \sum_{i \in \mathcal{I}_Q} \frac{|\mathcal{T}_{iQ}|}{|\mathcal{T}_{i\underline{q}(i)}|} \sum_{t \in \mathcal{T}_{i\underline{q}(i)}} \frac{\exp\left(\beta_j \ln(x_{it}) + \xi_{ijQ} + \alpha_M\right)}{\sum_{k=1}^J \exp\left(\beta_k \ln(x_{it}) + \xi_{ikQ} + \alpha_M\right)}. \tag{2}$$

Consider the case in which the set of transactions sizes realized in $\underline{q}(i)$ was the same as in Q. That would imply that the number of transactions in each period was the same, so $|\mathcal{T}_{iQ}| = |\mathcal{T}_{i\underline{q}(i)}|$, and the set of x_{it} was the same for the first and last period that i was in the data. In this case, $s_{jQ} = s_{jQ}^1$. The difference $s_{jQ} - s_{jQ}^1$ provides a measure of how changes in the distribution of transactions contributes to the change in market share $s_{jQ} - s_{j1}$.

Household-quarter-choice fixed effects: We capture the change in preferences within households with our household-quarter-choice fixed effects. In order to mute the effect of changing preferences, we fix household-quarter-choice fixed effects at the level of the first quarter the household is observed and then calculate the market share in the final quarter Q as:

$$s_{jQ}^2 = \frac{1}{\sum_{i \in \mathcal{I}_Q} |\mathcal{T}_{iQ}|} \sum_{i \in \mathcal{I}_Q} \sum_{t \in \mathcal{T}_{iQ}} \frac{\exp\left(\beta_j \ln(x_{it}) + \xi_{ij\underline{q}(i)} + \alpha_{m(i,t)}\right)}{\sum_{k=1}^J \exp\left(\beta_k \ln(x_{it}) + \xi_{ik\underline{q}(i)} + \alpha_{m(i,t)}\right)}. \tag{3}$$

In this case, $s_{jQ} - s_{jQ}^2$ provides a measure of the contribution of changes in household-quarter-choice fixed effects, and this term equals zero only if fixed effects are the same in the first and last period.

Number of transactions across households: As in the earlier young vs. older household example, the growth of card usage in this case could also be due to shifts in transactions from non-card to card users. To isolate this effect, we first calculate the household level choice probabilities, and when aggregating them to compute market share, we weight by the number of trips in the household's first quarter rather than the number of trips in the current quarter. Then, the last-quarter market share becomes:

$$s_{jQ}^3 = \frac{1}{\sum_{i \in \mathcal{I}_Q} |\mathcal{T}_{i\underline{q}(i)}|} \sum_{i \in \mathcal{I}_Q} \frac{|\mathcal{T}_{i\underline{q}(i)}|}{|\mathcal{T}_{iQ}|} \sum_{t \in \mathcal{T}_{iQ}} \frac{\exp\left(\beta_j \ln(x_{it}) + \xi_{ijQ} + \alpha_{m(i,t)}\right)}{\sum_{k=1}^J \exp\left(\beta_k \ln(x_{it}) + \xi_{ikQ} + \alpha_{m(i,t)}\right)}. \tag{4}$$

Entry: In this scenario, we focus on those households that remain in the dataset all the way from the first to the last quarter. These are households such that $i \in \mathcal{I}_1 \bigcup \mathcal{I}_Q$. The market share for these consumers in the final quarter is:

$$s_{jQ}^4 = \frac{1}{\sum_{i \in \mathcal{I}_1 \bigcup \mathcal{I}_Q} |\mathcal{T}_{iQ}|} \sum_{i \in \mathcal{I}_1 \bigcup \mathcal{I}_Q} \sum_{t \in \mathcal{T}_{iQ}} \frac{\exp\left(\beta_j \ln(x_{it}) + \xi_{ijQ} + \alpha_{m(i,t)}\right)}{\sum_{k=1}^{J} \exp\left(\beta_k \ln(x_{it}) + \xi_{ikQ} + \alpha_{m(i,t)}\right)}. \tag{5}$$

Exit: We consider a counterfactual scenario where no households leave the sample. Therefore, all households that ever show up in the sample stay until the last quarter. For those households that leave before the final quarter, we assume that their number of trips, the *transaction size* of each trip and fixed effects in the same in the final quarter as in the last quarter that they are observed in the data, i.e. $\mathcal{T}_{iQ} = \mathcal{T}_{i\overline{q}(i)} \forall i$. Letting N be the number of households in the data, the market share in the final quarter under this scenario is:

$$s_{jQ}^5 = \frac{1}{\sum_{i=1}^{N} |\mathcal{T}_{i\overline{q}(i)}|} \sum_{i=1}^{N} \sum_{t \in \mathcal{T}_{i\overline{q}(i)}} \frac{\exp\left(\beta_j \ln(x_{it}) + \xi_{ij\overline{q}(i)} + \alpha_{m(i,t)}\right)}{\sum_{k=1}^{J} \exp\left(\beta_k \ln(x_{it}) + \xi_{ik\overline{q}(i)} + \alpha_{m(i,t)}\right)}. \tag{6}$$

Then the contribution of each channel is the difference $s_{jQ} - s_{jQ}^k$ for each $k = \{1, 2, 3, 4, 5\}$. Note that the sum of these differences does not exactly equal $s_{jQ} - s_{j1}$, in part because of joint effects. By isolating each effect separately, we do not capture the role of simultaneous changes in channels, for instance, because in practice, ξ_{ijq} and x_{it} change jointly. Still, these differences give a first-order approximation of how much each type of change contributes to the overall change. Therefore, for demonstration purposes, we re-scale these differences so that the sum of them equals to $s_{jQ} - s_{j1}$.[7]

The results of the decomposition are shown in Figure 12. We see that changes in household payment preferences are the largest single factor in the growth of card use. Entry and exit of households also contribute. We find that changes in the number of transactions and changes in *transaction size* contribute negatively. That is, users of cash and check see increased *transaction sizes* and numbers of transactions over the sample, although these effects are smaller in magnitude than the other effects. Without the negative effects, the growth of card's market share would be 18.6 percentage points, rather than the 13.1 that we see in the data. We do not report standard errors for conciseness but they are low following the low standard errors we observe in Table 3. Of the 13.1 percentage point increase, the change in preferences accounts for 9.3 percentage points, 70.8%. But this percentage is amplified by the two negative effects. By taking the absolute value of the effects, we find changing household preferences explain only about a 38.5% of the change in card usage over time.

[7] In this sense, our measure is similar to Variance Partition Coefficients, as in Goldstein, Browne, and Rasbash (2002). See also Grömping (2007).

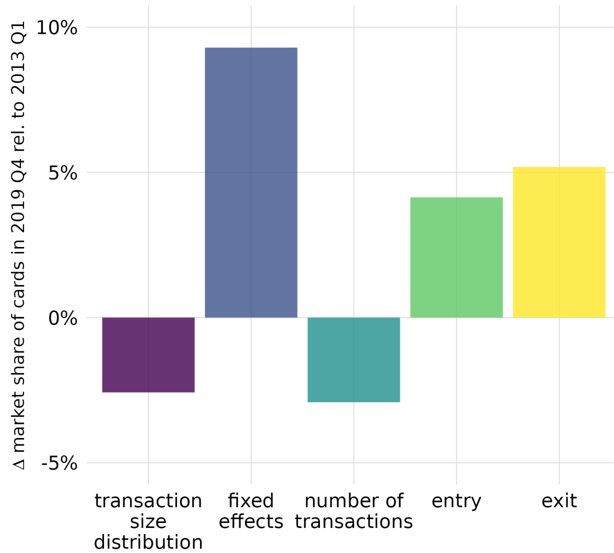


Figure 12: Long-term decomposition

6 Conclusion

Although the transition to digital payments has been one of the most significant developments in the payment industry in recent years, the continued prevalence of cash and check raise important policy questions. This paper studies the determinants of payment choice in the short and long term. Decomposing the drivers of shifts in payment patterns over the long term allows us to contribute significantly to the payment literature, which typically focuses only on short-term payment decisions. Key to this is our ability to capture in our model unobserved household preferences for payments, as well as how they change over time.

In our paper, we use a novel source of data on payment behavior: a transaction-level consumer panel survey. Although the data source is typically used to study household shopping behavior and responses to advertising, we show that these data can be usefully employed to study payment behavior. Doing so allows us to keep track of individual households' payment behavior over multiple years through the lens of high-frequency shopping trip data.

The results of our estimation shine new light both on short and long-term payment decisions. First, our results suggest that while transaction size is an important determinant of payment choice in the short term, its effect is smaller and more heterogeneous than previously estimated in papers not able to directly account for unobserved household payment preferences. Looking to the long term, we use our model to study the key factors driving the increase in card usage observed in our data. We find that while changes in household payment preferences are an important factor, they explain only 38.5% of the observed growth in card usage. Instead, the model finds that other important drivers of long-term changes in payments has been the entry of young households with stronger preferences for card payments, as well as exit of older households with stronger preferences for cash and check payments. Changes in the composition of transactions and transaction sizes have worked to reduce the growth in card usage over the seven-year period ending 2019. Future research may build on our approach to study further evolution of the payments. For example, an analysis of payment patterns during and following the COVID pandemic using our framework could disentangle the long-term effect the pandemic had on households' payment preferences and broad shopping patterns from the short-term changes driven by idiosyncratic circumstances.

References

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

Bayeh, B., E. Cubides, and S. O'Brien (2024). 2024 findings from the Diary of Consumer Payment Choice. Federal Reserve Bank of Atlanta.

Chen, M., I. Meeker, M. Rysman, and S. Wang (2025). An MM algorithm for fixed effects multinomial logit models. Unpublished manuscript, Boston University.

Dhaene, G. and K. Jochmans (2015). Split-panel jackknife estimation of fixed-effect models. *Review of Economic Studies 82*, 991–1030.

Dutkowsky, D. and M. Fusaro (2011). What explains consumption in the very short run? Evidence from checking account data. *Journal of Macroeconomics 33*, 542–552.

Goldstein, H., W. Browne, and J. Rasbash (2002). Partitioning variation in multilevel models. *Understanding Statistics 1*, 223–231.

Grömping, U. (2007). Estimators of relative importance in linear regression based on variance decomposition. *The American Statistician 61*, 139–147.

Jonker, N. and A. Kosse (2009). The impact of survey design on research outcomes: A case study of seven pilots measuring cash usage in the Netherlands. Working Paper 221/2009 Bank of Netherlands.

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

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

Rysman, M. (2007). Empirical analysis of payment card usage. *Journal of Industrial Economics 60*, 1–36.

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

Stango, V. and J. Zinman (2014). Limited and varying consumer attention: Evidence from shocks to the salience of bank overdraft fees. *Review of Financial Studies 27*, 990–1030.

Wakamori, N. and A. Welte (2017). Why do shoppers use cash? Evidence from shopping diary data. *Journal of Money, Credit and Banking 2017*, 115–169.

Wang, L. (2025). Regulating competing payment networks. Unpublished manuscript, Kellogg School of Management.

Wang, Z. and A. F. Wolman (2016). Payment choice and currency use: Insights from two billion retail transactions. *Journal of Monetary Economics 84*, 94–115.

White, K. J. (1975). Consumer choice and use of bank credit cards: A model and cross-section results. *Journal of Consumer Research 2*, 10–18.