Working Paper Series

Payment Choice and the Future of Currency: Insights from Two Billion Retail Transactions

WP 14-09R

Zhu Wang
Federal Reserve Bank of Richmond

Alexander L. Wolman
Federal Reserve Bank of Richmond

THE FEDERAL RESERVE BANK OF RICHMOND
RICHMOND ▪ BALTIMORE ▪ CHARLOTTE

Payment Choice and the Future of Currency: Insights from Two Billion Retail Transactions

Zhu Wang, Alexander L. Wolman*†

October, 2014

Federal Reserve Bank of Richmond Working Paper No. 14-09R

Abstract

This paper uses transaction-level data from a large discount retail chain together with zip-code-level explanatory variables to learn about consumer payment choices across location, time, and size of transaction. With three years of data from thousands of stores across the country, we identify important economic and demographic effects; weekly, monthly, and seasonal cycles in payments; as well as time trends and state fixed effects. We use these estimates to evaluate some implications of theories of money demand and payment choice, and to project future use of currency in retail transactions.

Keywords: Payment choice; Money demand; Consumer behavior

JEL Classification: E41; D12; G2

---

*Research Department, Federal Reserve Bank of Richmond; zhu.wang@rich.frb.org and alexander.wolman@rich.frb.org. The views expressed here are those of the authors and do not represent the views of the Federal Reserve Bank of Richmond, the Board of Governors of the Federal Reserve, or the Federal Reserve System. The data used in this study is proprietary and has been provided to us by the Payment Studies Group at the Federal Reserve Bank of Richmond. We are grateful to the members of that group for their efforts in putting together this data. In addition, we would like to thank members of FRB Richmond's IT department for their critical role in putting the data in manageable form. Sam Marshall and Joe Johnson have provided outstanding research assistance, and Liz Marshall and Sabrina Pellerin provided us with valuable input.

†For helpful comments, we would like to thank Fernando Alvarez, Dave Beck, Itay Goldstein, Tom Holmes, Marc Rysman, Joanna Stavins, Mark Watson and participants in the 2014 Economics of Payments VII conference hosted by the Federal Reserve Bank of Boston, the 2014 Federal Reserve System Applied Microeconomics conference hosted by the Federal Reserve Bank of Minneapolis, the 2014 Econometric Society North American Summer Meeting in Minneapolis, the 2014 International Banking, Economics and Finance Association Meeting in Denver, the 2014 Econometric Society European Meeting in Toulouse, France, and seminars at the Federal Reserve Bank of Richmond and the Federal Reserve Bank of Philadelphia.

1 Introduction

The U.S. payments system has been undergoing fundamental changes in the past few decades, migrating from paper payment instruments, namely cash and check, to faster and more efficient electronic forms, such as debit and credit cards. Amidst these changes, a large empirical literature has developed to study consumer payment choice at the retail point of sale, with the broader goals of understanding payments system functioning and the transactions demand for currency. For researchers and policymakers working on these issues, one major impediment is the lack of data on consumers' use of cash. Given the difficulties of tracking cash use, most studies have relied on data from consumer surveys.[1] The surveys typically provide information about consumers' characteristics, sometimes including their stated perceptions or preferences regarding the attributes of different payment instruments. While this research has improved our understanding of how consumers choose to pay, using consumer survey data has its limitations: Most surveys have relatively small samples (hundreds or thousands of participants at most) and lack broad coverage of location and time.

Our paper helps to fill the gap. We report new evidence on cash use in retail transactions, as well as credit card, debit card, and check use, based on a comprehensive dataset comprising merchant transaction records. The data, provided by a discount retail chain, covers every transaction over a three-year period in each of its thousands of stores across most of the country. In total, we have about 2 billion transactions, which involve a huge number of consumers. If we assume a consumer visits a store once a week, the data would cover more than twelve million consumers; even if we assume daily shopping, it would still cover almost two million consumers. The richness of the data allows us to estimate relationships between location-specific explanatory variables and payment choice, as well as time patterns of payment use associated with day of week, day of month, seasonal cycles and a trend. We use these estimates to evaluate some implications of theories of money demand and payment choice, and to project future use of currency in retail transactions.

A natural reference point for our work is Klee (2008), which also studied consumer payment choices at retail outlets using merchant transaction records. While we are interested in similar questions, there are some important distinctions. First, we look at a different type of store – discount retailer versus grocery store, and a more recent time period – 2010-13 versus 2001. Second, compared with Klee's data, we see richer geographic variation – several thousand zip codes versus 99 census tracts, and richer time variation – more than 1,000 days versus 90 days. We also assemble a larger set of economic and demographic variables, many of them motivated by theory, to help explain consumer payment choices. With this richer dataset, we are able to investigate not only the cross-sectional variation of payment composition, but also the monthly and longer-run time patterns not addressed in Klee's study. In addition, our analytical approach differs from Klee (2008). Because our data set is so large we do not work with the transaction data directly, instead aggregating it up to the shares of transactions for each payment type on each day in each zip code. Aggregation allows us to use all transactions, and it smooths out the "noise" in individual transactions. After providing results

[1] For example, Borzekowski et al. (2008), Borzekowski and Kiser (2008), Zinman (2009), Ching and Hayashi (2010), Arango et al. (2011), Koulayev et al. (2011), Cohen and Rysman (2012), Schuh and Stavins (2012).

for the overall variation in the payment composition across time and locations, we then group the data by transaction size and estimate separate models for each group, thereby allowing all coefficient estimates to vary across transactions of different sizes. This approach is consistent with a general theoretical framework (e.g. Prescott 1987, Lucas and Nicolini 2013), in which consumers each have a threshold transaction size below which they use cash and above which they use a non-cash form of payment that varies across consumers. The interpretation of the thresholds allows for corner cases: For some consumers, the thresholds could be close to zero so they rarely pay with cash; for some others, the thresholds could be prohibitively high so they always rely on cash. The share of each payment instrument for a given transaction size is then determined by the fraction of consumers whose thresholds lead to their use of that particular instrument, and it is a function of the location-specific variables and calendar time. In terms of estimation, we use the fractional multinomial logit model, which specifically handles the fractional multinomial nature of our dependent variables. As a robustness check, we have verified that transaction-level multinomial logit regressions on subsamples of the data yield consistent results (See Appendix C and Appendix D).

The fact that our data comes from a discount retailer means that transaction sizes tend to be small – the median sale value is around \$7. As such, Klee's grocery-store data may be more appropriate for estimating the value-weighted mix of payment instruments that characterizes point-of-sale transactions. However, for the specific purpose of learning about cash use in retail transactions our data is well-suited. Beyond illegal or overseas use of cash, there are two main reasons that the much-hyped "cashless society" has not arrived. First, cash has remained stubbornly popular for use in small-dollar transactions because of its convenience. Second, a nontrivial segment of the population remains unbanked or underbanked, thus without access to the primary alternatives to cash (though alternatives that do not require a bank account, e.g. EBT or prepaid cards, are now becoming more widely available). While our data does not address the underground economy or overseas cash holding, it has the desirable properties for studying cash use that (i) transactions tend to be small, and (ii) the stores are located in relatively low-income zip codes, suggesting that the customer base is more likely to be unbanked or underbanked than the population at large. In sum, although our data overstates the *proportion* of cash use in U.S. retail transactions, this very fact means that it provides valuable insights into the *nature* of cash use, which in turn can be used to forecast future cash use.

Our empirical model is necessarily reduced form, because we are not able to identify customers, only transactions. However, we link the empirical model to theories of money demand and payment choice by assuming that the demographic and economic characteristics of the zip code in which each store is located reflect the demographic characteristics of the store's customers and the economic environment in which they live. In virtually all models of money demand, dating back to Baumol (1952) and Tobin (1956) and including Sidrauski (1967) and Lucas (1982), foregone interest represents a main cost of holding cash, although those early models have only one means of payment. Prescott (1987), Freeman and Kydland (2000) and Lucas and Nicolini (2013) among others, have allowed for multiple means of payment in models where non-cash payments (such as check or payment cards) require a fixed per-transaction cost.[2] This fixed cost then implies a consumer-specific threshold transaction size below which cash is used, with the threshold depending on the consumer's characteristics, as well as the economic environment. Motivated by these theories, we include variables in our empirical model that proxy for the costs of using cash relative to non-cash payment means.

Recently, a different strand of literature has incorporated financial or payment innovations and heterogeneous households into Baumol-Tobin style models that explicitly account for the sequential interplay between payments and cash balances (e.g. Alvarez and Lippi 2009, 2014). In turn, these models have predictions for how cash use varies over time, in relation to the shopper's cash inventory. To the extent that there are systematic relationships between day-of-week or day-of-month and cash inventory, these models then motivate the inclusion of the time effects discussed above in our empirical model. Time effects may also arise because of time variation in the composition of customers.

In contrast to the progress that has been made in modeling payment choice when consumers have access to multiple means of payment, there has been relatively little theoretical work done on the consumer's decision to adopt a new form of payment (Recent exceptions include McAndrews and Wang 2012, Koulayev et al. 2012). Nonetheless, a nontrivial fraction of the U.S. population is unbanked or underbanked and thus does not have easy access to non-cash means of payment.[3] Therefore, we include several zip-code-level variables in the empirical model that are likely to be correlated with consumers' adoption of non-cash payments. Our empirical model also includes demographic variables (education, race, etc.) that may be related to both the choice of how to pay and the choice of whether to adopt non-cash means of payment.

Our empirical model fits the data well and allows us to evaluate implications of the theoretical models introduced above. In terms of zip-code-level variables, we find that banks per capita and the robbery rate are associated with a low fraction of cash transactions, while bank branches per capita go in the opposite direction. These findings are consistent with the theories of money demand and payment choice: More banks per capita entail more banking competition and hence lower banking fees and/or better deposit terms, which increases consumers' opportunity costs of using cash relative to other payment instruments; results for the robbery rate have a similar interpretation. In contrast, conditioning on banks per capita, more bank branches per capita reduce consumers' costs of replenishing cash balances, encouraging more frequent use of cash. We also find that median household income, deposits per capita, population density, education level, and the white and Asian population shares are positively related to non-cash transactions. Presumably, these can be explained in large part by considerations related to the adoption of non-cash payment instruments.

Our findings also reveal significant state and time fixed effects. States with the lowest fractions of cash payments tend to have the highest fraction of debit payments, while states with the lowest debit card use tend to be the top states for credit and check use. Turning to the time effects, there are interesting patterns for

[2] For example, consumers may face certain fees, restrictions or risks of identify theft that are related to using non-cash payment means. Those are typically fixed per-transaction costs regardless of the transaction size.

[3] According to the *2011 FDIC National Survey of Unbanked and Underbanked Households*, 8.2 percent of U.S. households are unbanked and 20.1 percent are underbanked. In total, 29.3 percent of U.S. households do not have a savings account, while about 10 percent do not have a checking account.

day-of-week, day-of-month, and month-of-sample. Over the course of the week, the cash and debit fractions are nearly mirror images of each other: Cash falls and debit rises from Monday through Thursday, then cash rises and debit falls on Friday and Saturday. This pattern is consistent with cash-inventory considerations given that many customers of this retailer likely receive their wages on Friday. Within the month, however, it is credit that comes closer to mirroring cash. Early in the month, the cash share is at its highest. Afterwards, cash falls while credit rises. This pattern is likely driven by customers who have monthly paychecks. Early in the month these customers may be financially unconstrained, and thus spend cash, whereas late in the month they rely more on credit while anticipating the next paycheck. Finally, our month-of-sample dummies identify seasonal cycles and long run trends in the payment mix. In particular, the fractions of cash and check transactions decline steadily, while debit and credit rise over the long run.

We also find that as transaction size increases, in a given zip-code location the fraction of cash payments decreases but those of debit, credit and check increase. This is consistent with the threshold hypothesis mentioned above: At a higher transaction size, there will be a higher fraction of consumers whose thresholds of switching from cash to non-cash payments have been crossed. On the other hand, the cross-location dispersion of the payment mix increases with transaction size. We show this is primarily driven by changes in the coefficients on the zip-code-level variables. As transaction sizes increase, consumers in locations with easier access to non-cash payment options will switch increasingly further away from cash compared to locations that do not have those options.

Our results indicate that the fraction of transactions made with cash fell at a rate of between 1.3 and 3.3 percentage points per year, depending on the size of transactions. Taking into account the size distribution of payments, we project that the cash fraction of transactions will decline by 2.46 percentage points per year. A relatively small portion of this decline can be attributed to forecasted changes in the zip-code-level variables. Our projections can also be used to assess whether the *level* of cash use in retail transactions will increase or decrease. The answer depends on assumptions about the current share of cash in overall transactions and the growth rate of in-person retail sales. However, a plausible scenario has the level of cash use declining over the period of our study and continuing to decline in coming years.

The paper proceeds as follows. In section 2 we describe the transactions data and the zip-code-level explanatory variables. Section 3 presents our empirical model and estimated marginal effects for the overall variation in payment shares across time and location. In Section 4 we turn to the separate models by transaction size, and discuss the sources and implications of payment variation across transaction size. In Section 5 we use the estimated coefficients together with projections of some of the explanatory variables to generate forecasts for the future composition of payments at the retailer, and we discuss the future of currency use more generally. Section 6 concludes and suggests directions for future research.

2 Data

The transactions data is from a large discount retailer with thousands of stores, covering most U.S. states. The stores sell a wide variety of goods in various price ranges, with household consumables such as food and health and beauty aids accounting for a majority of sales. The unit of observation is a transaction, and the time period is April 1, 2010 through March 30, 2013. For each transaction, the data includes means of payment, time, location, and amount. We include only transactions that consist of a sale of goods, with one payment type used, where the payment type is cash, credit card, debit card, or check – the four general-purpose means of payment.[4] The retailer also provides cash-back services, and the purchase components of cash-back transactions are included in our analysis. In contrast, transactions made with special-purpose means of payment such as EBT, coupons and store return cards are excluded. All told, our empirical analysis covers 94% of the total transactions (or 97% of the transactions that use just one payment type) in the sample period. Our summary of the data in this section will refer to all stores; the zip-code-level data introduced below and used in the empirical analysis covers most of those stores' zip codes, but we will need to omit some of the retail outlets from that analysis because the zip-code-level data is unavailable.

2.1 Transactions Data

Figure 1 summarizes the data at the daily level, displaying the fraction of all the transactions accounted for by each payment type. Note that while cash is measured on the left axis, and debit, credit, and check are all measured on the right axis, both axes vary by 0.35 from bottom to top, so fluctuations for each payment type are displayed comparably. The figure shows that cash is the dominant payment instrument at this retailer, followed by debit, credit and check. Over the long term, the fractions of cash and check are trending down, with debit and credit trending up. There seems to be a weekly pattern in both the cash and debit shares, with the two moving in opposite directions. Credit displays a monthly pattern, rising over the course of the month. We will devote more attention to both the time trend and the weekly and monthly patterns below – their presence in the raw data will need to be accommodated by the econometric model.

In Figure 1 we aggregated the data to focus on time variation in means of payment. We turn now to the variation across zip codes. Figure 2 restricts attention to the last full month of the sample, March 2013, aggregates the data by zip code, and displays smoothed estimates of the density functions for fraction of transactions conducted with cash, debit, credit, and check. We use only one month because of the time trend evident in Figure 1. The ranking from Figure 1 is also apparent in Figure 2: Cash is the dominant form of payment, followed by debit, credit, and check. More importantly, there is significant variation across locations in cash and debit use, and to a lesser extent in credit use as well. This variation highlights the need for including location-specific variables in our econometric model.

[4] As in Klee (2008), the transactions we classify as credit card may include some signature debit card transactions. However, the patterns for credit card and debit card transactions in our data are sufficiently different from each other that this measurment issue appears quantitatively unimportant.

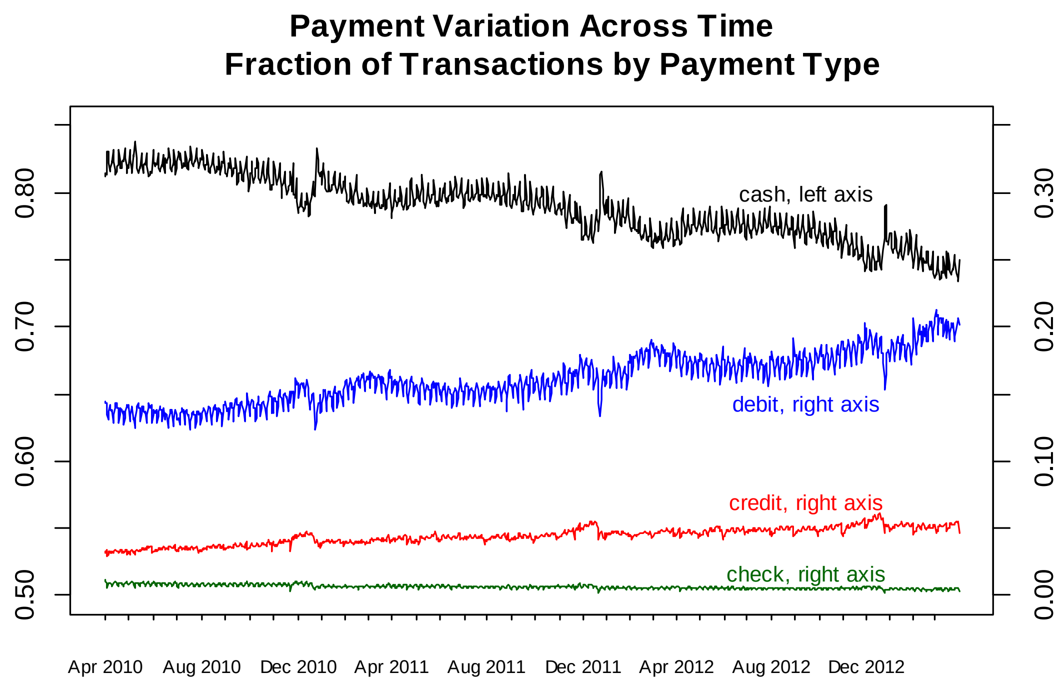

Figure 1.

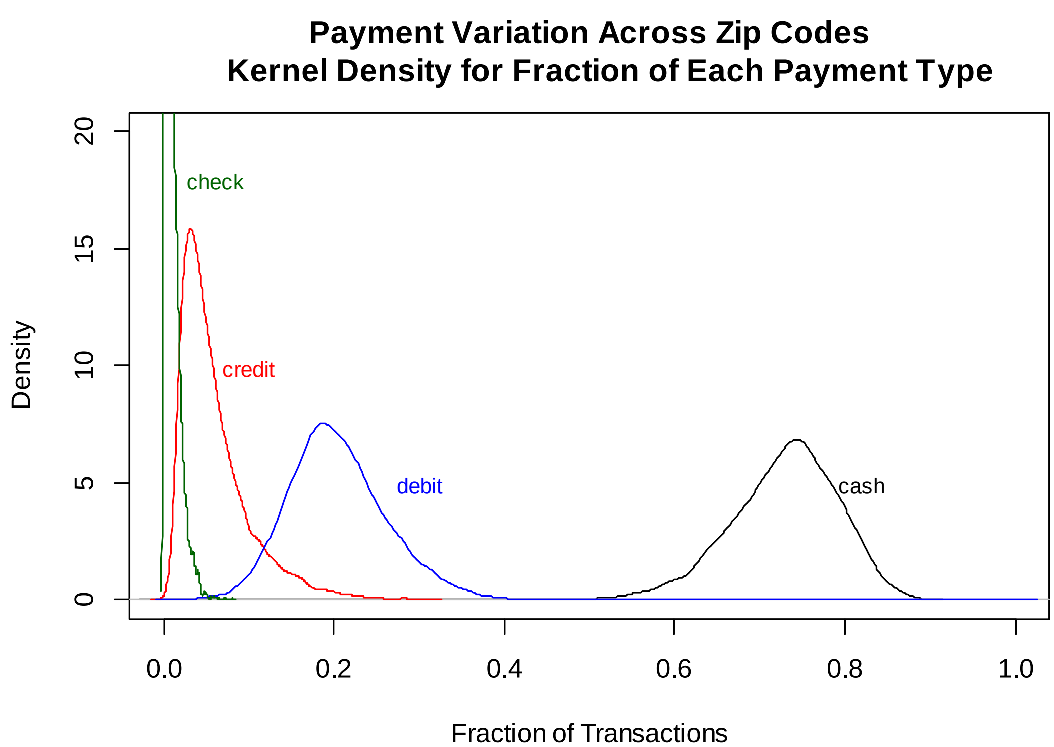

Figure 2.

In Figure 3 we show how the payment mix varies with transaction size, again restricting attention to March 2013. To construct Figure 3, for each zip-code day we group the data by transaction size, using $1 bins between $1 and $15, $5 bins between $15 and $50, and combining all transactions above $50 into one bin. These categories were chosen to ensure a sufficient number of transactions in each bin. For transactions in a given size bin, we calculate the shares of the four payment types on each zip-code day. The solid lines represent the median across zip-code days of the payment shares, and the dashed lines represent the 5th and 95th percentiles of the distribution. The overall message of Figure 3 is that cash is relatively more important for small transactions, whereas non-cash means of payment become relatively more important for large transactions. While hardly surprising, these properties of the data are consistent with the theories that we refer to in the introduction. The top left panel shows that for transactions $1 and below, the median zip-code day had 93 percent of payments made in cash, and, notably, even for transactions in the $50 range the median zip-code day had almost half the payments made in cash. The predominance of cash even for large transactions makes this retailer atypical relative to overall retail sales, suggesting that a significant fraction of this retailer's customer base may not have easy access to other means of payment. However, the prevalence of cash also renders the data especially revealing about the trend in cash use. A final feature of Figure 3 worth noting is that the distribution of payment shares across zip-code days exhibits increasing dispersion for higher transaction sizes, as shown by the fanning out of the 5th and 95th percentiles. We will explore this phenomenon in more detail below.

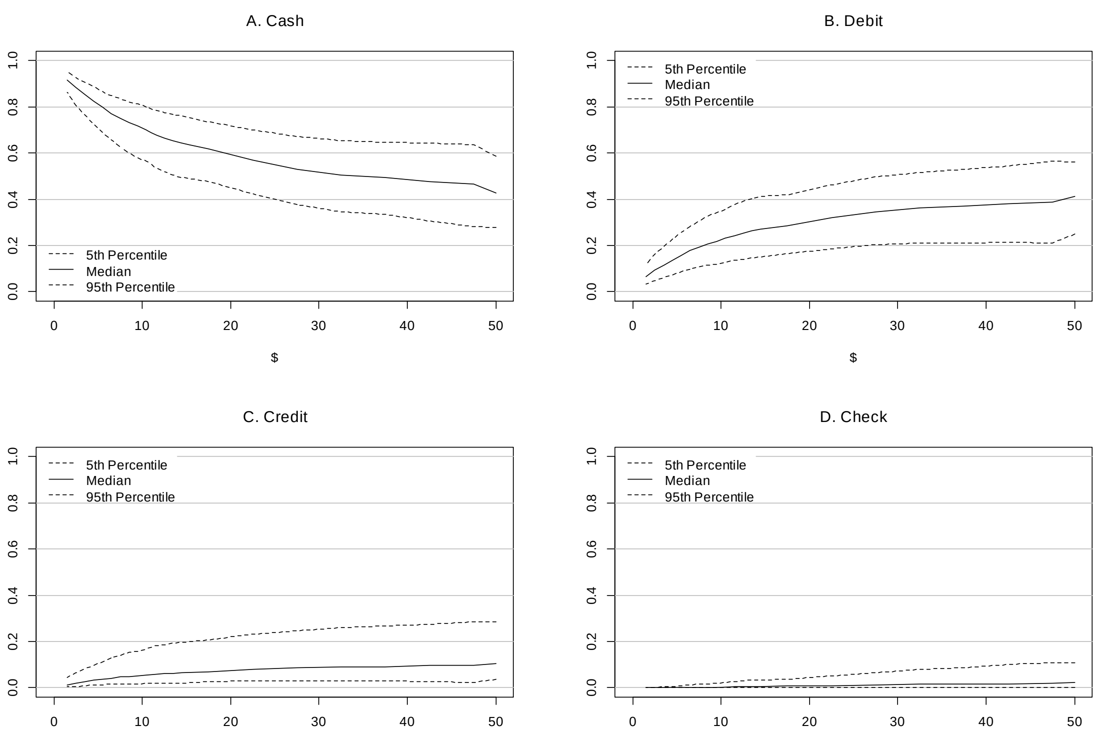

Figure 3. Payment variation across transaction sizes.

Figure 3 shows that means of payment varies systematically with transaction size. Thus, the overall payment mix should be related to the transaction size distribution. Figure 4 provides information about the size distribution of transactions in March 2013, without regard for means of payment. Figure 4A displays a smoothed density function, by sale value, for all 74,465,100 transactions in our sample in March 2013. The prevalence of small transactions helps to explain the large fraction of cash transactions in Figures 1 and 2. Figure 4B plots the distribution of median transaction sizes across zip-code days, also for March 2013 (representing 178,315 zip-code days). Figure 4B complements Figure 2 in showing that there is substantial heterogeneity across location and time with respect to size of transaction, as well as payment mix. Transaction size thus needs to be taken into account in our empirical model(s) of the payment mix.

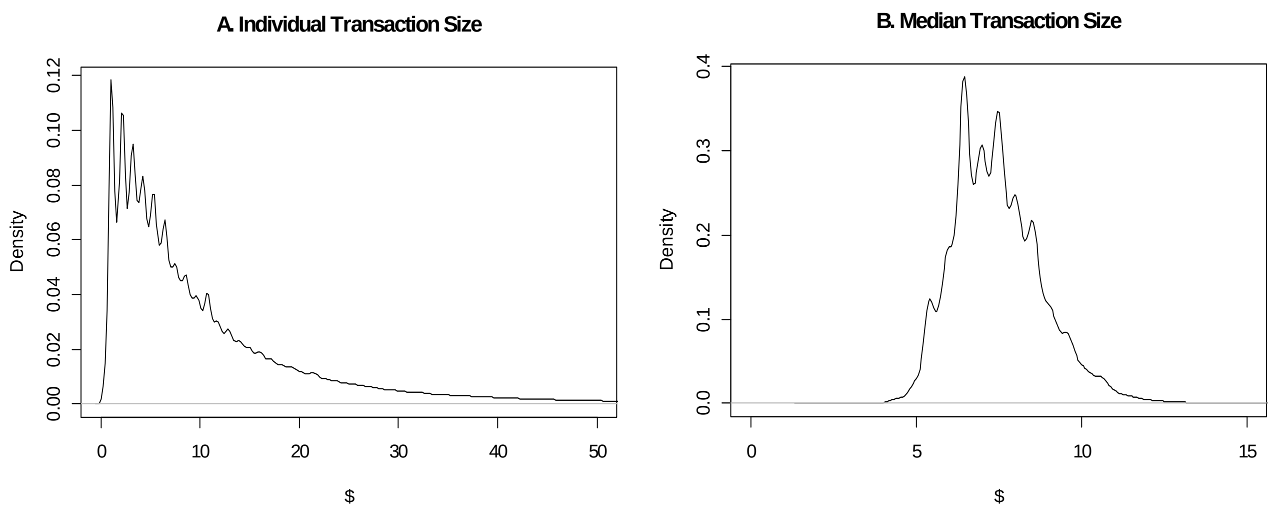

Figure 4. Kernel Densities of transaction size in March 2013.

2.2 Zip-code-level Explanatory Variables

Figures 2 and 3 show a great deal of heterogeneity in the payment mix across zip codes, suggesting the quantitative importance of including location-specific variables in an econometric model of means of payment. We include variables that describe the economic environment consumers face, as well as the characteristics of households. Many of these variables can be linked to theories of cash holding behavior or payment choice, or both. In either case, the variables are relevant for determining the threshold transaction size below which a customer uses cash and above which they use an alternative means of payment. Whether the alternative means of payment is credit, debit or check may depend on the consumer's characteristics, the economic environment, transaction size and calendar time. The distribution of those factors then determines the share of each means of payment in a given zip code. We also include demographic variables that are not explicitly linked to theory. Table 1 lists the zip-code-level explanatory variables we will use in the regressions (fixed at their values of 2011), and contrasts the distribution of those variables in our sample of zip codes to their distribution in the United States as a whole.[5]

2.2.1 Economic environment: cash holding and payment choice

Several of the explanatory variables represent aspects of the economic environment that have direct bearing on the cash holding and payment choice behavior considered in the theoretical literature.[6] These include banks per capita, bank branches per capita, and robbery rate. The robbery rate is measured at the county level, and we discard some zip codes from our transactions data because of missing robbery data. According to the hypotheses referred to in our introduction, the threshold transaction size below which a consumer uses cash should decrease in banks per capita and the robbery rate, but increase in bank branches per capita. Table 1 shows that in our sample of zip codes, the average number of banks and bank branches per capita are less than half their values in the entire U.S., and the differences are highly statistically significant. The robbery rate in our sample is not appreciably different than in the nation as a whole.

2.2.2 Household characteristics: adoption of non-cash payments

Adoption rates of non-cash payments are also important factors in explaining the usage pattern of payment means, and two of the variables we include, median household income and deposits per capita, may be correlated with the likelihood that consumers have bank accounts or own credit or debit cards. There is a clear sense in which adoption represents the extensive margin, compared to the intensive margin associated with cash holding and payment choice. For our purposes however, a consumer who has not adopted any non-cash forms of payment can simply be thought of as having an extremely high threshold transaction size. When we aggregate across the transactions of heterogeneous consumers, the fraction of cash transactions will be increasing in the fraction of non-adopters.

The mean value of median household income is 20 percent lower in our sample than in the U.S. as a whole. Figure 5 delves deeper into the difference in median household income, plotting kernel smoothed density functions for median income in our sample of zip codes and in the United States.[7] Although the modes are similar for the two densities, there is much less mass above the mode in the zip codes where our retail outlets are located. Mean deposits per capita are dramatically lower in our sample than in the entire country, but the nationwide value is driven by a small number of zip codes with extremely large bank branches.

Our classification of variables should not be taken as exclusive; banking competition, prevalence of bank branches, and the robbery rate may also affect household's choices of whether to adopt non-cash forms of payment. Likewise, while we classified deposits and income as "adoption" variables, to the extent that they

[5] Data sources: Most of our zip-code-level variables come from the U.S. Census's American Community Survey and the FDIC's Summary of Deposits. The robbery rate data is from the FBI's Uniform Crime Report.

[6] In thinking about the role of these variables, one is naturally drawn to inventory-theoretic considerations. While inventory theory was the basis for the money demand models of Baumol (1952) and Tobin (1956), most of the work we cite on payment choice takes a reduced-form approach, with an exception being Alvarez and Lippi (2009, 2013).

[7] The red density function in Figure 5 is estimated fairly precisely, as there are several thousand zip codes in our sample.

proxy for the opportunity costs of households' time, they may also fall into the intensive margin category: Households with a high opportunity cost of time face higher costs of replenishing their cash balances, and will therefore use cash less often.

We also include population density as an explanatory variable. As McAndrews and Wang (2012) point out, replacing traditional paper payments with electronic payments requires merchants and consumers to each pay a fixed cost but reduces marginal costs for doing transactions. Therefore, the adoption and usage of electronic payment instruments tend to be higher in areas with a high population density or more transaction activities. The zip codes in our sample are somewhat less densely populated than in the broader U.S.

2.2.3 Demographics

Relative to the United States average, the zip codes in our sample have a low percentage of owner-occupied dwellings, with little variation. The racial composition of these zip codes also differs markedly from the rest of the country: There is a higher percentage of blacks, Hispanics, and Native Americans and a lower percentage of whites and Asians. Also, there is a relatively low percentage of college graduates. However, the age, gender and marriage profiles of our sample are not that different from the nation as a whole.

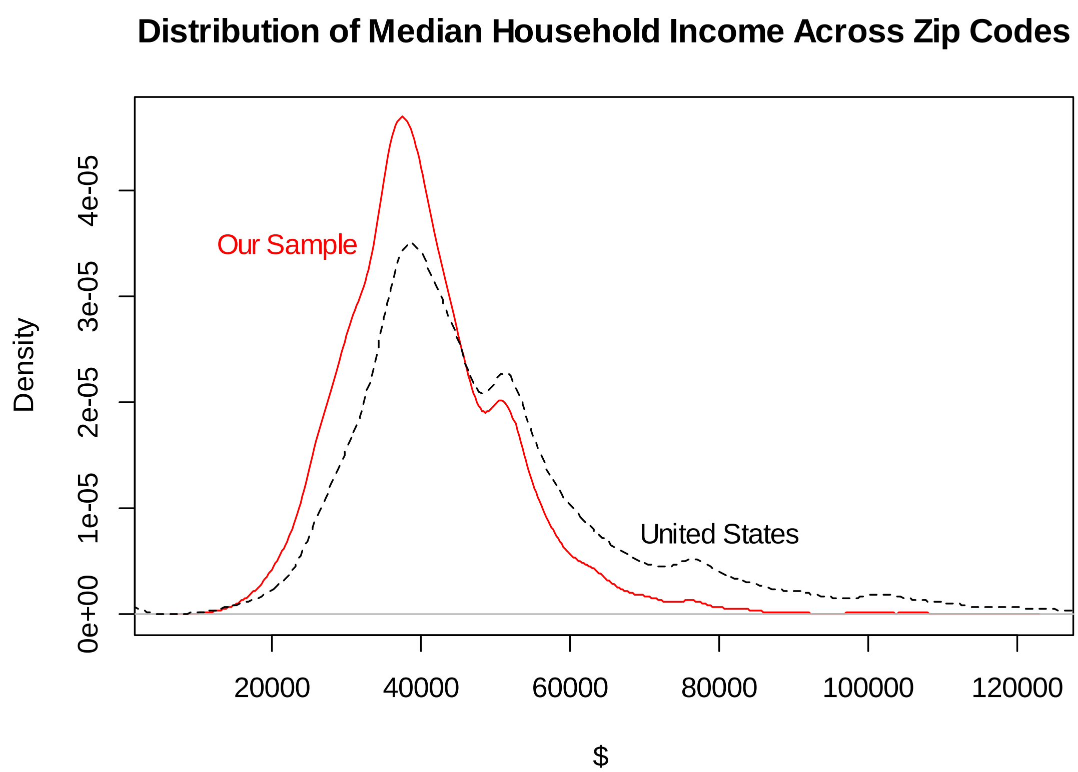

Figure 5.

Table 1. Summary statistics for zip-code-level variables

	Our sample		Entire U.S.	
Variable (unit)	Mean (S.D.)	1% - 99%	Mean (S.D.)	1% - 99%
Cash holding and payment choice				
Banks per capita (%)	0.040 (0.213)	0.0041 - 0.1484	0.091 (0.98)	0.0044 - 0.69
Branches per capita (%)	0.047 (0.214)	0.0045 - 0.1658	0.098 (1.084)	0.0047 - 0.72
Robbery rate ($1/10^5$)	13.88 (29.60)	0 - 179.15	14.12 (29.96)	0 - 179.15
Adoption of non-cash payments				
Median household income ($)	40,623 (11,389)	19,370 - 76,850	50,011 (21,475)	20,001 - 128,961
Deposits per capita ($)	2712 (20,158)	35.09 - 15,765	16,153 (1,205,581)	27.85 - 55,296
Population density (per mile^2)	1436 (2643)	4.2 - 12,021	1782 (5815)	1.8 - 21,159
Demographics (%)				
Family households	66.23 (8.41)	36.47 - 83.52	67.22 (9.93)	28.24 - 85.73
Housing: Renter-occupied	30.18 (11.81)	10.04 - 67.46	26.47 (14.38)	6.21 - 77.63
Owner-occupied	56.67 (12.62)	19.34 - 80.18	60.29 (15.68)	9.86 - 87.28
Vacant	13.14 (8.18)	3.93 - 46.96	13.24 (10.77)	2.81 - 60.23
Female	50.66 (2.58)	39.38 - 55.16	50.21 (2.89)	37.61 - 54.94
Age < 15	19.71 (3.78)	10.07 - 29.45	18.90 (4.15)	6.0 - 29.7
15-34	26.64 (5.93)	15.59 - 48.91	24.98 (7.52)	13.08 - 55.3
35-54	26.28 (2.79)	18.07 - 32.53	27.06 (3.70)	15.47 - 34.94
55-69	17.34 (3.74)	9.13 - 28.35	18.42 (4.47)	7.88 - 31.94
≥ 70	10.03 (3.78)	3.25 - 21.42	10.64 (4.36)	2.27 - 23.93
Race white	73.17 (22.70)	5.24 - 98.29	80.91(20.27)	11.93 - 99.02
black	16.53 (21.26)	0.13 - 90.64	9.09 (16.25)	0 - 79.82
Hispanic	14.12 (19.66)	0.56 - 91.72	10.18 (15.64)	0.3 - 78.69
Native	1.22 (4.53)	0.07 - 17.56	1.08 (4.39)	0 - 16.11
Asian	1.55 (2.43)	0.06 - 12.50	2.73 (5.89)	0 - 31.41
Pac-Islr	0.07 (0.22)	0 - 0.68	0.11 (0.69)	0 - 1.15
other	5.07 (7.03)	0.07 - 32.87	3.76 (6.32)	0 - 31.87
multiple	2.39 (1.31)	0.55 - 6.77	2.32 (1.97)	0.27 - 7.82
Educ below high school	18.16 (8.88)	4.60 - 47.10	15.20 (11.38)	0 - 54.0
high school	34.07 (7.48)	15.30 - 50.90	34.60 (13.18)	0 - 70.6
some college	21.38 (4.41)	10.90 - 31.70	20.91 (8.89)	0 - 49.6
college	26.39 (10.50)	8.70 - 57.70	29.30 (16.71)	0 - 80.4

3 Estimating Payment Shares by Location and Time

In the preceding section we documented substantial variation in the composition of payments across time and location, as well as transaction size. We turn now to an empirical model aimed primarily at explaining the variation across time and location, aggregating transactions by zip-code day. We include median payment size for each zip-code day as an explanatory variable, because the payment mix is sensitive to transaction size (Figure 3), and the distribution of transaction sizes varies across time and location (Figure 4B). The empirical model provides a good fit to the data and the estimated marginal effects are generally consistent with theory. In the next section, we will shift our attention to payment variation across transaction sizes, splitting the data into bins according to size of transaction before aggregating up to the zip-code day level, and running separate regressions for each bin.

3.1 Empirical Model

The data is analyzed using a fractional multinomial logit model (FMLogit). The dependent variables are the fractions of each of the four payment instruments used in transactions at stores in one zip code on one day between April 1, 2010, and March 31, 2013.[8] The explanatory variables comprise the economic and demographic variables listed above, plus time dummies (day of week, day of month, and month of sample) and state-level dummies.

The FMLogit model addresses the multiple fractional nature of the dependent variables, namely that the usage fractions of each payment instrument should remain between 0 and 1, and the fractions need to add up to 1.[9] The FMLogit model is a multivariate generalization of the method proposed by Papke and Wooldridge (1996) for handling univariate fractional response data using quasi maximum likelihood estimation. Mullahy (2010) provides more econometric details.

Formally, consider a random sample of $i = 1, ..., N$ zip code-day observations, each with M outcomes of payment shares. In our context, $M = 4$, which correspond to cash, debit, credit, and check. Letting s_{ik} represent the $k-th$ outcome for observation i, and x_i, $i = 1, ..., N$, be a vector of exogenous covariates. The nature of our data requires that

$$s_{ik} \in [0,1] \qquad k = 1, ..., M;$$

$$\Pr(s_{ik} = 0 \mid x_i) \geq 0 \quad \text{and} \quad \Pr(s_{ik} = 1 \mid x_i) \geq 0;$$

$$\text{and} \quad \sum_{m=1}^{M} s_{im} = 1 \quad \text{for all } i.$$

Given the properties of the data, the FMLogit model provides consistent estimation by enforcing condi-

[8] Most zip codes in our sample have only one store.

[9] Note that when dealing with fractional responses, linear models do not guarantee that their fitted values lie within the unit interval nor that their partial effect estimates for regressors' extreme values are good. The log-odds transformation, $\ln[y/(1-y)]$, is a traditional solution to recognize the bounded nature, but it requires the responses to be strictly inside the unit interval. The approach we take directly models the conditional mean of the responses as an appropriate nonlinear function, so that it can provide a consistent estimator even when the responses take the boundary values.

tions (1) and (2),

$$E[s_k \mid x] = G_k(x;\beta) \in (0,1), \quad k = 1, ..., M; \tag{1}$$

$$\sum_{m=1}^{M} E[s_m \mid x] = 1; \tag{2}$$

and also accommodating conditions (3) and (4),

$$\Pr(s_k = 0 \mid x) \geq 0 \qquad k = 1, ..., M; \tag{3}$$

$$\Pr(s_k = 1 \mid x) \geq 0 \qquad k = 1, ..., M. \tag{4}$$

where $\beta = [\beta_1, ..., \beta_M]$. Specifically, the FMLogit model assumes that the M conditional means have a multinomial logit functional form in linear indexes as

$$E[s_k \mid x] = G_k(x;\beta) = \frac{\exp(x\beta_k)}{\sum_{m=1}^{M} \exp(x\beta_m)}, \qquad k = 1, ..., M. \tag{5}$$

As with the familiar multinomial logit estimator, one needs to normalize $\beta_M = 0$ for identification purposes. Therefore, Eq (5) can be rewritten as

$$G_k(x;\beta) = \frac{\exp(x\beta_k)}{1 + \sum_{m=1}^{M-1} \exp(x\beta_m)}, \qquad k = 1, ..., M-1; \tag{6}$$

and

$$G_M(x;\beta) = \frac{1}{1 + \sum_{m=1}^{M-1} \exp(x\beta_m)}. \tag{7}$$

Finally, one can define a multinomial logit quasi-likelihood function $L(\beta)$ that takes the functional forms (6) and (7), and uses the observed shares $s_{ik} \in [0,1]$ in place of the binary indicator that would otherwise be used by a multinomial logit likelihood function, such that

$$L(\beta) = \prod_{i=1}^{N} \prod_{m=1}^{M} G_m(x_i;\beta)^{s_{im}}. \tag{8}$$

The consistency of the resulting parameter estimates $\hat{\beta}$ then follows from the proof in Gourieroux et al. (1984), which ensures a unique maximizer. In the following analysis, we use Stata code developed by Buis (2008) for estimating the FMLogit model.

Similar to the Multinomial logit (MLogit) model, the FMLogit model imposes some restrictions on the substitution patterns between the categories of the dependent variables. However, our empirical results are robust to alternative ways of grouping the payment options. We show in Appendix B that our findings on cash use remain essentially unchanged when we re-group payment types into two: cash and non-cash (combining debit, credit and check).

3.2 Estimates

We report the estimation results in Table 2. The coefficient estimates are expressed in terms of marginal effects evaluated at the means of the explanatory variables, which we have rescaled to facilitate comparisons of the coefficient estimates.[10]

3.2.1 Cash Holding and Payment Choice Considerations

As suggested by theory, we assume that each consumer has a threshold transaction size (possibly time-varying), below which they only use cash. Aggregating transactions within a zip-code day, we then expect to find that an upward shift in the size distribution of transactions corresponds to a lower share of cash transactions. Using median transaction size as a convenient summary of the size distribution, we find the expected result: Evaluating at the mean of median sale value (at the zip-code-day level) of \$6.86, the marginal effects indicate that a \$1 increase in the median sale value reduces the cash share by 1.7 percentage points but raises debit by 1.2 percentage points, credit by 0.5 percentage points, and checks by 0.05 percentage points. While these results reflect the sensitivity of the payment mix to the distribution of transaction sizes, in Section 4 we use a similar framework to investigate in detail how the payment mix varies across individual transaction sizes.

Our results also confirm the predictions for variables that we classified above as relating to cash holding and payment choice behavior: Recall that a higher opportunity cost of holding cash or a higher cost of replenishing cash balances were predicted to reduce each consumer's threshold transaction size, and therefore reduce the fraction of cash transactions. We find that a higher number of banks per capita corresponds to a lower cash share, mainly replacing it with credit and debit. One additional bank per thousand residents reduces the cash share by 2.3 percentage points, but raises debit's share by 1.3 percentage points and credit by 1.1 percentage points. The robbery rate also significantly reduces overall consumer cash usage. In an area with a higher robbery rate, people tend to use debit cards more frequently. Our estimates show that an increase in the robbery rate of one per thousand residents reduces the cash share by 0.46 percentage points but raises debit by 0.63 percentage points.[11] In contrast, higher bank branches per capita are associated with a higher cash share, mainly at the expense of debit and credit: One additional bank branch per thousand

[10] For continuous variables, the marginal effects are calculated at the means of the independent variables. For dummy variables, the marginal effects are calculated by changing the dummy from zero to one, holding the other variables fixed at their means. Branches per capita is defined as the number of bank branches per 100 residents in a zip code. Median household income is measured in the unit of \$100,000 per household. Banks per capita is defined as the number of banks per 100 residents in a zip code. Deposits per capita is measured in the unit of \$10,000 deposits per resident in a zip code. Population density is measured in the unit of 100,000 residents per square mile in a zip code. Robbery rate is defined as the number of robberies per 100 residents in a county. All the demographic variables are expressed as fractions.

[11] Consistent with our results, Judson and Porter (2004) find that local crime seems to depress overall demand for currency, as measured by payment and receipt growth at 37 Federal Reserve Cash Offices.

residents increases the cash share by 2.4 percentage points but reduces debit by 1.3 percentage points and credit by 1.1 percentage points.

3.2.2 Adoption of Non-cash Payments

For the variables that we classified as relating to the adoption decision, our coefficient estimates also have the expected signs. The fractions of debit and credit card purchases increase with income while the fraction made with cash decreases. The magnitude of these effects imply that for a $10,000 increase in median household income from its mean, the cash share drops by 0.48 percentage points while credit and debit rise by 0.42 percentage points and 0.15 percentage points respectively.[12] Similarly, A $10,000 increase of deposits per capita reduces the cash share by 3.6 percentage points, but it raises debit by 3.5 percentage points and credit by 1.6 percentage points.

On the other hand, we find that higher population density is associated with lower shares of paper payments, especially checks, and higher shares of card payments. This is consistent with McAndrews and Wang's (2012) theory of the scale economies of adopting relatively new payment instruments. An increase of 10,000 population per square mile reduces the check share by 1.4 percentage points and cash by 0.39 percentage points, but it raises debit by 0.90 percentage points and credit by 0.97 percentage points. Although the stores in our sample accept both credit and debit cards, consumers' adoption decisions should be related to the policies of other stores, and those may vary systematically with population density.

3.2.3 Demographics

Previous research using consumer survey and diary studies has found that demographic characteristics such as age, gender, and education play an important role in determining consumer payment choices (e.g. Cohen and Rysman, 2012; Koulayev et al., 2013). Our findings are consistent with that research, but based on a data set with much wider coverage of consumers, locations, and time. We interpret the demographic variables as proxying for consumer characteristics that affect the threshold transaction size below which cash is used, the preferred non-cash means of payment, and the likelihood of adopting non-cash means of payment.

We find that a higher percentage of family households is associated with greater use of card payments in place of paper payments. This again may reflect the scale economies of adopting new payment instruments. Our estimates show that as the fraction of family households increases by 1 percentage point, the cash share falls by 0.093 percentage points and check falls by 0.008 percentage points, while debit rises by 0.09 percentage points and credit rises by 0.013 percentage points.

Comparing with renters, we find that a high percentage of homeowners is associated with greater use of credit and checks, but lower use of cash and debit. However, the magnitude is quite small: A one percentage

[12] The relatively small magnitude could partially reflect the fact that our marginal effects are evaluated at the median sale value $6.86, and consumers tend to favor cash for small dollar transactions. In addition, it may be that the customer base of this retailer varies less across store locations than would be implied by the variation in median income across those locations.

point higher fraction of homeowners is only associated with changes of each payment type in the range of 0.1-0.9 basis points.

In terms of gender differences, we find that a high female population is associated with a high debit share in place of cash. Evaluating at the mean fraction of females, 50.69 percent, the marginal effects indicate that a 1 percentage point increase in the female fraction reduces the cash share by 0.08 percentage points but raises debit by 0.10 percentage points. This could reflect a greater preference for safety by females (which may relate to our earlier discussion of robbery) or a male's preference for anonymity on certain consumption goods (e.g. Klee (2008) argues that certain types of items are more likely to be purchased with cash than with other forms of payments).

Age statistics also are related to the prevalence of different payment types. A higher presence of older age groups is associated with greater use of payment cards relative to the baseline age group, under 15. This might be simply because minors do not have access to non-cash payments, or because families with children tend to use more cash or checks. However, the age profile with respect to cash and checks is non-monotonic. A higher presence of the age group 55-69 is associated with a significantly higher cash fraction, while a higher presence of people at age 70 and older is associated with a higher check fraction. These findings suggest that the age variables may be standing in primarily for cohort effects: Older people tend to be cash users not because they are older but because they did not have access to cards when they first reached adulthood. When we forecast future cash use in Section 5, we will use the cohort interpretation of these estimates, with one exception for the youngest age group.

We also find some interesting racial patterns associated with payment choices. A higher presence of Native American, black, or Hispanic people is associated with a higher cash share relative to the baseline race, white. In contrast, a higher presence of Asian or Pacific Islanders is associated with a lower cash share. However, there are also subtle differences in the substitution patterns: Comparing with white, a high Asian population predicts more credit use in place of cash and checks, whereas a high population of Pacific Islanders predicts debit replacing cash.

Turning to the education results, a more highly educated population (i.e. high school and above) is associated with a lower cash fraction relative to the baseline education group (below high school). The effect is substantial: A one percentage point higher fraction of high-school-and-above population is associated with a 0.20-0.34 percentage point lower cash share. While there are some differences between high school and college groups, they are small compared with the differences from the below-high-school group.

Table 2. Marginal effects for zip-code-level variables

Variable	Cash	Debit	Credit	Check
Cash holding and payment choice				
Median sale value	-0.017* (0.000)	0.012* (0.000)	0.005* (0.000)	0.001* (0.000)
Banks per capita	-0.234* (0.004)	0.128* (0.003)	0.109* (0.002)	-0.002* (0.000)
Branches per capita	0.243* (0.004)	-0.133* (0.003)	-0.113* (0.002)	0.003* (0.000)
Robbery rate	-0.046* (0.001)	0.063* (0.001)	-0.006* (0.000)	-0.011* (0.000)
Adoption of non-cash payments				
Median household income	-0.048* (0.000)	0.015* (0.000)	0.042* (0.000)	-0.009* (0.000)
Deposits per capita	-0.036* (0.001)	0.035* (0.001)	0.016* (0.001)	-0.014* (0.000)
Population density	-0.039* (0.001)	0.090* (0.001)	0.097* (0.001)	-0.148* (0.000)
Demographics				
Family households	-0.093* (0.001)	0.088* (0.001)	0.013* (0.000)	-0.008* (0.000)
Owner-occupied	-0.007* (0.001)	-0.003* (0.000)	0.001* (0.000)	0.009* (0.000)
Vacant housing	-0.019* (0.001)	-0.005* (0.000)	0.017* (0.000)	0.006* (0.000)
Female	-0.080* (0.001)	0.101* (0.001)	0.005* (0.001)	-0.026* (0.000)
Age 15-34	-0.186* (0.002)	0.169* (0.002)	0.035* (0.001)	-0.017* (0.000)
35-54	-0.174* (0.002)	0.134* (0.002)	0.061* (0.001)	-0.022* (0.000)
55-69	0.039* (0.002)	-0.003 (0.002)	-0.014* (0.001)	-0.022* (0.000)
≥ 70	-0.034* (0.002)	-0.030* (0.002)	0.058* (0.001)	0.006* (0.000)
Race black	0.056* (0.000)	-0.026* (0.000)	-0.020* (0.000)	-0.010* (0.000)
Hispanic	0.022* (0.000)	-0.019* (0.000)	0.004* (0.000)	-0.007* (0.000)
Native	0.145* (0.001)	-0.081* (0.001)	-0.059* (0.000)	-0.006* (0.000)
Asian	-0.010* (0.001)	0.000 (0.001)	0.030* (0.001)	-0.020* (0.000)
Pac-Islr	-0.363* (0.011)	0.597* (0.008)	-0.185* (0.007)	-0.050* (0.002)
other	0.088* (0.001)	-0.039* (0.001)	-0.047* (0.000)	-0.002* (0.000)
multiple	-0.123* (0.003)	0.138* (0.003)	0.023* (0.001)	-0.038* (0.000)
Edu high school	-0.202* (0.001)	0.137* (0.001)	0.059* (0.000)	0.006* (0.000)
some college	-0.342* (0.001)	0.246* (0.001)	0.097* (0.000)	-0.001* (0.000)
college	-0.227* (0.001)	0.140* (0.001)	0.081* (0.000)	0.006* (0.000)
Time & State	included	included	included	included
Pseudo R-squared	0.59	0.57	0.59	0.57
Zip-day observations	4,505,642	4,505,642	4,505,642	4,505,642

Robust standard errors in parentheses. *Significant at 1%. Units of regression variables are defined in footnote 10.

3.2.4 State Effects

The estimates for state dummies reveal marked variation in consumer payment choices across states. Figure 6 plots histograms of state dummies for each payment type. Conditioning on the other variables in the regression, the cross-state variation appears largest in the fraction of debit, with a maximum difference of 14.8 percentage points. Credit ranks second with a maximum difference of 9.56 percentage points, and cash ranks third with a maximum difference of 9.52 percentage points. The cross-state variation is smallest for checks with a maximum difference of merely 0.75 percentage points, reflecting the fact that checks only account for 2 percent of all transactions (cf. Figure 1).

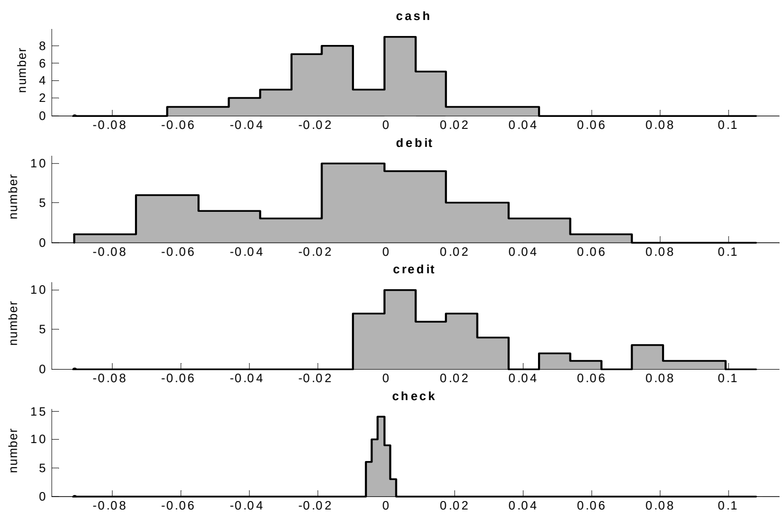

Figure 6. Histograms of state effects.

Table 3. Rankings of state effects

	Cash	Debit	Credit	Check
Top States				
	New Jersey	Arizona	Minnesota	South Dakota
	New York	Idaho	North Dakota	North Dakota
	Michigan	Nevada	South Dakota	Minnesota
	Vermont	New Mexico	Oklahoma	Oklahoma
	Delaware	Florida	Ohio	Colorado
Bottom States				
	Florida	Maryland	Iowa	New Hampshire
	Texas	New York	Arkansas	New York
	New Mexico	North Dakota	Nevada	Arizona
	Idaho	South Dakota	Mississippi	Delaware
	Arizona	Minnesota	New Jersey	New Jersey

The state effects also show interesting substitution patterns between payment types. Table 3 lists the top five and the bottom five states based on the ranking of using each payment type. Conditioning on other variables in the regression, the states that have the smallest fraction of cash use, such as Arizona, Idaho, Florida, New Mexico, turn out to be the top states for debit use. The bottom states for debit use, such as Minnesota, South Dakota, and North Dakota, appear as the top states for credit and check. New Jersey, which ranks the highest in terms of cash use, has the smallest fraction of credit and check. These patterns suggest that there may exist systematic variation in payments systems or regulatory environments at the state level. High cash use may also in part reflect a relatively large underground economy, driven by high tax rates. We investigated this hypothesis by examining the relationship between the cash marginal effects and a measure of state income tax rates. The correlation is positive, around 0.2, providing limited support for the hypothesis.

3.2.5 Time Effects

Figure 1 revealed weekly and monthly cycles in our data, as well as a time trend and what appear to be seasonal cycles. To account for the weekly and monthly patterns, we included day-of-week and day-of-month dummies in our regression. To account for the time trend and any seasonality we also included month-of-sample dummies. Our month-of-sample dummies will pick up regular seasonal variation and idiosyncratic monthly variation as well as any pure time trend. While we cannot perfectly disentangle these three components, with three full years of data it will be possible to begin to identify them. In interpreting each of the sets of time dummies, it will be important to keep in mind that our data do not allow us to distinguish time variation in the behavior of a given set of customers from time variation in the composition of customers.

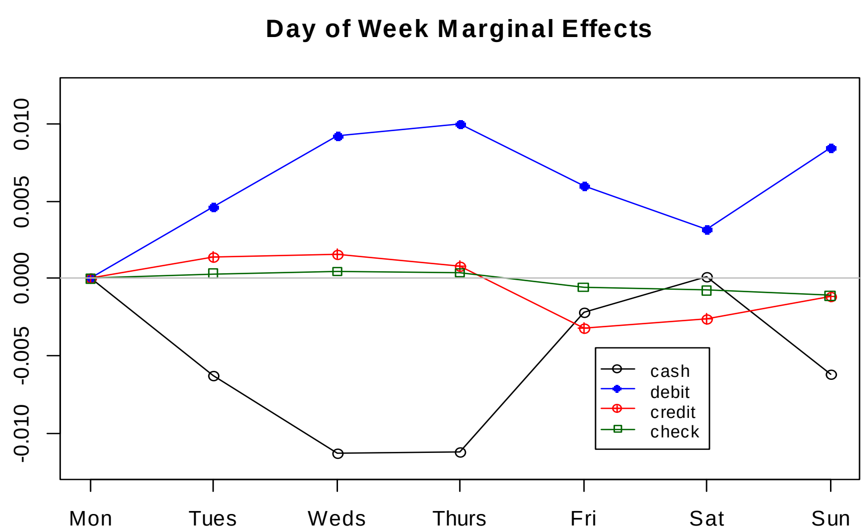

Figure 7.

Figure 7 plots the marginal effects associated with our estimated day-of-week dummies. Just as with the state-level dummies, for each of the time dummies marginal effects will refer to the change in the dependent variable associated with the dummy changing from zero to one, holding all other variables fixed at their means. The cash and debit effects are nearly mirror images of each other: Cash falls and debit rises from Monday through Thursday, then cash rises and debit falls on Friday and Saturday, and the pattern reverses again on Sunday. Although credit displays less variation than cash or debit, there are noticeable movements in credit from Friday through Sunday. From Monday through Thursday, the fall in cash and offsetting rise in debit likely reflects the pattern of cash use predicted by Alvarez and Lippi's (2013) model: Households may visit ATM machines on the weekend and spend cash early in the week when they have it, substituting debit for cash as their cash inventory falls over the week. The spike in cash from Thursday to Friday may reflect the prevalence of Friday as a pay day and a day for ATM visits. Note also that credit actually falls more than debit from Thursday to Friday, suggesting customers are indeed becoming less financially constrained on Friday – consistent with the payday explanation.

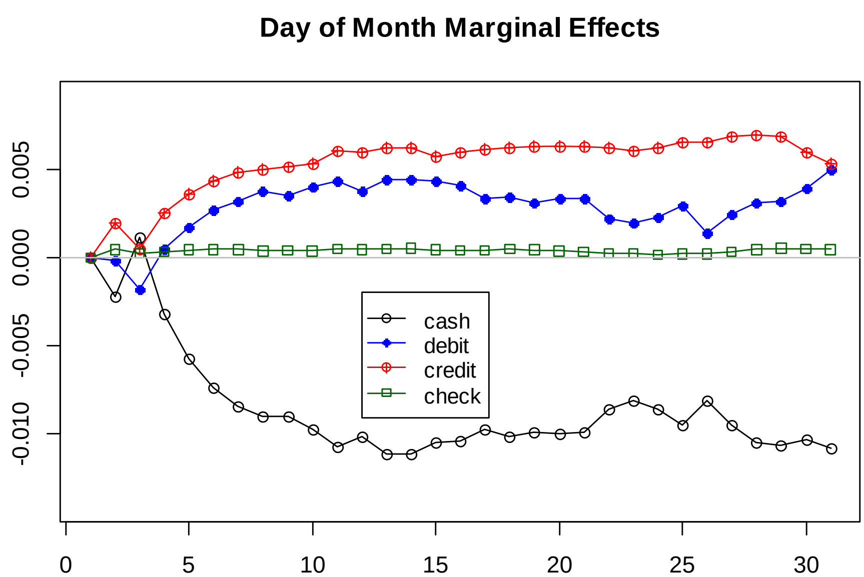

Figure 8.

Figure 8 plots the marginal effects associated with our day-of-month dummies. Whereas most of the "substitution" within the week occurred between cash and debit, within the month the substitution with cash comes from both credit and debit, especially credit. Early in the month, cash is at its highest and credit and debit are at their lowest. Over the month, cash generally falls and credit rises. Debit has a similar pattern to credit, although the variation is smaller. Just as the weekly pattern seemed influenced by paydays, it is also likely that the monthly pattern is driven by customers who have monthly paychecks, for example those who receive certain government benefits. One notable feature of the monthly pattern is a transitory reversal of the broad trends on the 3rd day of the month. In fact, many recipients of Social Security and Supplemental Security Income are usually paid on the 3rd of the month.[13] Early in the month these customers may be financially unconstrained, and thus spend cash, whereas late in the month they rely more on credit while anticipating the next paycheck. It is not clear how the rise in debit early in the month fits with this story, but it may be that even unconstrained customers use the occasion of a monthly paycheck to replenish their cash balances, switching to debit as they draw down their cash over the course of the month. Supporting this conjecture, we find that cash-back transactions peak in the beginning of the month in our data. Finally, the composition of customers likely shifts over the month toward those with access to cards.

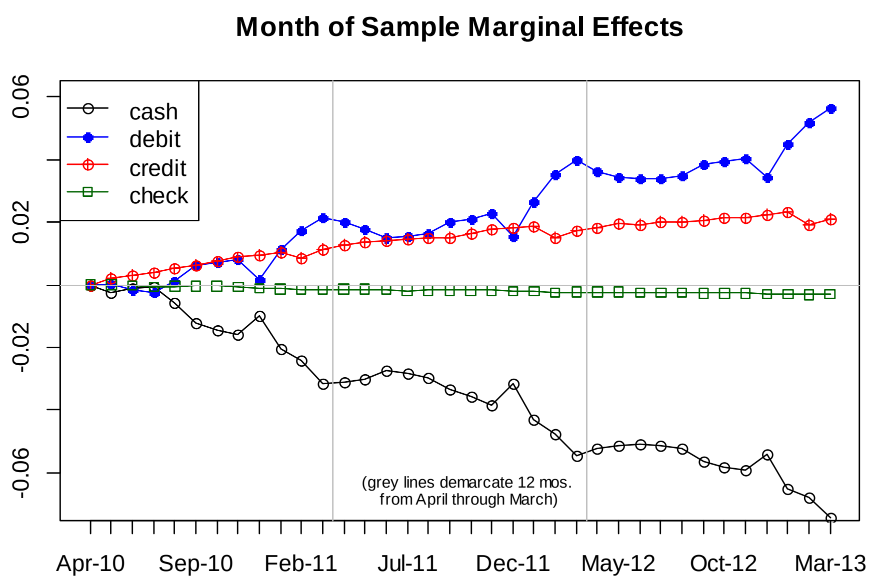

Figure 9.

Figure 9 plots the marginal effects for month-of-sample dummies. As mentioned earlier, these effects combine seasonality with a time trend and idiosyncratic monthly variation. The vertical lines lie between March and April, and thus divide our sample into three 12-month periods. Comparing these periods, both the seasonal and trend are striking, but it is challenging to disentangle them with the naked eye. To separate trend, seasonal, and idiosyncratic components, we regress the four time series plotted in Figure 9 on a linear time trend. The estimated annual time trends are -2.3 percentage points for cash, 1.73 percentage points for debit, and 0.70 percentage points for credit. The four panels in Figure 10 then plot a simple decomposition of the deviations from the time trends into seasonal and idiosyncratic components, for each payment type. The solid lines in these figures represent the average deviation from time trend for each month of the year,

[13] Our rough estimates suggest that more than one million individuals fall into this category.

averaging over the three years in the sample. Actual deviations from trend are represented by the symbols, black for April 2010 through March 2011, red for 2011-12, and blue for 2012-13. While the seasonal patterns contain interesting features – for example, cash and debit are nearly mirror images, with a spike (drop) in cash (debit) in December – note that the overall magnitude of seasonal variation is relatively small: The maximum seasonal effect for any of the payment types is on the order of 1 percentage point. In generating the seasonal effects, we have assumed that the time trends for each payment type are constant over our sample. If this is a good assumption, then the deviations from trend in Figure 10 should be randomly distributed around the seasonal (solid line). There is clearly some serial correlation in the deviations from seasonal, but the only obvious changes in the time trend across years occur for credit. It appears that the growth in credit was higher from April 2011 to March 2012 than in the other two years.

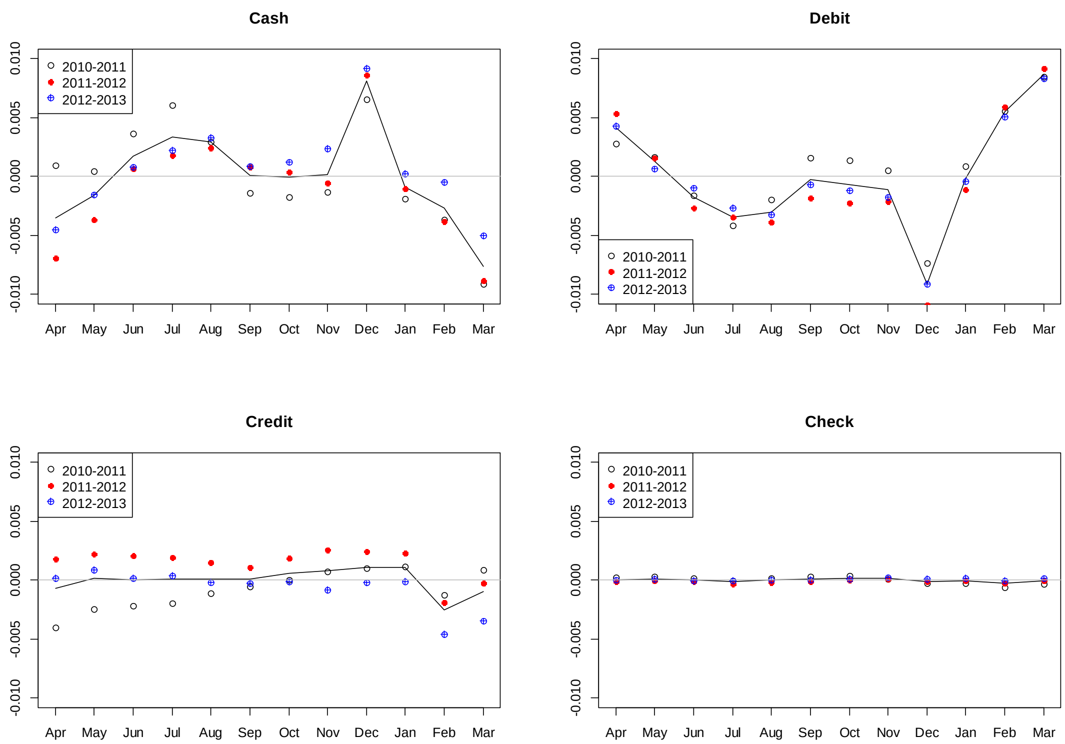

Figure 10. Seasonal and Idiosyncratic Monthly Variation

3.2.6 Further Remarks

Figure 11 displays the distributions of both our model's predicted payment fractions and the actual payment fractions for the entire sample. Together with the pseudo R-squared statistics (calculated as the square of the correlation between the model predicted values and the actual data), which range around 0.57-0.59, Figure 11 indicates that the model does a good job at capturing variation in the composition of payments across time and location. Because our focus thus far has been on payment shares at the zip-code day level, we have implicitly been able to incorporate all transactions into our estimation procedure. With a small subset of the data, it is feasible to estimate a transaction-level multinomial logit regression using nearly identical explanatory variables – we replace median transaction size with individual transaction size, resulting in a specification similar to the one used by Klee (2008). We have verified that such a regression on a random subset of transactions yields similar results to those reported above; see Appendix C for results of a multinomial logit (MLogit) regression on a randomly selected subsample of 4.4 million individual transactions in our three-year data set.

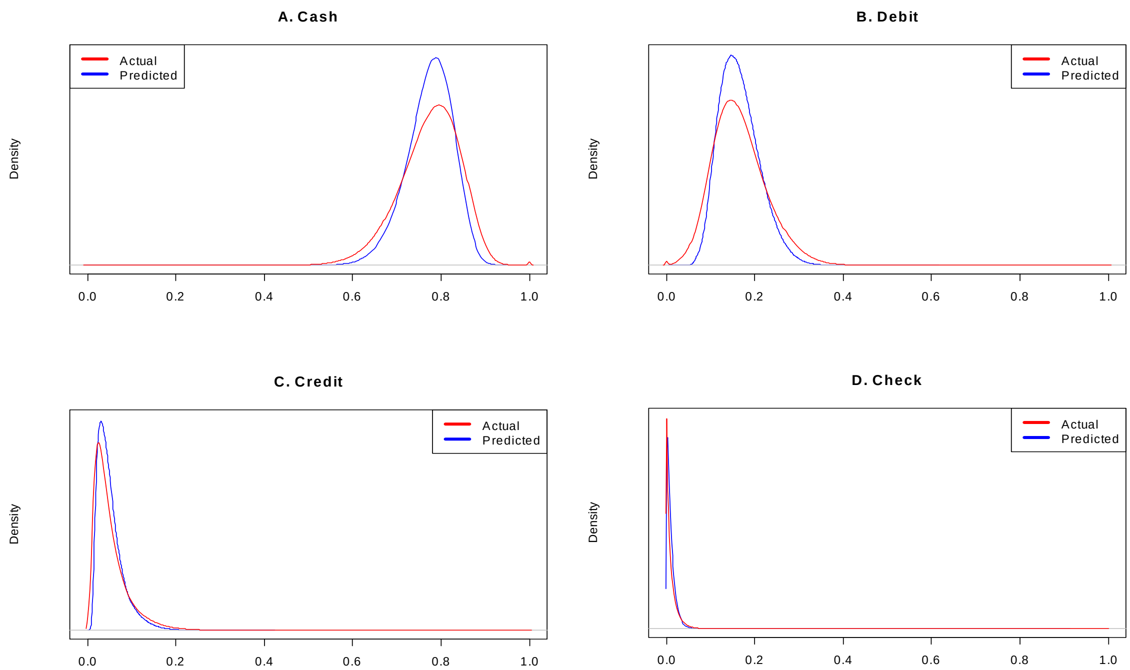

Figure 11. Distribution of actual and predicted payment fractions.

4 Estimating Payment Shares by Transaction Size

Two salient patterns in our data, shown in Figure 3, are that (1) for a given zip-code-day, the share of cash (non-cash) payments decreases (increases) in transaction size, and (2) the dispersion of the payment mix across zip-code days increases in transaction size. To study those patterns, we estimate separate regressions for the 22 transaction size bins used in Figure 3, again aggregating to the zip-code day level. While the approach we took in Section 3 had a role for the size distribution to affect the payment mix, our interest there was in the overall payment mix and we could not use our estimates to produce an estimated counterpart to Figure 3. We will be able to produce that counterpart with the approach in this section. Just as in Section 3, we have verified using a subsample of the data that the payment-share FMLogit regressions yield results consistent with a transaction-level multinomial logit regression, as shown in Appendix D.

4.1 Empirical Specification

We subdivide the sample by transaction size class before aggregating to the day and zip-code level. This allows us to analyze composition of payment mix using FMLogit regressions, just as before, but based on subsamples according to different transaction sizes. In the background, we continue to be motivated by theories in which each consumer has a threshold transaction size below which they use cash. In Section 3, when we aggregated all transactions in a zip-code day, the payment shares were determined by consumer payment choice at each transaction size together with the transaction size distribution, possibly through the explanatory variables we included that capture consumer characteristics, location/time fixed effects, and the median transaction size. Here, by conditioning on transaction size we are essentially looking at "marginal" payment shares instead of the "total" payment shares. With heterogeneous consumers, the payment share at a particular transaction size still depends on the distribution of consumer characteristics, the economic environment and calendar time: The fact that those variables affect a consumer's threshold transaction size means that the payment shares for a given transaction size will depend on the same variables.

An alternative more restrictive version of the approach we take here would impose common coefficients on zip-code level variables across each transaction size regression, allowing only the constant terms to vary. We will see below that these restrictions appear inconsistent with the data. The sensitivity of both level and dispersion of payment shares, shown in Figure 3, is attributed overwhelmingly to variation across transaction size in the coefficients on zip-code-level variables.

For the sake of space, we report only the cash results in this section, leaving the others for Appendix A. We report marginal effects for zip code variables for size classes \$1-\$2, \$5-\$6, \$10-\$11, \$15-\$20, \$25-\$30, \$40-\$45 and above \$50 in Table 4 and in Tables A1–A3 in Appendix A, but Figures 12 and A1-A3 plot the complete marginal effects for all size classes. We highlight the findings from the estimates by transaction size in what follows. As a robustness check, we report in Appendix D the results of a transaction-level multinomial logit regression based on a subsample of the data, specifically all \$6-\$7 transactions in March, 2013, about 3.4 million transactions. In principle the transaction-level regression could yield different results. Even though the explanatory variables are identical for every transaction in a given zip code on a given day, the two regression approaches implicitly use different weights on zip-code days: The payment-share FMLogit regressions weight each zip-code day the same, whereas the transaction-level regressions weight each zip-code day according to the number of transactions. In practice, however, the results differ little across the two approaches.

4.2 Marginal Effects and Amplification

First, most zip-code-level explanatory variables show a sign consistent with our estimates for the overall zip-code-day shares, but the marginal effects amplify significantly as transaction size increases. Comparing Table 2 and Tables 4 and A1-A3 shows that our overall marginal-effect estimates fall between the estimates for \$5-\$6 transactions and \$10-\$11 transactions (recall that the mean value of zip-code-day level median sales is \$6.86). Therefore, the discussion of our marginal-effect estimates for the overall payment mix in Section 3 also applies here for the appropriate size transactions. Moreover, as transaction size increases, the marginal effects for most explanatory variables are increasing in absolute value. For example, comparing cash use between \$1-\$2 transactions and \$40-\$45 transactions, the marginal effects for median household income, deposits per capita, robbery rate, college education, Pacific Islander and Native American rise by a factor of 5.5 to 9.8. Marginal effects for banks per capita, branches per capita, age 35-54, high school education, family household are amplified even more, rising by a factor of 11 to 49. Similar patterns are found for debit, credit and check.[14]

Second, the marginal effects for state dummies show a consistent sign across transaction sizes and amplify as transaction size increases. In Tables 5 and A4-A6, we list the top and the bottom five states based on their marginal effect on using each payment type across transaction sizes. The ranking of states across transaction sizes is generally consistent with the Section 3 results, which suggests that the cross-state differences in payment choices are mainly driven by state fixed effects, rather than state-specific composition of transaction sizes. We also find that the state effects display some amplification as transaction size increases. Taking cash as an example, Figure 13 shows that the maximum cross-state variation is 4 percentage points for \$1-\$2 transactions, but rises to 12-14 percentage points for transactions above \$10. Similar patterns are found for debit, credit and check (see Figures A4-A6 in Appendix A).

Third, the marginal effects of time dummies also show a consistent pattern as before, but tend to be larger in absolute value the larger the transaction size. Comparing Figure 7 to Figure 14 reveals that while the overall cash day-of-week pattern is close to the ones estimated for \$5-\$6 and \$10-\$11 payments, it is slightly different than the patterns estimated for very small and very large payment sizes. The magnitudes of day of week effects are also increasing in transaction size: For transactions in the \$1-\$2 range, the debit marginal effects vary by less than 1 percentage point over the week, whereas that variation is more than 3 percentage points for debit transactions in the \$40-\$45 range (Figure 14). For day-of-month marginal effects, there are also differences across payment size, although the qualitative patterns are common within each payment type. For the most part, the within-month patterns are amplified for larger payment sizes (Figures 15 and A8); this is especially noticeable for cash transactions, where \$40-\$45 transactions have within-month variation of more than 4 percentage points, compared to less than half a percentage point for transactions in the \$5-\$6 range. Turning last to the month-of-sample dummies, these too exhibit interesting variation across the size-specific regressions (Figures 16 and A9). For small-value transactions, the month-of-sample effects are dominated by a stable time trend, whereas the larger transactions display more pronounced seasonal variation. The trends will be discussed further below.

[14] In very few cases, the marginal effects flip signs across transaction sizes. For example, population density shows a negative effect on cash use in small-dollar transactions, but a positive effect in higher-value transactions. However, a careful look into the results in Appendix A show that population density mainly affects the substitution between cards and checks, while cash only captures a small residual effect. In fact, the marginal effects of population density have a consistent sign across all transaction sizes for debit, credit and check.

Table 4. Cash: marginal effects by transaction size

Variable	$1-$2	$5-$6	$10-$11	$15-$20	$25-$30	$40-$45	above $50
Cash holding and payment choice							
Banks per capita	-0.029*	-0.143*	-0.289*	-0.380*	-0.476*	-0.567*	-0.582*
Branches per capita	0.032*	0.151*	0.300*	0.393*	0.491*	0.583*	0.597*
Robbery rate	-0.011*	-0.044*	-0.076*	-0.094*	-0.105*	-0.108*	-0.114*
Adoption of non-cash payments							
Median household income	-0.017*	-0.039*	-0.060*	-0.072*	-0.098*	-0.119*	-0.164*
Deposits per capita	-0.008*	-0.035*	-0.061*	-0.051*	-0.058*	-0.075*	-0.093*
Population density	-0.061*	-0.085*	-0.085*	-0.054*	-0.017*	0.000	0.035*
Demographics							
Family households	-0.005*	-0.080*	-0.141*	-0.184*	-0.216*	-0.247*	-0.236*
Owner-occupied	0.009*	0.008*	-0.009*	-0.029*	-0.045*	-0.049*	-0.062*
Vacant housing	0.008*	0.000	-0.026*	-0.054*	-0.078*	-0.101*	-0.117*
Female	-0.061*	-0.089*	-0.092*	-0.079*	-0.065*	-0.096*	0.000
Age 15-34	-0.038*	-0.155*	-0.250*	-0.310*	-0.361*	-0.431*	-0.403*
35-54	-0.003	-0.128*	-0.264*	-0.352*	-0.430*	-0.558*	-0.512*
55-69	0.084*	0.077*	0.015*	-0.056*	-0.132*	-0.213*	-0.216*
≥ 70	0.051*	0.021*	-0.066*	-0.137*	-0.210*	-0.292*	-0.289*
Race black	0.003*	0.049*	0.079*	0.098*	0.106*	0.119*	0.122*
Hispanic	-0.001*	0.012*	0.025*	0.043*	0.059*	0.076*	0.088*
Native	0.037*	0.120*	0.162*	0.190*	0.219*	0.244*	0.256*
Asian	-0.019*	-0.034*	-0.013*	0.029*	0.042*	0.054*	0.073*
Pac-Islr	-0.118*	-0.338*	-0.448*	-0.440*	-0.547*	-0.648*	-0.918*
other	0.029*	0.076*	0.119*	0.140*	0.127*	0.067*	0.028*
multiple	-0.132*	-0.164*	-0.062*	0.037*	0.124*	0.291*	0.391*
Edu high school	-0.018*	-0.162*	-0.269*	-0.332*	-0.380*	-0.401*	-0.384*
some college	-0.088*	-0.304*	-0.437*	-0.506*	-0.554*	-0.581*	-0.546*
college	-0.045*	-0.199*	-0.293*	-0.344*	-0.374*	-0.372*	-0.356*
Time & state dummies	included	included	included	included	included	included	included
Pseudo R-squared	0.10	0.15	0.11	0.22	0.11	0.05	0.10
Zip code-days (1,000)	4,505	4,505	4,498	4,505	4,483	4,045	4,405
Transactions (1,000)	198,700	129,299	67,465	132,108	50,800	16,425	37,905

*Significant at 1%. Units of regression variables are defined in footnote 10.

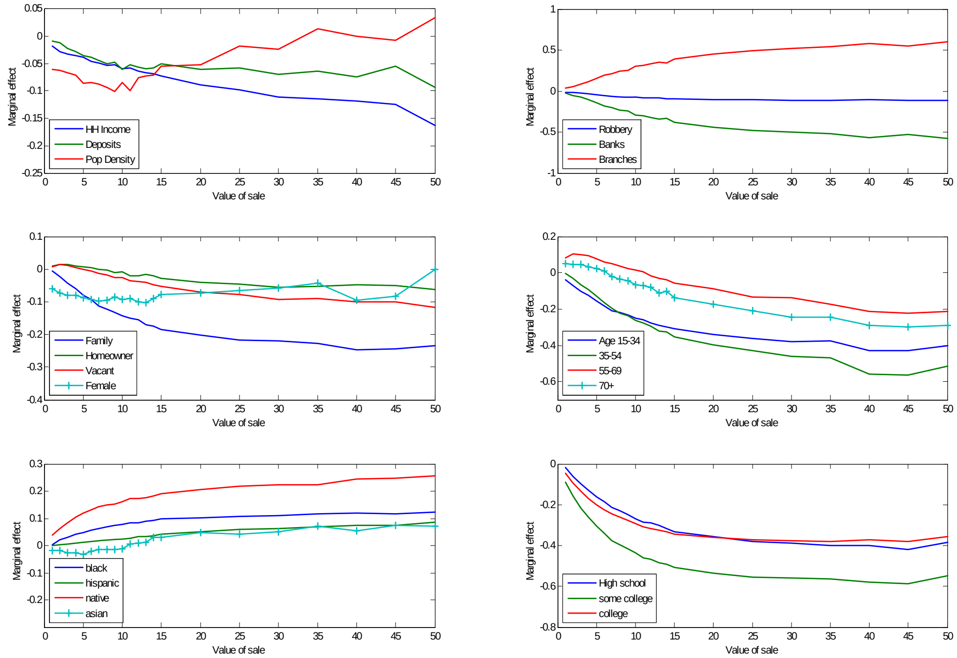

Figure 12. Cash marginal effects by transaction size.

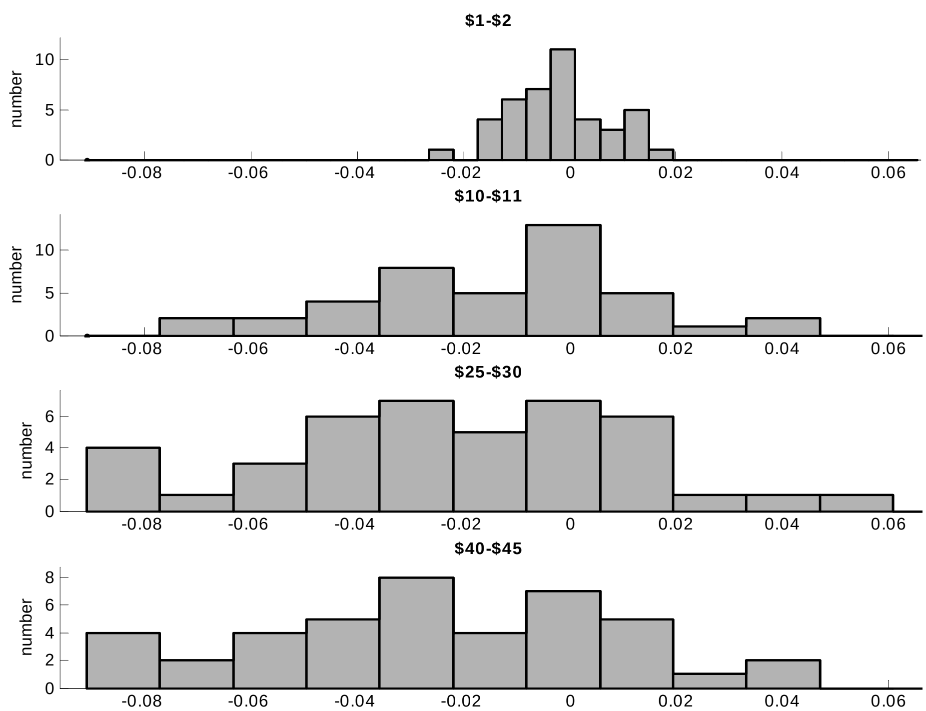

Figure 13. Cash: histogram of state effects.

Table 5. Ranking of cash state effects

	$1-$2	$10-$11	$25-$30	$40-$45
Top States				
	Delaware	New Jersey	New York	New York
	Minnesota	New York	New Jersey	New Jersey
	New Jersey	Michigan	Michigan	Michigan
	Vermont	Vermont	Mississippi	Mississippi
	Wisconsin	Delaware	Delaware	Maine
Bottom States				
	Idaho	New Mexico	New Mexico	New Mexico
	New Mexico	North Dakota	Nevada	North Dakota
	Nevada	Nevada	North Dakota	Arizona
	Florida	Idaho	Arizona	Nevada
	Arizona	Arizona	Idaho	Idaho

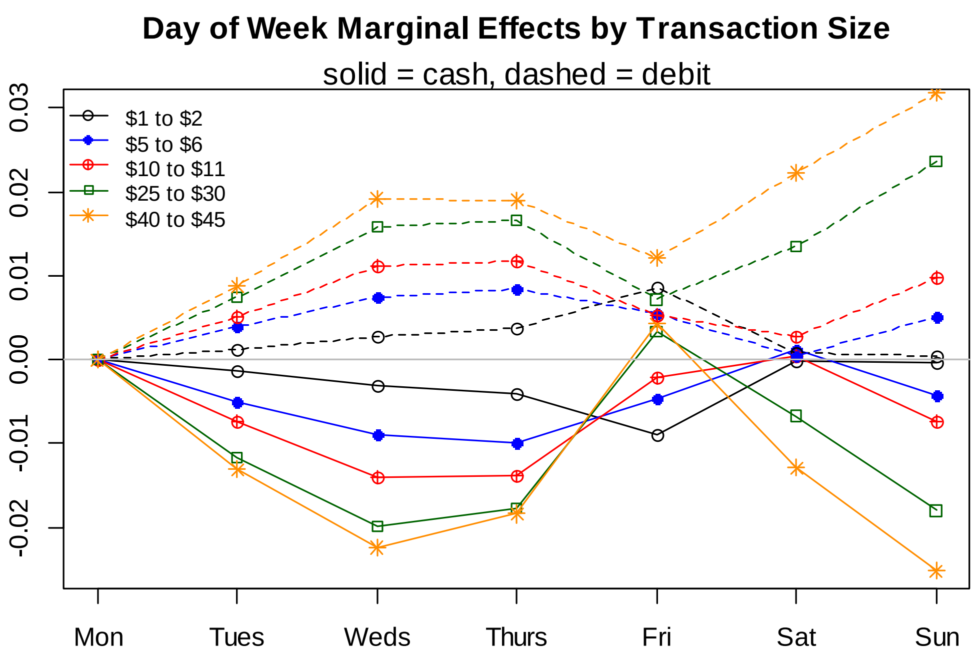

Figure 14.

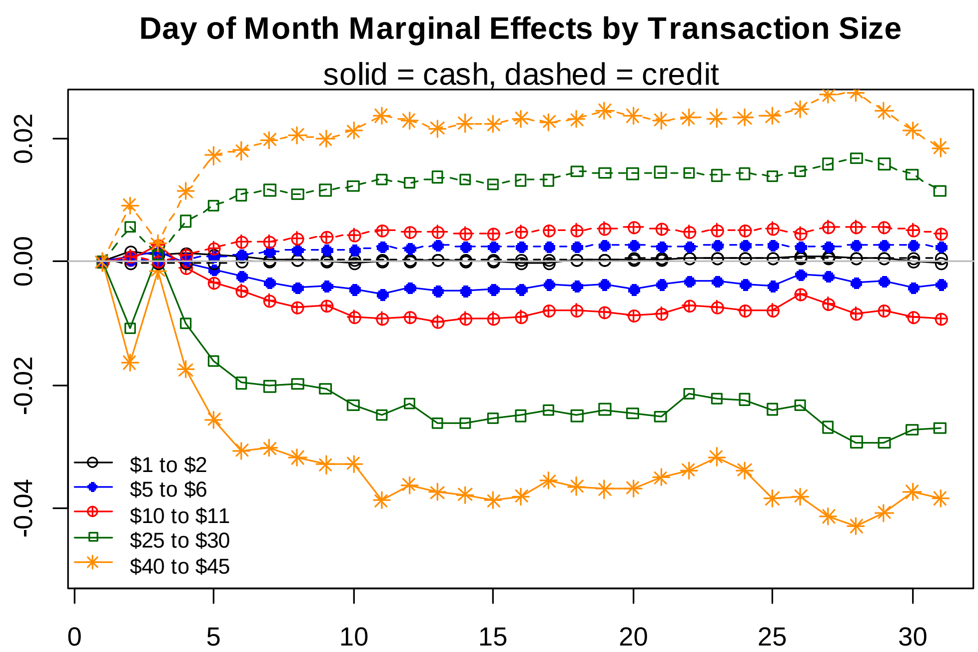

Figure 15.

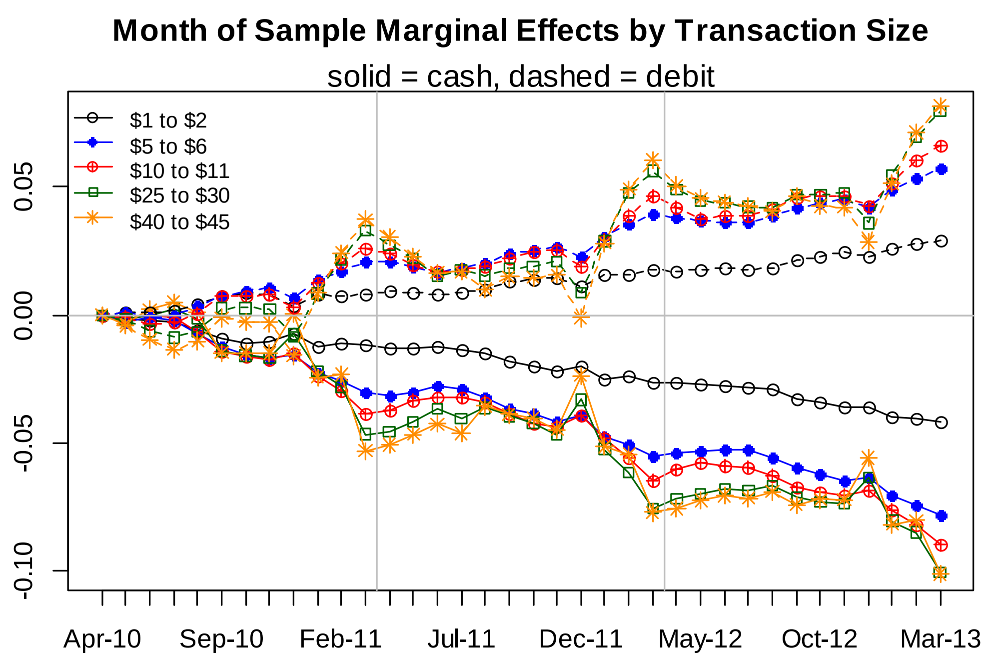

Figure 16.

4.3 Payment Variation by Transaction Size

Figure 17 displays the estimated counterpart to the raw data of Figure 3. For each size class, we plot the median, 5th, and 95th percentiles of the distribution of predicted values for the four payment shares. Comparing the two figures, it is clear that the estimated models for each transaction size are successful at replicating both (i) the relationship between transaction size and the level of payment composition, and (ii) the relationship between transaction size and the dispersion of payment composition across zip-code days. We now discuss how those relationships are related to the amplifying effects of explanatory variables.

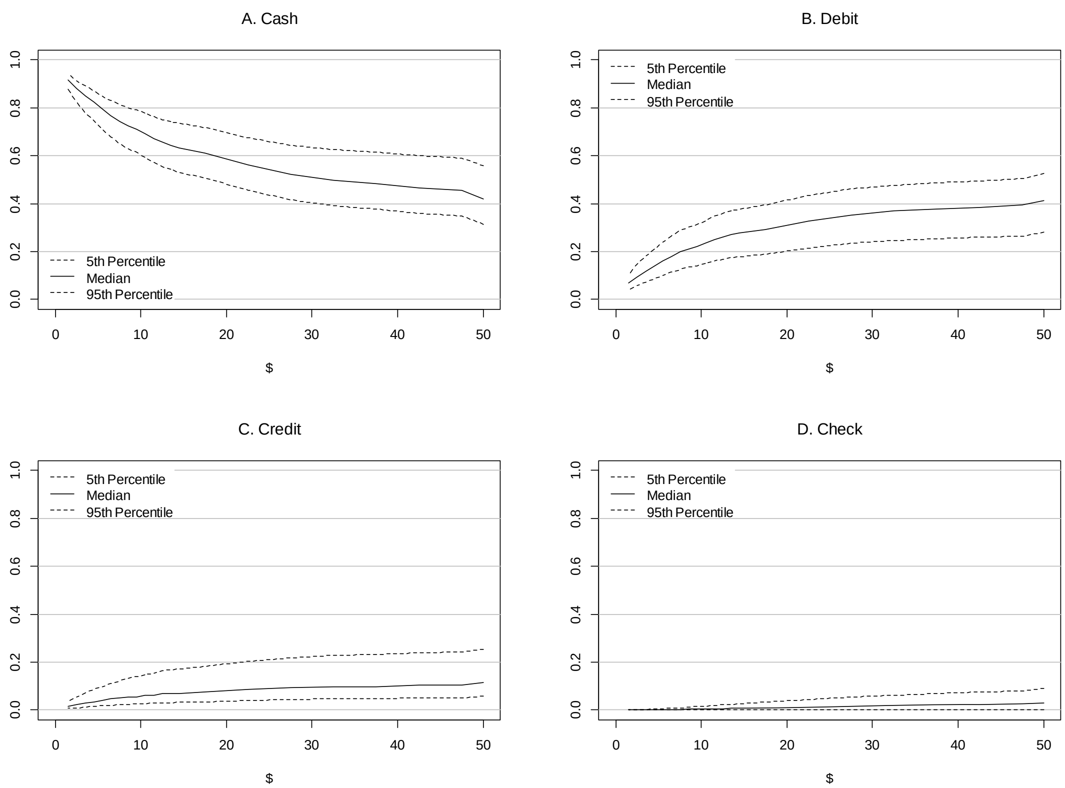

Figure 17. Predicted payment variation across transaction sizes.

While we have found above that marginal effects increase in transaction size for most explanatory variables, that does not mean all those variables are quantitatively important in explaining payment variations across transaction sizes. Even in a linear framework, the quantitative importance would depend on the combination of coefficients and the variation in each explanatory variable. In the FMLogit model, which is nonlinear, there is an added degree of complexity: The coefficients on all variables interact with each other and with the data in determining the marginal effect of a given variable. Amplification of marginal effects does not necessarily reflect amplification of coefficients, and it is the coefficients which matter for the quantitative contributions of different variables. Nonetheless it is possible to quantify those contributions, and we do so as follows.

We first divide the explanatory variables into two groups: One comprises constant terms, which include the intercept and time and state-level fixed effects, and the other comprises all zip-code-level variables. We wish to quantify the relative contributions of the two groups of variables to the levels and dispersions of payment mix across transaction sizes. Several questions can be asked: Does the negative relationship between cash share and transaction size reflect lower constant terms for higher transaction sizes, or changing coefficients on the zip-code-level variables? Why does the dispersion of the payment mix across zip codes increase with transaction size? In Section 3 we found that zip-code days with larger transactions were associated with a lower share of cash payments, but that finding is consistent with either the constant terms or zip-code-level variables driving the share levels in the transaction size-class regressions. Likewise, our theoretical framework of individual-specific threshold transaction sizes does not tell us which effect should dominate.[15]

We then decompose the level and dispersion of the payment mix across transaction sizes into components associated with the constant terms and the coefficients on zip-code-level variables. The decomposition uses the \$1-\$2 regression as a benchmark. First we allow the constant terms to take on their estimated values in each of the size-class regressions, holding fixed the coefficients on zip-code-level variables at the \$1-\$2 benchmark. Then we allow the coefficients on the zip-code-level variables to take on their estimated values in each of the size-class regressions, holding fixed the constant terms at their \$1-\$2 benchmark. The results of this decomposition are shown in Figure 18. For each size class, we plot the median, 5th, and 95th percentiles of the distribution of counterfactual values for each payment fraction. The lines marked with “x” come from the first exercise described above – allowing only the constant terms to vary, and the lines marked with “o”s come from the second experiment – holding fixed the constant terms and allowing the other coefficients to vary. Note that we do not re-estimate the model subject to restrictions; we simply use different combinations of the estimates from the \$1-\$2 regressions and the other size-class regressions.

Because of the nonlinearity inherent in the FMLogit model, the decomposition is not additive. In addition, there is no guarantee that it will unambiguously assign the change in the payment mix as transaction size changes to one or the other set of coefficients. However, Figure 18 shows that the decomposition turns out to be relatively clean: It is changes in the coefficients on zip-code-level variables, rather than changes in constants, that overwhelmingly account for changes in the level and dispersion of each payment type.

As we have emphasized throughout, the fraction of cash payments at a given transaction size represents the fraction of transactions made by consumers with thresholds above that size. Thus, the distribution of thresholds on each location day pins down the associated fraction of cash payments. The negative relationship between transaction size and the level of cash fractions is a straightforward implication of theory: For any distribution of thresholds, a higher transaction size corresponds to a higher fraction of consumers whose threshold for switching away from cash has been crossed. In principle, our econometric model could account

[15] If we made parametric assumptions about (1) the function matching characteristics to the threshold transaction size, and (2) the distribution of characteristics, those assumptions would imply restrictions on the roles of constant terms and zip-code-level variables, with the latter standing in for the zip-code-level distribution of characteristics. However, we choose to view our FMLogit model as a low-order approximation to arbitrary threshold functions and distributions of characteristics.

for that negative relationship with various combinations of changes across transaction sizes in the constant terms or the zip-code-level coefficients; in practice, the decomposition presented above attributes the negative relationship entirely to changes in the zip-code-level coefficients.

The increasing relationship between dispersion of payment fractions and transaction size is not necessarily implied by theory. However, that relationship is intuitive: For higher transaction sizes, the fixed costs of using non-cash instruments become less important, and thus consumers in locations with better access to those instruments behave increasingly differently than consumers in locations with worse access. The fact that our decomposition attributes the relationship to the zip-code-level coefficients reveals that is primarily the zip-code-level variables that proxy for access to non-cash payments.

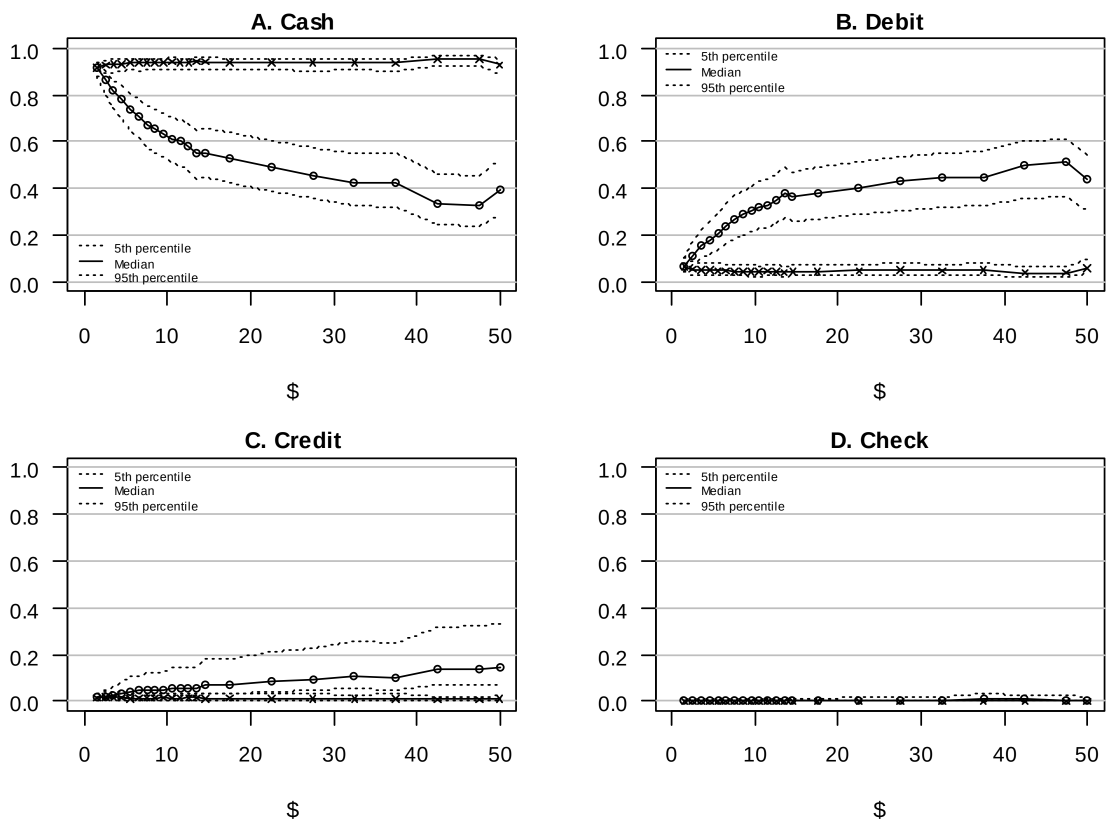

Figure 18. Decomposition of payment variation.

(Fixed zip-code-level coefficients (x) Fixed constants(o))

4.4 Long-run Trends by Transaction Size

In Section 4.2 we reviewed the time marginal effects from the transaction size-class regressions. Here we discuss how the predicted payment mix varies from the beginning to the end of the sample period, as well as the time trends implied by the estimated time effects.

Figure 19 compares the predicted payment mix at the mean values of the explanatory variables for the first and last months of our sample; the lines marked with x's represent April 2010, and the lines marked with o's represent March 2013. For each transaction size, the x's and the o's are from the same set of regressions, simply evaluated at different values of the time dummies. In contrast, the different transaction sizes represent different regressions. There is a marked downward shift in the predicted cash and check fractions, and corresponding upward shifts in the predicted debit and credit fractions. The size of the shift is generally increasing in transaction size.

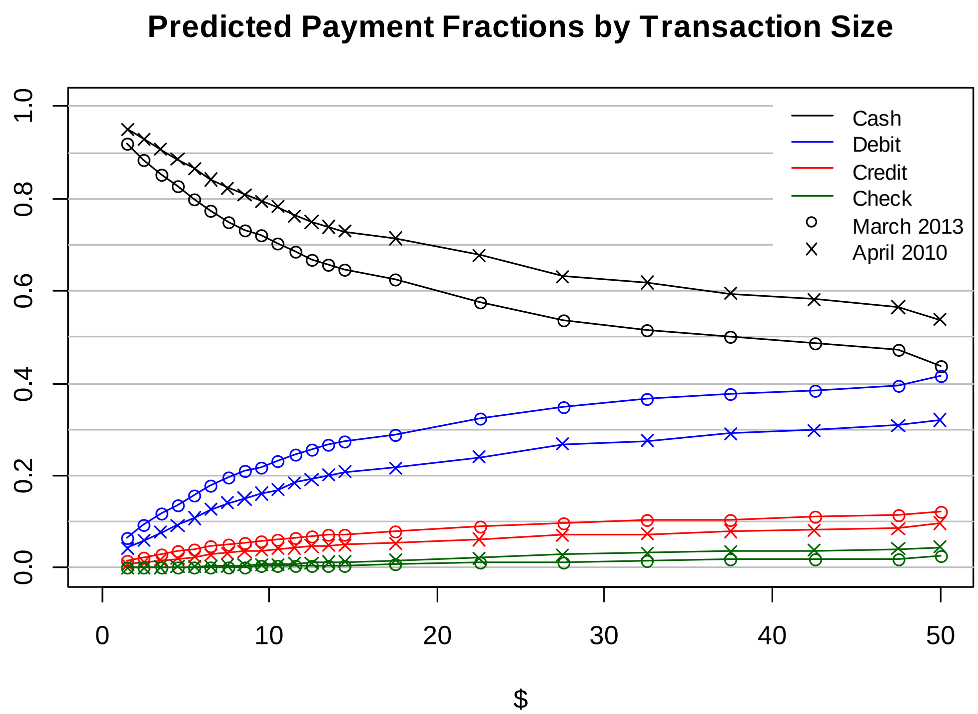

Figure 19.

As with the regression of overall payment shares in Section 3, we estimated linear time trends for each payment size within each payment type. The resulting linear trends are plotted as annual percentage point changes in Figure 20. In almost all cases, the time trends are greater in absolute value for larger payment sizes. For cash, the time trends range from a decrease of 1.3 percentage points per year for \$1-\$2 transactions to a decrease of 3.32 percentage points per year for \$20-\$25 transactions; for debit, the trends range from an increase of less than 1 percentage point per year for \$1-\$2 transactions to 2.6 percentage points per year for transactions greater than \$50. In general the time trends indicate replacement of cash with debit. However, roughly one-third of the decline in cash is accounted for by an increase in credit. The increase in credit ranges from 0.45 to 1.13 percentage points per year across transaction sizes.

The estimated time trends for each payment size can be used as a foundation to forecast the future consumer payments mix. One application of particular interest involves forecasting currency use. We turn to this topic in the next section.

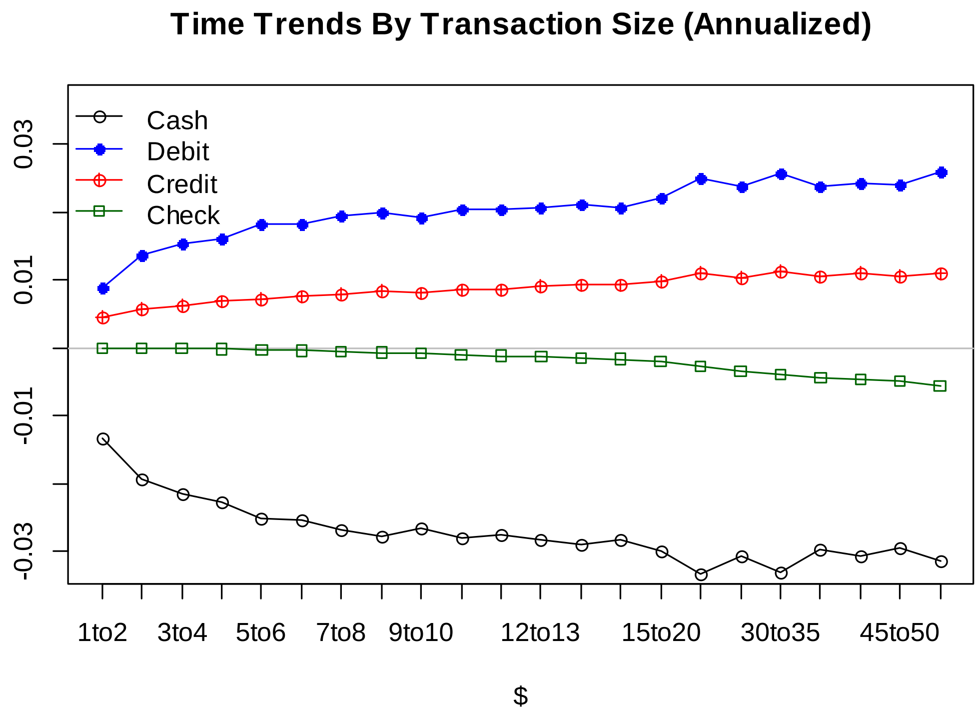

Figure 20.

5 Forecasting the Mix of Payments and the Future of Currency

Our econometric model can be used to forecast the future composition of payments at the discount retailer, and presumably the forecast would be similar for other retailers in the same market segment. The cash component of those forecasts is related to the level of currency use in transactions, which in turn has implications for money demand. Below we first present the forecasts specific to the discount retailer. We then discuss how those forecasts can be informative about the level of overall currency use going forward, even though the discount retailer represents a small fraction of the total value of retail sales.

5.1 Currency's Share in Discount Retail

In order to forecast the retailer's payments mix, we begin with the predicted mix for March 2013, as shown in Figure 19. We then incorporate a time trend by assuming the payment mix will change each year at an exogenous rate implied by the marginal effects associated with our estimated coefficients on month-of-sample dummies. The time trends we impose are those represented by the black open circles in Figure 20, for each transaction size bin. In Figure 21, the blue and red lines without symbols represent the estimated cash fractions from our transaction-size regressions, evaluated at the means of the explanatory variables, but with the time dummies set at March of 2011 and 2013. The black lines with open and closed circles display the forecasted cash fractions for 2015 and 2020 implied by the estimated time trends. For the smallest transactions ($1-$2), cash accounted for 91.9 percent of the total in March 2013, and we predict that cash will fall to 89.2 percent in 2015 and 82.6 percent in 2020. For transactions in the $5-$6 range, cash accounted for 80.0 percent of transactions in March 2013, and we predict that it will fall to 75.0 percent in 2015 and 62.4 percent in 2020. And for transactions in the $40-$45 range, the predicted decline in cash is from 48.6 percent of transactions in 2013 to 42.5 percent in 2015 and 27.1 percent in 2020.

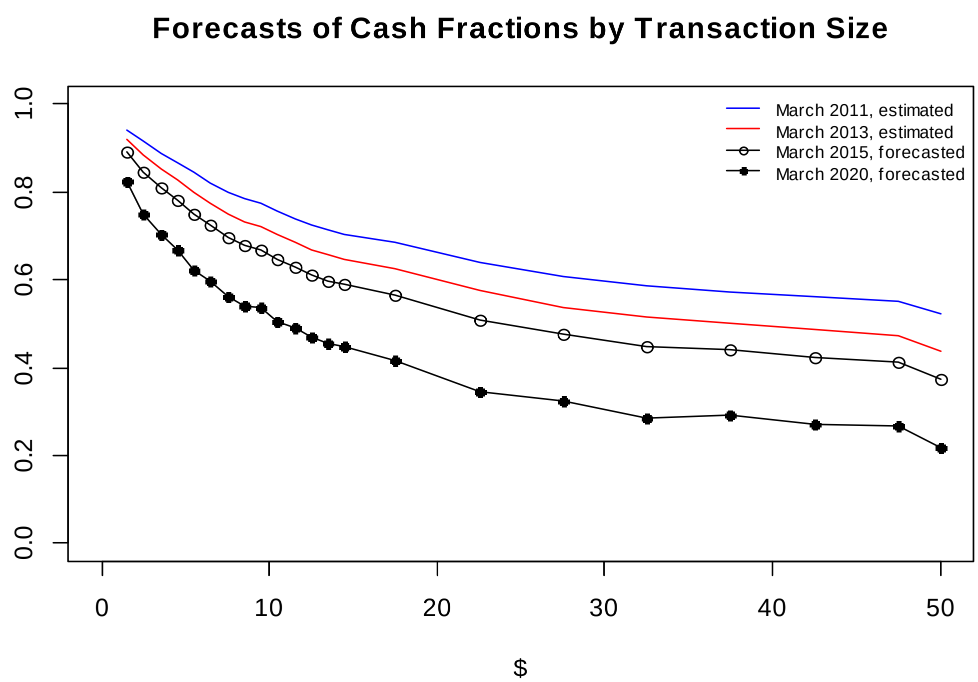

Figure 21.

Several forces may be driving the time trend, with prime candidates being technological progress and changing consumer perceptions of the attributes of each payment instrument. These attributes include setup costs; marginal cost of transactions; speed of transactions; security; record keeping; merchant acceptance; ease of use and possibly other attributes, none of which are directly included in our regressions. Stavins (2013) provides an extensive discussion of the role that consumer perceptions of security seem to play in the adoption and use of each payment instrument, and Stavins (2014) provides evidence from the Survey of Consumer Payment Choice showing continuous improvement in consumers' perceptions of card security relative to cash in recent years.[16]

Of course, while our regressions do not include measures of payment attributes, they do include a large number of zip-code-level variables. Forecasted changes in those variables also imply forecasted changes in the cash fractions. However, it would be inappropriate to incorporate forecasts for the zip-code-level variables

[16] The Survey of Consumer Payment Choice is a longitudinal panel survey conducted by the Federal Reserve Bank of Boston every year since 2008, covering approximately 2,000 consumers.

in addition to the time trend. Recall that we treated all zip-code-level variables as fixed at their 2011 values across time in the regressions. Therefore any time trend is picked up by the month of sample dummies, even if some of the trend is actually associated with time variation in the zip-code-level variables. While we cannot add together the effects of forecasted changes in zip-code-level variables with our estimated time trend, we can use those forecasted changes together with our regression estimates to gain some insight into the role of demographic and other zip-code-level changes in accounting for time trend.

We forecast the zip-code-level variables as follows. For racial composition and age composition, we use the United States Census Department's projections, adjusted for the level differences between the means of our sample and the national averages.[17] We interpret the age/cohort effects as primarily representing cohort. That is, for cohorts that were 15 or above in 2011, we assume that they carry with them their estimated coefficient through 2020.[18] The exception is the group age 14 and below; we assume that the regression estimates for this group represent an age and not a cohort effect. We forecast median nominal household income to grow at a 2.5 percent annual rate, which is approximately equal to the 20-year national average. Educational attainment has been rising, and we forecast that it will continue to increase but at a slowing rate: The mean percentage of college graduates in our sample zip codes was 26.24 percent in 2011, and we forecast that it will reach 29.04 percent in 2015 and 32.04 percent in 2020. Bank branches per capita are forecasted to increase at 1 percent per year based on a trend identified from the FDIC's Summary of Deposits. The housing vacancy rate is forecasted to decline from 13.16 percent in 2011 to 12.25 percent in 2015 and 11.75 percent in 2020. All other zip-code-level explanatory variables are projected to remain constant at their zip-code-level means. We hold the day-of-week and day-of-month dummies fixed at their means. Holding fixed the month-of-sample dummies at March 2011, this procedure gives us forecasts for the payment mix based solely on changes in the zip-code-level variables. Note that there is a separate forecast associated with each of the payment size regressions.

The "contributions" of demographic and other zip-code-level changes to our forecasts of cash use are displayed in Figure 22. Each of the lines denoted by a square or a circle plots the difference between (i) a forecasted cash fraction that is based on a particular forecasted change in zip-code-level variables, and (ii), the estimated cash fractions for March 2011 (the top line in Figure 21). For the sake of comparison, we also include in Figure 22 the overall forecasted declines in cash use implied by our estimated time trends. There are two main messages from Figure 22. First, a majority of the decline in cash use that we can attribute to changes in zip-code-level variables is due to the cohort effect: For 2015, between 53% and 75% of the zip-code-level effects represent cohort effects, across transaction sizes, and for 2020 these numbers rise to 71% and 79%, respectively. Second, while the cohort effects are important relative to other zip-code-level effects, the overall effects of zip-code-level variables are small relative to the time trends: For our 2020 forecasts, the effects of forecasted changes in zip-code-level variables represent between 11.6% and 15.2% of the changes

[17] The Census projections are available at http://www.census.gov/population/projections/data/national/2012/summarytables.html. Forecasts for all demographic variables are available upon request.

[18] For example, in constructing the 2020 forecast, we apply the estimated coefficients for age 15-34 to the fraction of the population that is forecasted to be age 40 in 2020.

in cash use implied by our estimated time trends.

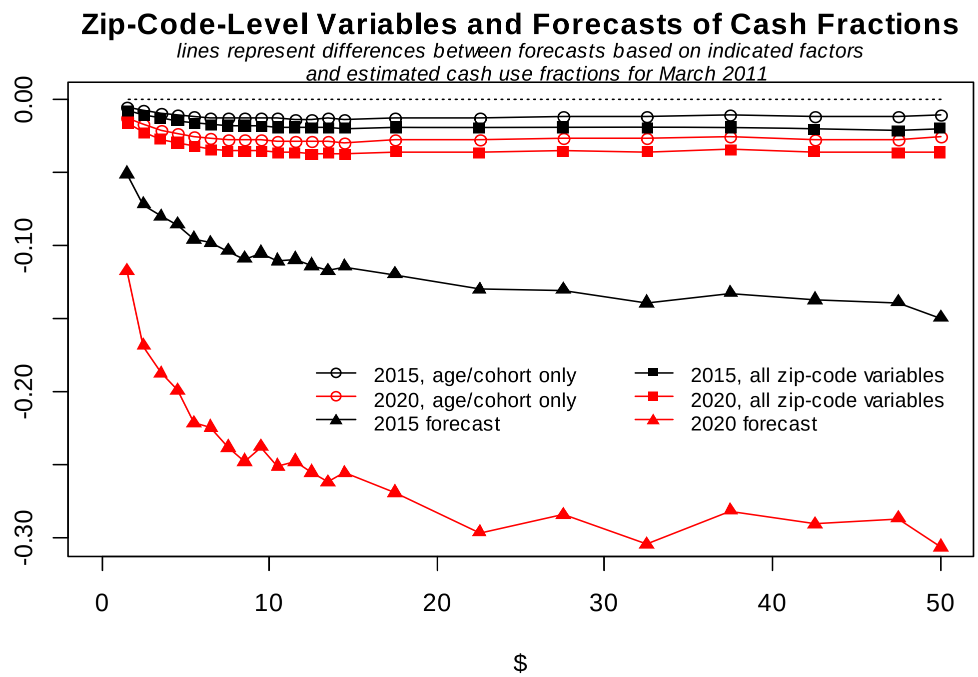

Figure 22.

Returning now to the time trends, in order to predict overall cash use at this retailer, we can combine the forecasts in Figure 21, for cash use at each transaction size, with the size distribution of transactions. For March 2013 this yields cash transactions as 75.0 percent of the total. The forecast for 2015 is that cash will account for 70.1 percent of transactions, and for 2020 cash will account for 57.8 percent of transactions. From 2013 to 2020 then, we forecast that the cash share of transactions will decline by 2.46 percentage points per year. These forecasts assume that the size distribution of transactions will remain constant. If the size distribution were to shift upward, as one might expect given our forecast of 2.5 percent nominal income growth, then the cash fraction of transactions would likely decline more. To illustrate the additional effects that could come from a shifting size distribution, consider the following crude experiment: Suppose that by 2020 the cumulative distribution of payment size shifts to the right exactly one bin, so that, for example, the fraction of transactions less than \$7 in 2020 is identical to the fraction of transactions less than \$6 in 2012. Under this additional assumption, instead of forecasting a 57.2 percent cash share in 2020 we would forecast a 54.1 percent cash share, representing a decrease of 3.0 percentage points per year. This experiment may be conservative: In 2010 the fraction of transactions less than \$7 was 0.53, and by 2013 the fraction of transactions less than \$8 was just above that level, at 0.54.

5.2 The Future of Currency Use in Retail Transactions

The forecasts displayed in Figure 21 and discussed above assume that the time trend observed in our sample of 36 months continues over the next seven years. Whether the trend will continue is of course uncertain, but the presence of that trend in our data is quite clear. We argued in the introduction that the uniquely cash-intensive nature of our data, while rendering it unrepresentative of the U.S. economy, made it particularly well-suited to studying the behavior of cash. As such, we can use our forecasts to think about the future of currency use more broadly.

Nominal retail sales in the United States grew at a 3.7 percent rate in 2013.[19] However, currency is a feasible payment instrument only for in-person sales, and the in-person component of retail sales grew only 2.5 percent in 2013. The future of currency as a means of payment in legitimate transactions is a race between, on the one hand, the growth of in-person nominal retail sales, and on the other hand, the decline in currency's share of in-person sales, as predicted in Figure 21. In general, suppose the cash share of in-person retail transactions is s in some initial period (i.e. 2013); suppose overall in-person retail is growing at annual rate μ, and the cash share of in-person retail is falling at rate δ, where δ is measured in percentage points per year. If we denote total in-person transactions in period t by R_t, then the level of cash use, C_t, in the initial period is given by $C_0 = sR_0$, and in subsequent periods we have

$$C_t = (s - \delta t)\, R_{t-1}\,(1+\mu)\,,\ t = 1, 2, ...$$

It follows that the level of cash use will fall after the initial period ($C_1 < C_0$) if the following condition holds:

$$C_1/C_0 < 1\ \Rightarrow \frac{(s-\delta)\,(1+\mu)}{s} < 1\ \Rightarrow s < \frac{\delta\,(1+\mu)}{\mu}. \tag{9}$$

Assuming that $\mu = 0.025$ (the growth rate of in-person retail in 2013), and given our estimated $\delta = 0.0246$, it follows from (9) that the *level* of cash use must be falling regardless of the overall cash share (note that $0.025 < (0.0246 \times 1.025)$). Even for our discount retailer, with a relatively high cash share of 0.75, the fact that the decline in the share of cash transactions outpaced the nominal growth rate of in-person retail sales implies an absolute decrease in cash use. Furthermore, there may be reasons to adjust *upward* the threshold in (9). First, the growth rate used for in-person retail sales refers to nominal value, but the rest of our analysis is in terms of number of transactions. It seems likely that the number of transactions is growing more slowly than the value of retail transactions. Another reason for adjusting upward the threshold for s is that new forms of electronic payments may lead to a faster decline in the cash share. In particular, mobile payments are just emerging and may experience strong growth in coming years, especially for small dollar transactions and at the expense of cash. Additionally, short-term nominal interest rates have been close to zero for the entire time period covered by our study. If interest rates rise in the coming years, theory suggests

[19] These and related numbers that follow are taken from the U.S. Census Department's monthly retail sales report, available at http://www.census.gov/retail/.

that the increased opportunity cost of holding cash will cause households to further reduce cash use. Finally, there is the question of the overall cash share of in-person retail (transactions, not value), as of 2013. As a conservative estimate the discount retailer's 0.75 share seems reasonable: Its transactions are small and cash-intensive relative to grocery stores or department stores, but presumably the overall distribution of in-person transactions (as opposed to value) is heavily weighted toward small transactions (drinks, snacks, etc.). Summing up, this line of reasoning suggests that the number of legitimate cash transactions is likely to begin declining in the next few years, if it is not declining already.

6 Conclusion

Using data on almost 2 billion transactions from a discount retailer, we have studied the variation in payment mix across location, time, and transaction size. There is large variation in the payment mix across each of these dimensions, and our empirical model is quite successful in accounting for that variation. Our analysis identifies important economic and demographic effects, weekly, monthly and seasonal cycles in payments, as well as time trends and state fixed effects. We show that changes in the coefficients on the zip-code-level variables account for most of the variation in the payment mix across transaction sizes, affecting both level and dispersion.

We also use the estimated model to forecast how the mix of consumer payments will evolve and to forecast future currency use. The key input to those forecasts comes from the marginal effects associated with our estimated month-of-sample dummy variables. These marginal effects indicate that the fraction of transactions conducted with cash has been declining at a rate of between 1.3 and 3.3 percentage points per year, depending on the size of transactions being considered. Combining the time trends with information about the size distribution of payments, we project that the cash share of transactions will decline at 2.46 percentage points per year, from its current level of 75 percent. A relatively small portion of this decline can be attributed to forecasted changes in the zip-code-level variables.

Although the retailer we study represents a small fraction of the value of U.S. retail sales, in absolute terms it has a large number of cash transactions – more than half a billion per year. As such, our projections are useful for considering the future of currency more generally. The trend decrease in the cash share of transactions in our data implies that the number of above-ground cash transactions is currently falling and will continue to fall over the next several years.

In future research with this data it would be interesting to investigate in more detail the residual variation in payment mix across states, which is not explained by the location-specific explanatory variables. To the extent that the cross-state variation is associated with different legal and regulatory environments, it may provide useful information for evaluating policy.

Our findings have implications for the continuing development of theories of money demand, cash holding and payment choice. It is clear from the time- and location-specific variation in payment choice that operational versions of those theories must incorporate consumer heterogeneity, in both adoption and use of different means of payment. Work along these lines is being done, for example by Alvarez and Lippi (2009, 2013), but it remains at an early stage. As that work continues, our findings may be useful for pinning down the parameters of models. That process would be facilitated if we could complement our data with information on the behavior of individual consumers. In particular, to better understand the patterns in our time dummies and relate that behavior to inventory theory, we need information about households' balance sheets over the course of the week and month. Tracking consumers would also reveal the extent to which time variation in payment choice reflects time variation in the composition of customers paying, as opposed to time variation in the payment choices of a fixed set of customers.

References

[1] Alvarez, F.E. and F. Lippi (2013). "Cash Burns." Mimeo, Department of Economics, University of Chicago.

[2] Alvarez, F. E. and F. Lippi (2009). "Financial Innovation and the Transactions Demand for Cash." *Econometrica*, 77(2) 363-402.

[3] Arango, C., K. P. Huynh, and L. Sabetti (2011). "How Do You Pay? The Role of Incentives at the Point-of-Sale." Working Paper 2011-23, Bank of Canada.

[4] Baumol, W.J. (1952). "The Transactions Demand for Cash: An Inventory Theoretic Approach." *Quarterly Journal of Economics* 66(4), 545–556.

[5] Borzekowski, R., E. K. Kiser, and S. Ahmed (2008). "Consumers' Use of Debit Cards: Patterns, Preferences, and Price Response" *Journal of Money, Credit, and Banking*, 40(1), 149-172.

[6] Borzekowski, R., E. K. Kiser (2008). "The Choice at the Checkout: Quantifying Demand across Payment Instruments." *International Journal of Industrial Organization*, 26(4), 889–902.

[7] Buis, M. L. *FMLogit: Stata module fitting a fractional multinomial logitmodel by quasi maximum likelihood*. Statistical Software Components, Department of Economics, Boston College, June 2008.

[8] Ching, A. T. and F. Hayashi (2010). "Payment Card Rewards Programs and Consumer Payment Choice," *Journal of Banking and Finance*, 34(8), 1773-1787.

[9] Cohen, M. and M. Rysman (2012). "Payment Choice with Consumer Panel Data." Memo, Department of Economics, Boston University.

[10] Freeman, S. and F. E. Kydland (2000). "Monetary Aggregates and Output." *American Economic Review* 90(5), 1125-1135.

[11] Gourieroux, C., A. Monfort, and A. Trognon (1984). "Pseudo Maximum Likelihood Methods: Theory." *Econometrica* 52: 681-700.

[12] Judson, R. A. and R. D. Porter (2004).“Currency Demand by Federal Reserve Cash Office: What Do We Know?” *Journal of Economics and Business*, 56 (4), 273-285.

[13] Klee, E. (2008). “How People Pay: Evidence from Grocery Store Data.” *Journal of Monetary Economics*, 55, 526-541.

[14] Koulayev, S., M. Rysman, S. Schuh and J. Stavins (2012). “Explaining Adoption and Use of Payment Instruments by U.S. Consumers.” Federal Reserve Bank of Boston Working Paper No. 12-14.

[15] Lucas, R.E. Jr. (1982). “Interest Rates and Currency Prices in a Two-country World.” *Journal of Monetary Economics*, 10(3), 335-359.

[16] Lucas, R.E. Jr. and J.P. Nicolini (2013). “On the Stability of Money Demand.” Memo, Department of Economics, University of Chicago.

[17] McAndrews, J. and Z. Wang (2012). “The Economics of Two-Sided Payment Card Markets: Pricing, Adoption and Usage.” Research Working Paper 12-06, Federal Reserve Bank of Richmond.

[18] Mullahy, J. (2010). “Multivariate Fractional Regression Estimation of Econometric Share Models.” NBER Working Papers 16354.

[19] Papke, L. E. and J. M. Wooldridge (1996). “Econometric Methods for Fractional Response Variables with An Application to 401(k) Plan Participation Rates.” *Journal of Applied Econometrics* 11(6), 619–32.

[20] Prescott, Edward C. (1987). “A Multiple Means of Payment Model,” in William A. Barnett and Kenneth N. Singleton, eds., *New Approaches to Monetary Economics*, Cambridge University Press: 42-45.

[21] Sidrauski, Miguel (1967). “Rational Choice and Patterns of Growth in a Monetary Economy.” *American Economic Review* 57 (2): 534–544.

[22] Stavins, J. (2014). “Comments on ‘Payment Choice and the Future of Currency’ by Wang and Wolman.” IBEFA Meetings, Denver.

[23] Stavins, J. (2013). “Security of Retail Payments: The New Strategic Objective.” Federal Reserve Bank of Boston Public Policy Discussion Paper No. 13-9.

[24] Schuh, S. and J. Stavins (2012). “How Consumers Pay: Adoption and Use of Payments.” Federal Reserve Bank of Boston Working Paper 12-2.

[25] Tobin, J. (1956). “The Interest Elasticity of the Transactions Demand for Cash.” *The Review of Economics and Statistics* 38 (3), 241–247.

[26] Zinman, J. (2009). “Debit or Credit?” *Journal of Banking and Finance*, 33(2), 358–366.

Appendix A.

Supplementary Material for Size Class Regressions

For the sake of space, we only report the FMLogit estimation results regarding cash in Section 4. In this Appendix, we provide the remaining FMLogit estimation results that are related to debit, credit and check. The results are shown in the following order.

- Table A1. Debit: marginal effects by transaction size
- Table A2. Credit: marginal effects by transaction size
- Table A3. Check: marginal effects by transaction size
- Figure A1. Debit marginal effects by transaction size
- Figure A2. Credit marginal effects by transaction size
- Figure A3. Check marginal effects by transaction size
- Figure A4. Debit: histograms of state effects by transaction size
- Figure A5. Credit: histograms of state effects by transaction size
- Figure A6. Check: histograms of state effects by transaction size
- Table A4. Debit: rankings of state effects
- Table A5. Credit: rankings of state effects
- Table A6. Check: rankings of state effects
- Figure A7. Day of week marginal effects by transaction size (credit vs check)
- Figure A8. Day of month marginal effects by transaction size (debit vs check)
- Figure A9. Month of sample marginal effects by transaction size (credit vs check)

Table A1. Debit: marginal effects by transaction size

Variable	\$1-\$2	\$5-\$6	\$10-\$11	\$15-\$20	\$25-\$30	\$40-\$45	above \$50
Cash holding and payment choice							
Banks per capita	0.004	0.067*	0.171*	0.223*	0.282*	0.335*	0.340*
Branches per capita	-0.007*	-0.072*	-0.178*	-0.231*	-0.292*	-0.345*	-0.350*
Robbery rate	0.015*	0.051*	0.087*	0.113*	0.142*	0.145*	0.174*
Adoption of non-cash payments							
Median household income	0.004*	0.011*	0.018*	0.022*	0.039*	0.053*	0.081*
Deposits per capita	0.007*	0.026*	0.050*	0.051*	0.070*	0.096*	0.112*
Population density	0.044*	0.058*	0.092*	0.156*	0.270*	0.363*	0.480*
Demographics							
Family households	0.008*	0.071*	0.126*	0.164*	0.194*	0.215*	0.200*
Owner-occupied	-0.007*	-0.009*	-0.002	0.007*	0.009*	0.002	0.004
Vacant housing	-0.009*	-0.013*	-0.002	0.008*	0.014*	0.020*	0.020*
Female	0.052*	0.089*	0.116*	0.133*	0.152*	0.185*	0.137*
Age 15-34	0.035*	0.135*	0.221*	0.276*	0.326*	0.375*	0.344*
35-54	-0.001	0.096*	0.205*	0.275*	0.335*	0.430*	0.381*
55-69	-0.060*	-0.039*	0.021*	0.086*	0.152*	0.192*	0.182*
$\geq$ 70	-0.051*	-0.058*	-0.016*	0.011	0.039*	0.101*	0.079*
Race black	0.004*	-0.028*	-0.042*	-0.049*	-0.043*	-0.042*	-0.030*
Hispanic	0.000	-0.012*	-0.023*	-0.035*	-0.045*	-0.058*	-0.066*
Native	-0.025*	-0.073*	-0.090*	-0.102*	-0.111*	-0.127*	-0.122*
Asian	0.007*	0.011*	-0.007	-0.025*	-0.028*	-0.011	-0.008
Pac-Islr	0.114*	0.455*	0.722*	0.913*	1.147*	1.316*	1.449*
other	-0.018*	-0.040*	-0.054*	-0.056*	-0.029*	0.019*	0.056*
multiple	0.118*	0.152*	0.091*	0.054*	0.069*	0.060*	0.126*
Edu high school	0.008*	0.114*	0.189*	0.226*	0.243*	0.247*	0.235*
some college	0.064*	0.222*	0.314*	0.360*	0.384*	0.401*	0.384*
college	0.030*	0.130*	0.185*	0.208*	0.209*	0.195*	0.173*
Time & state dummies	included	included	included	included	included	included	included
Pseudo R-squared	0.12	0.17	0.12	0.23	0.12	0.05	0.10
Zip code-days (1,000)	4,505	4,505	4,498	4,505	4,483	4,045	4,405
Transactions (1,000)	198,700	129,299	67,465	132,108	50,800	16,425	37,905

*Significant at 1%. Units of regression variables are defined in footnote 10.

Table A2. Credit: marginal effects by transaction size

Variable	\$1-\$2	\$5-\$6	\$10-\$11	\$15-\$20	\$25-\$30	\$40-\$45	above \$50
Cash holding and payment choice							
Banks per capita	0.024*	0.079*	0.123*	0.164*	0.194*	0.218*	0.222*
Branches per capita	-0.025*	-0.083*	-0.128*	-0.170*	-0.202*	-0.227*	-0.232*
Robbery rate	-0.004*	-0.003*	0.000	0.001	-0.002	0.003	-0.001
Adoption of non-cash payments							
Median household income	0.013*	0.031*	0.051*	0.068*	0.090*	0.104*	0.124*
Deposits per capita	0.002*	0.014*	0.022*	0.024*	0.034*	0.029*	0.051*
Population density	0.019*	0.069*	0.126*	0.177*	0.239*	0.272*	0.331*
Demographics							
Family households	-0.003*	0.011*	0.023*	0.035*	0.048*	0.066*	0.081*
Owner-occupied	-0.002*	-0.001	0.003*	0.004*	0.004*	0.000	0.001
Vacant housing	0.001*	0.011*	0.023*	0.034*	0.043*	0.051*	0.059*
Female	0.010*	0.008*	0.000	-0.008*	-0.008	0.017	0.003
Age 15-34	0.004*	0.026*	0.045*	0.063*	0.085*	0.121*	0.147*
35-54	0.004*	0.039*	0.078*	0.114*	0.160*	0.217*	0.259*
55-69	-0.023*	-0.032*	-0.014*	0.011*	0.044*	0.094*	0.123*
≥ 70	0.000	0.035*	0.077*	0.112*	0.153*	0.173*	0.204*
Race black	-0.007*	-0.018*	-0.026*	-0.030*	-0.033*	-0.039*	-0.043*
Hispanic	0.001*	0.002*	0.004*	0.006*	0.011*	0.016*	0.023*
Native	-0.012*	-0.045*	-0.067*	-0.079*	-0.093*	-0.100*	-0.112*
Asian	0.012*	0.029*	0.036*	0.037*	0.048*	0.045*	0.053*
Pac-Islr	0.006	-0.104*	-0.234*	-0.393*	-0.453*	-0.432*	-0.304*
other	-0.011*	-0.036*	-0.063*	-0.079*	-0.091*	-0.087*	-0.091*
multiple	0.015*	0.021*	0.006	-0.015*	-0.051*	-0.122*	-0.209*
Edu high school	0.011*	0.045*	0.073*	0.094*	0.115*	0.125*	0.125*
some college	0.024*	0.082*	0.123*	0.148*	0.175*	0.189*	0.185*
college	0.015*	0.066*	0.102*	0.124*	0.145*	0.155*	0.163*
Time & state dummies	included	included	included	included	included	included	included
Pseudo R-squared	0.08	0.16	0.14	0.28	0.15	0.07	0.11
Zip code-days (1,000)	4,505	4,505	4,498	4,505	4,483	4,045	4,405
Transactions (1,000)	198,700	129,299	67,465	132,108	50,800	16,425	37,905

*Significant at 1%. Units of regression variables are defined in footnote 10.

Table A3. Check: marginal effects by transaction size

Variable	\$1-\$2	\$5-\$6	\$10-\$11	\$15-\$20	\$25-\$30	\$40-\$45	above \$50
Cash holding and payment choice							
Banks per capita	-0.000*	-0.003*	-0.005*	-0.006*	0.000	0.014*	0.020*
Branches per capita	0.000*	0.003*	0.006*	0.008*	0.003	-0.011	-0.014*
Robbery rate	-0.000*	-0.004*	-0.012*	-0.020*	-0.034*	-0.041*	-0.060*
Adoption of non-cash payments							
Median household income	-0.000*	-0.003*	-0.009*	-0.018*	-0.031*	-0.039*	-0.041*
Deposits per capita	0.000	-0.005*	-0.011*	-0.025*	-0.045*	-0.050*	-0.070*
Population density	-0.002*	-0.041*	-0.134*	-0.279*	-0.491*	-0.635*	-0.846*
Demographics							
Family households	-0.000*	-0.002*	-0.007*	-0.014*	-0.025*	-0.034*	-0.046*
Owner-occupied	0.000*	0.003*	0.008*	0.017*	0.032*	0.047*	0.058*
Vacant housing	0.000*	0.002*	0.005*	0.012*	0.021*	0.031*	0.038*
Female	-0.000*	-0.008*	-0.024*	-0.047*	-0.079*	-0.106*	-0.139*
Age 15-34	-0.000*	-0.005*	-0.016*	-0.029*	-0.050*	-0.064*	-0.087*
35-54	-0.000*	-0.007*	-0.020*	-0.037*	-0.064*	-0.089*	-0.129*
55-69	-0.000*	-0.007*	-0.021*	-0.041*	-0.063*	-0.074*	-0.089*
≥ 70	0.000	0.002*	0.006*	0.014*	0.018*	0.018*	0.006
Race black	-0.000*	-0.004*	-0.010*	-0.019*	-0.030*	-0.038*	-0.048*
Hispanic	-0.000*	-0.002*	-0.007*	-0.014*	-0.024*	-0.033*	-0.045*
Native	-0.000*	-0.002*	-0.005*	-0.010*	-0.016*	-0.018*	-0.022*
Asian	0.000	-0.006*	-0.015*	-0.040*	-0.061*	-0.088*	-0.118*
Pac-Islr	-0.001	-0.013*	-0.040*	-0.081*	-0.147*	-0.236*	-0.226*
other	0.000	0.000	-0.002*	-0.005*	-0.007*	0.000	0.006*
multiple	-0.001*	-0.010*	-0.035*	-0.076*	-0.142*	-0.229*	-0.308*
Edu high school	0.000	0.002*	0.006*	0.013*	0.022*	0.029*	0.024*
some college	0.000	0.000	0.000	-0.002*	-0.006*	-0.008*	-0.023*
college	0.000	0.002*	0.006*	0.012*	0.019*	0.023*	0.020*
Time & state dummies	included	included	included	included	included	included	included
Pseudo R-squared	0.003	0.04	0.06	0.19	0.11	0.06	0.11
Zip code-days (1,000)	4,505	4,505	4,498	4,505	4,483	4,045	4,405
Transactions (1,000)	198,700	129,299	67,465	132,108	50,800	16,425	37,905

*Significant at 1%. Units of regression variables are defined in footnote 10.

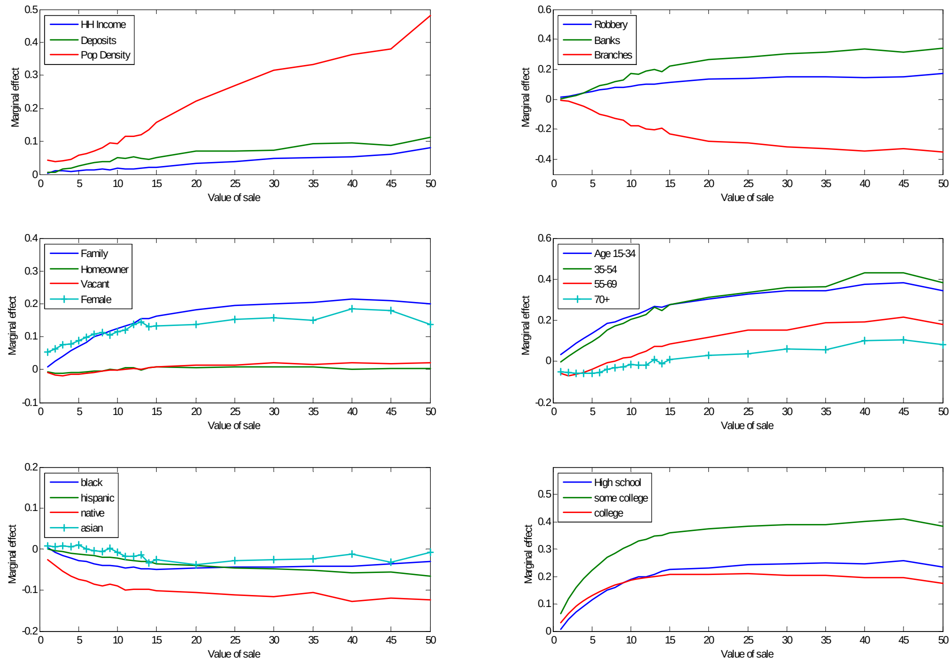

Figure A1. Debit marginal effects by transaction size.

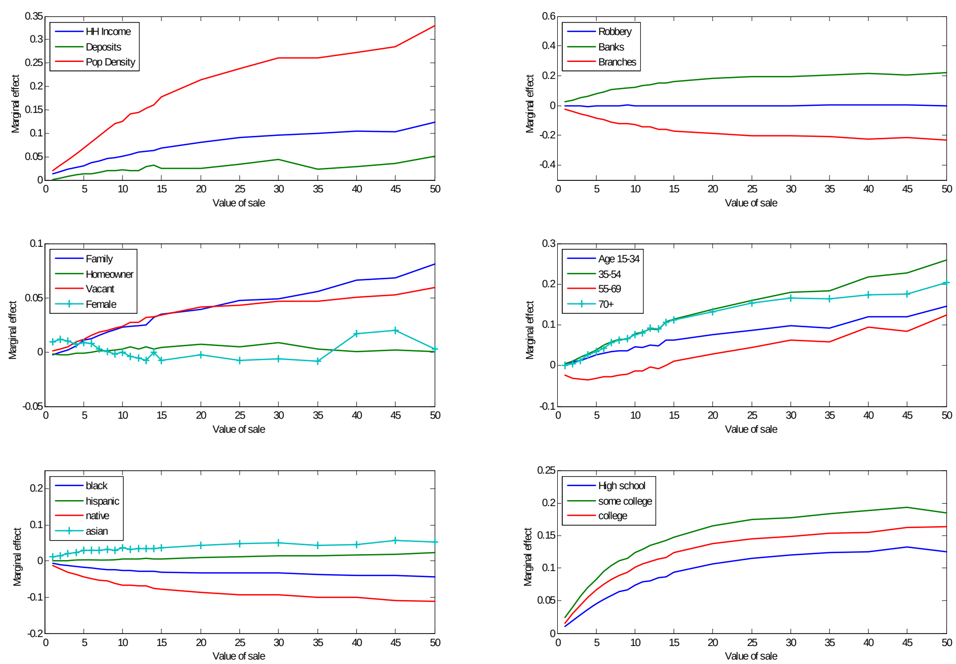

Figure A2. Credit marginal effects by transaction size.

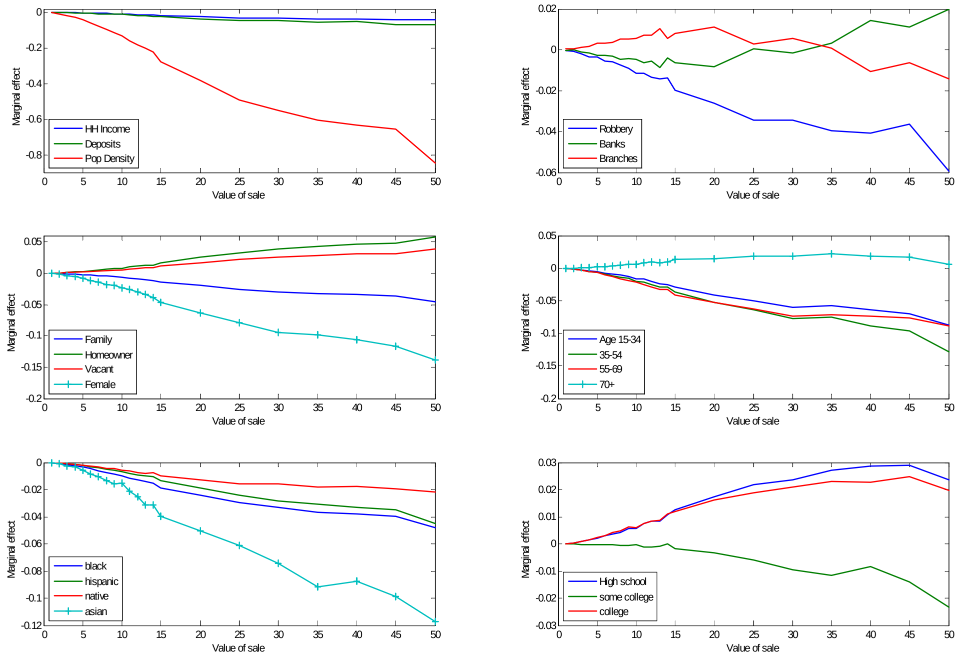

Figure A3. Check marginal effects by transaction size.

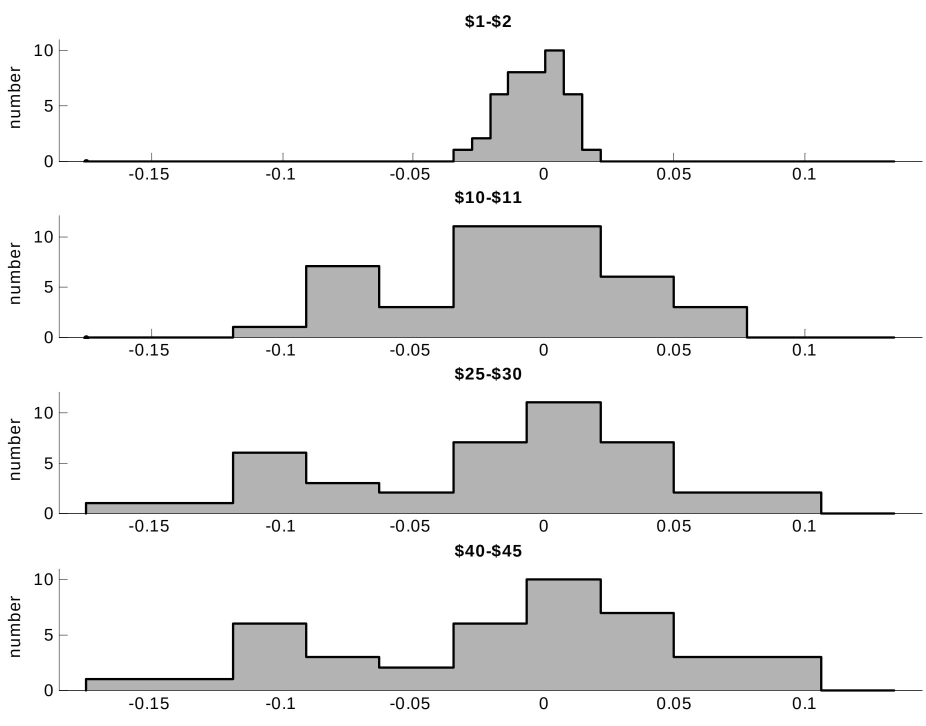

Figure A4. Debit: histograms of state effects by transaction size.

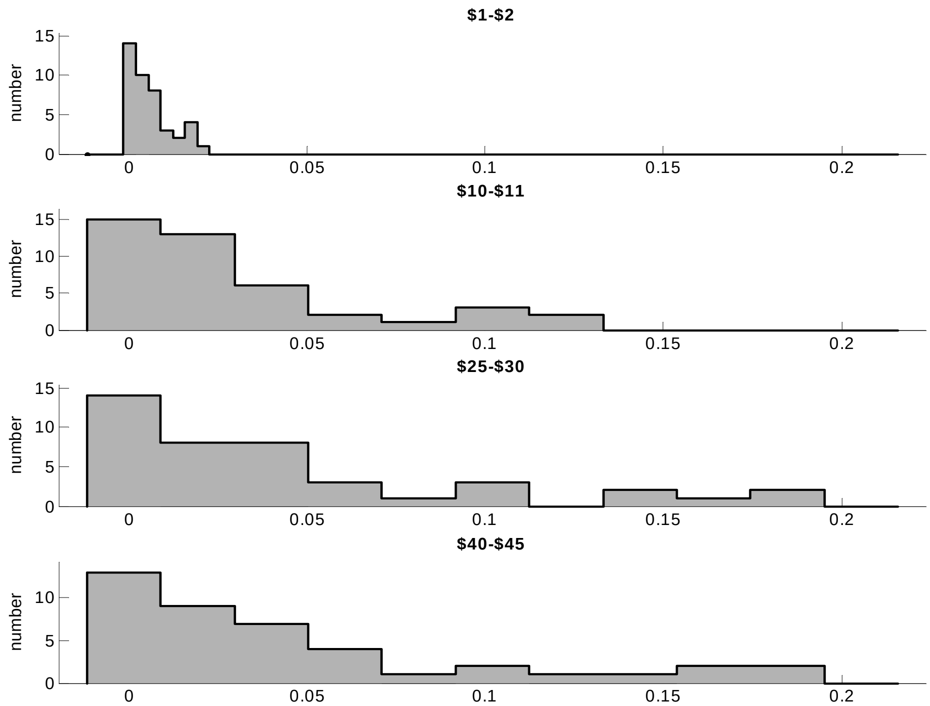

Figure A5. Credit: histograms of state effects by transaction size.

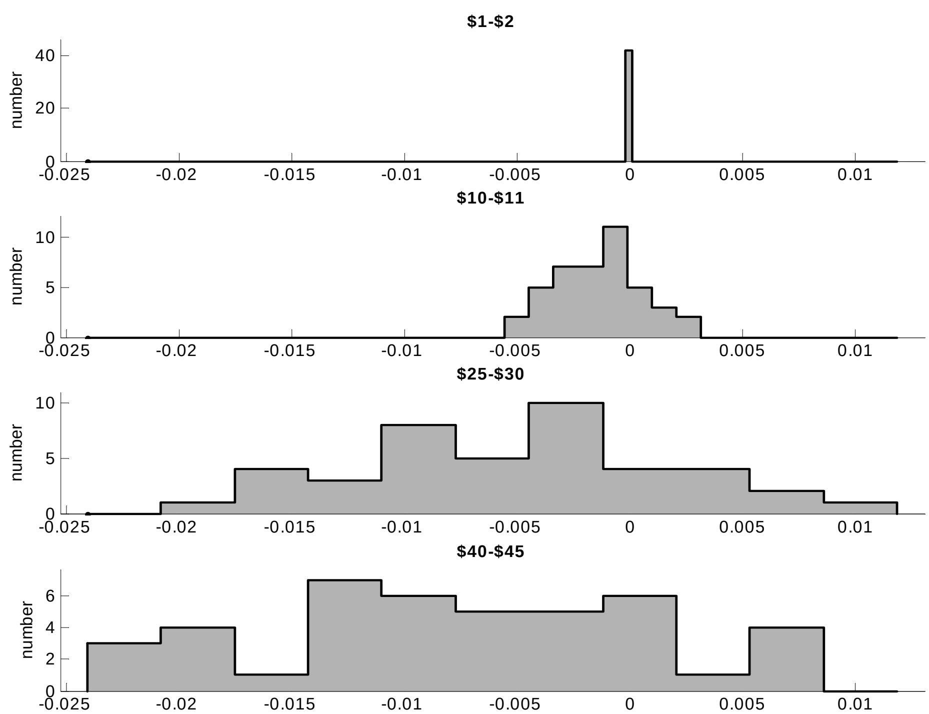

Figure A6. Check: histograms of state effects by transaction size.

Table A4. Debit: rankings of state effects

	\$1-\$2	\$10-\$11	\$25-\$30	\$40-\$45
Top States	Arizona	Arizona	Nevada	Nevada
	Nevada	Idaho	Arizona	Arizona
	New Mexico	Nevada	Idaho	Idaho
	Florida	New Mexico	New Mexico	New Mexico
	Idaho	Florida	Florida	Florida
Bottom States				
	Wisconsin	Maryland	North Dakota	Ohio
	Maryland	Ohio	Ohio	North Dakota
	North Dakota	New York	Oklahoma	Oklahoma
	South Dakota	South Dakota	South Dakota	South Dakota
	Minnesota	Minnesota	Minnesota	Minnesota

Table A5. Credit: rankings of state effects

	\$1-\$2	\$10-\$11	\$25-\$30	\$40-\$45
Top States	Ohio	North Dakota	Minnesota	Minnesota
	Kentucky	Minnesota	North Dakota	North Dakota
	Oklahoma	South Dakota	South Dakota	South Dakota
	Minnesota	Oklahoma	Oklahoma	Oklahoma
	South Dakota	Ohio	Ohio	Ohio
Bottom States				
	Alabama	Iowa	Nevada	New Jersey
	New Jersey	California	Arkansas	California
	Arkansas	Arkansas	Iowa	Arkansas
	California	New Jersey	New Jersey	Iowa
	Mississippi	Mississippi	Mississippi	Mississippi

Table A6. Check: rankings of state effects

	\$1-\$2	\$10-\$11	\$25-\$30	\$40-\$45
Top States	South Dakota	North Dakota	South Dakota	South Dakota
	North Dakota	South Dakota	North Dakota	Oklahoma
	Wyoming	Minnesota	Minnesota	North Dakota
	Minnesota	Wyoming	Colorado	Minnesota
	Colorado	Colorado	Oklahoma	Colorado
Bottom States				
	Florida	Pennsylvania	New Hampshire	New Hampshire
	New York	New York	New York	New York
	Arizona	Arizona	Arizona	Arizona
	Delaware	Delaware	Delaware	Delaware
	New Jersey	New Jersey	New Jersey	New Jersey

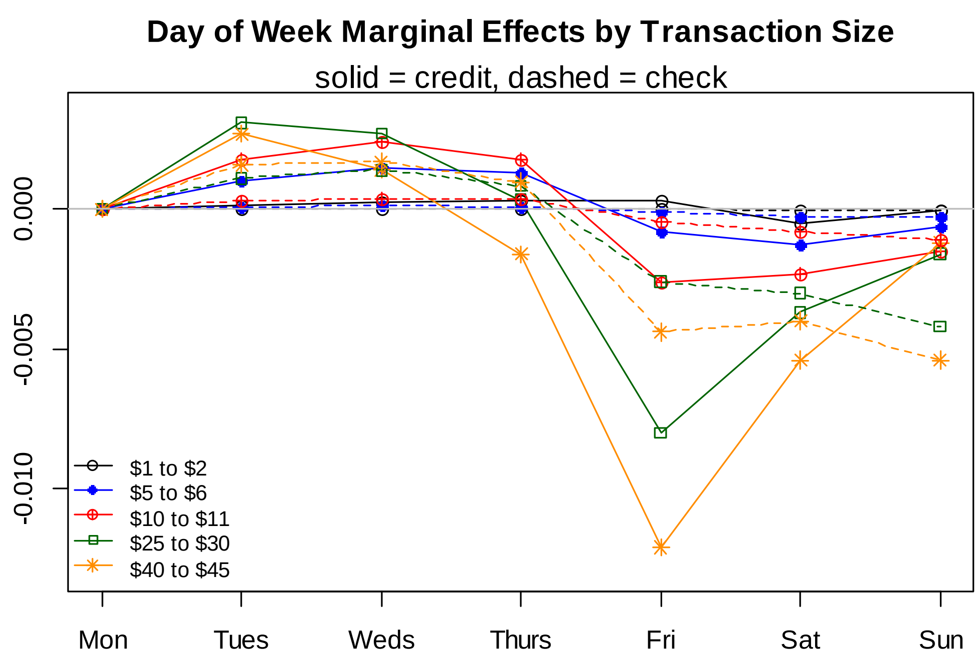

Figure A7.

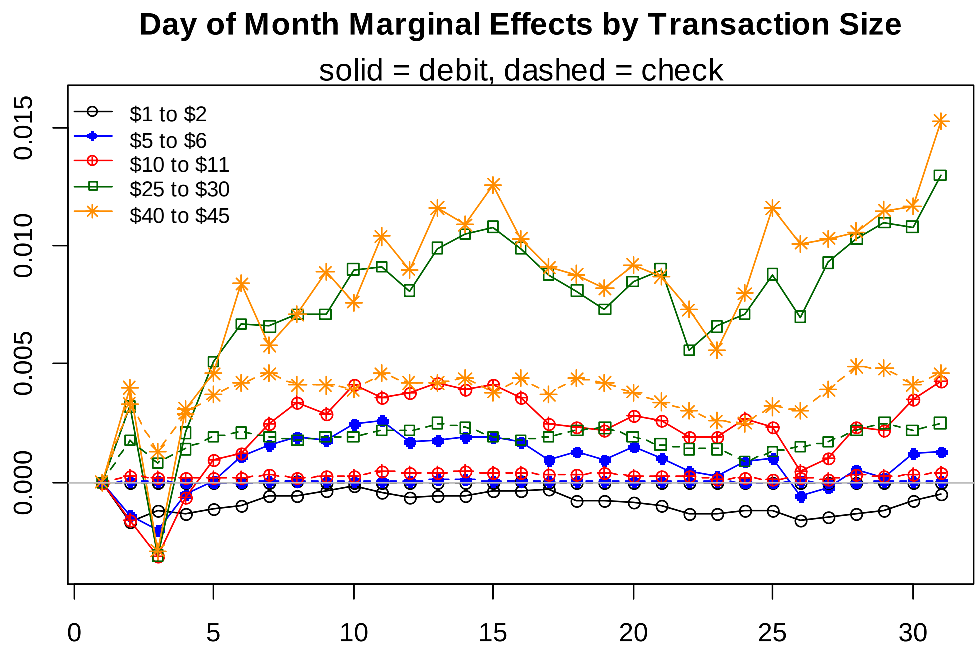

Figure A8.

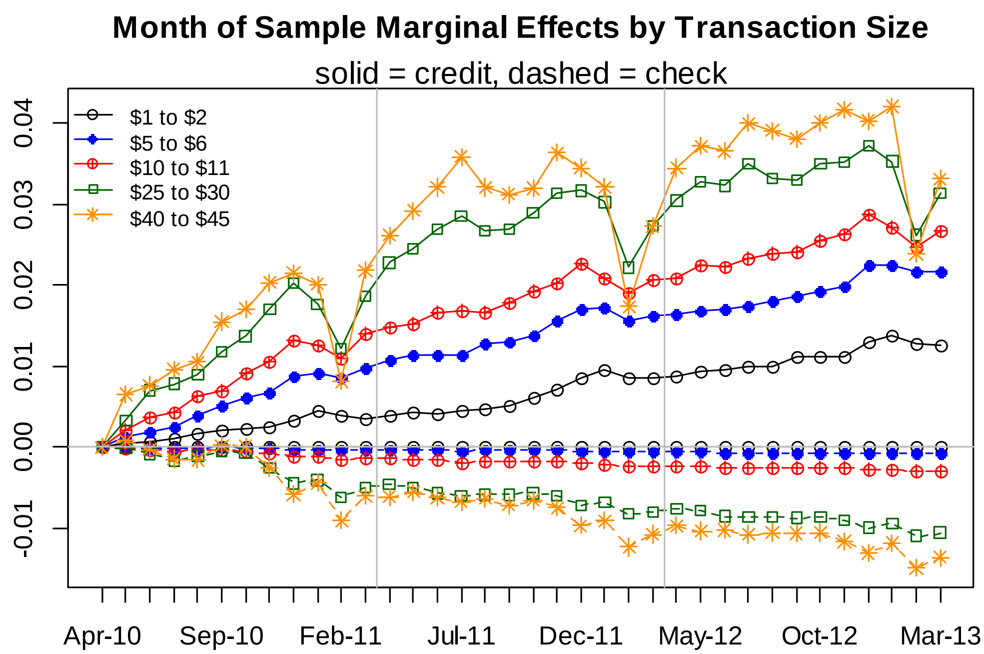

Figure A9.

Appendix B.

Robustness Checks: Cash vs. Non-cash Payments

As pointed out in Section 3, the FMLogit model is similar to the Multinomial logit model in the sense that they impose some restrictions on the substitution patterns between the categories of the dependent variables. For robustness checks, we re-group payment types into two: cash and non-cash (combining debit, credit and check) and re-run the FMLogit model based on the new categories. The results show that our findings on cash use reported in Section 3 remain essentially unchanged.

The results are shown in the following order.

- Table B1. Marginal effects for zip-code-level variables
- Figure B1. Histograms of state effects
- Table B2. Rankings of state effects
- Figure B2. Day of week marginal effects
- Figure B3. Day of month marginal effects
- Figure B4. Month of sample marginal effects

Table B1. Marginal effects for zip-code-level variables

Variable	Cash	Non-Cash
Cash holding and payment choice		
Median sale value	-0.018* (0.000)	0.018* (0.000)
Banks per capita	-0.229* (0.004)	0.229* (0.004)
Branches per capita	0.237* (0.004)	-0.237* (0.004)
Robbery rate	-0.062* (0.001)	0.062* (0.001)
Adoption of non-cash payments		
Median household income	-0.049* (0.000)	0.049* (0.000)
Deposits per capita	-0.042* (0.001)	0.042* (0.001)
Population density	-0.131* (0.001)	0.131* (0.001)
Demographics		
Family households	-0.121* (0.001)	0.121* (0.001)
Owner-occupied	0.001 (0.001)	-0.001 (0.001)
Vacant housing	-0.020* (0.001)	0.020* (0.001)
Female	-0.086* (0.001)	0.086* (0.001)
Age 15-34	-0.215* (0.002)	0.215* (0.002)
35-54	-0.207* (0.002)	0.207* (0.002)
55-69	0.016* (0.002)	-0.016* (0.002)
$\geq$ 70	-0.061* (0.002)	0.061* (0.002)
Race black	0.056* (0.000)	-0.056* (0.000)
Hispanic	0.027* (0.000)	-0.027* (0.000)
Native	0.143* (0.001)	-0.143* (0.001)
Asian	-0.006* (0.001)	0.006* (0.001)
Pac-Islr	-0.546* (0.011)	0.546* (0.011)
other	0.083* (0.001)	-0.083* (0.001)
multiple	-0.008* (0.003)	0.008* (0.003)
Edu high school	-0.205* (0.001)	0.205* (0.001)
some college	-0.328* (0.001)	0.328* (0.001)
college	-0.235* (0.001)	0.235* (0.001)
Time & State	included	included
Pseudo R-squared	0.59	0.59
Zip-day observations	4,505,642	4,505,642

Robust standard errors in parentheses. *Significant at 1%. Units of regression variables are defined in footnote 10.

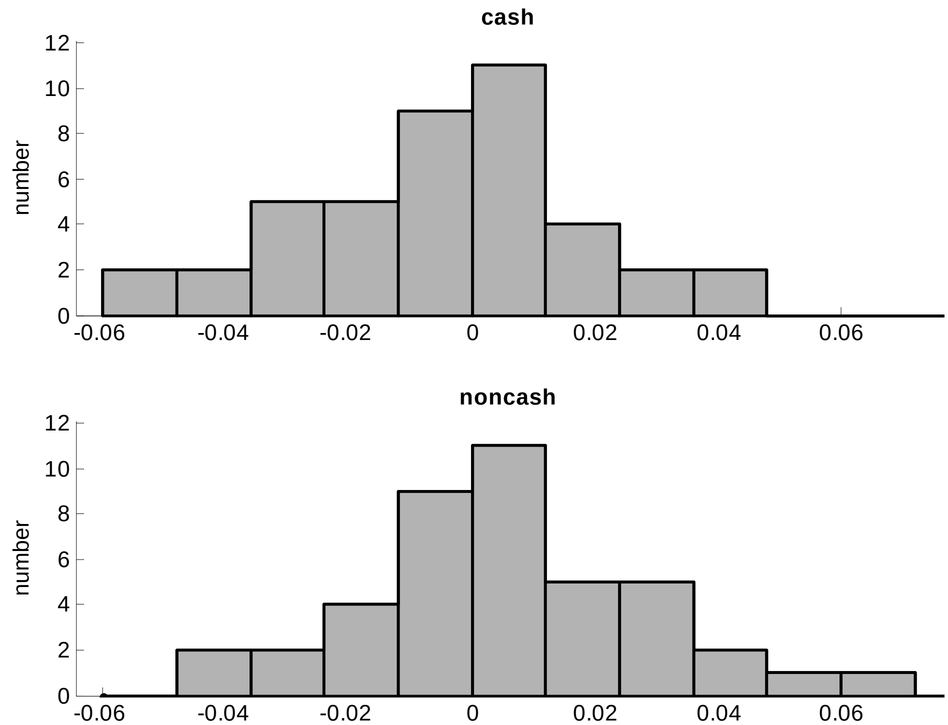

Figure B1. Histograms of state effects.

Table B2. Rankings of state effects

	Cash	Non-Cash
Top States		
	New York	Arizona
	New Jersey	Idaho
	Michigan	New Mexico
	Maryland	Texas
	Vermont	Nevada
Bottom States		
	Nevada	Vermont
	Texas	Maryland
	New Mexico	Michigan
	Idaho	New Jersey
	Arizona	New York

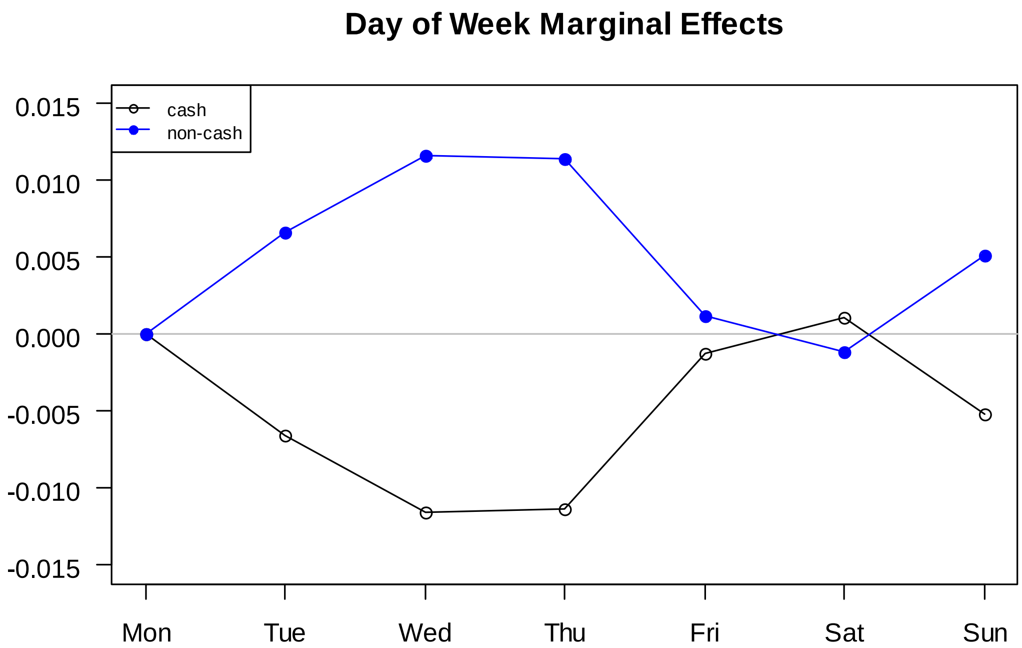

Figure B2.

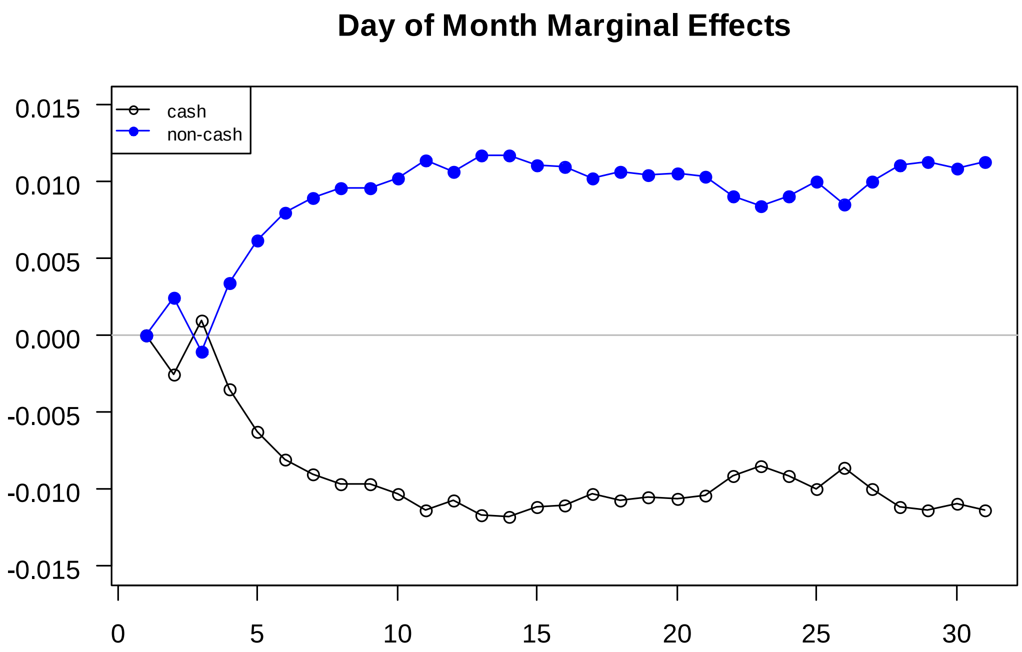

Figure B3.

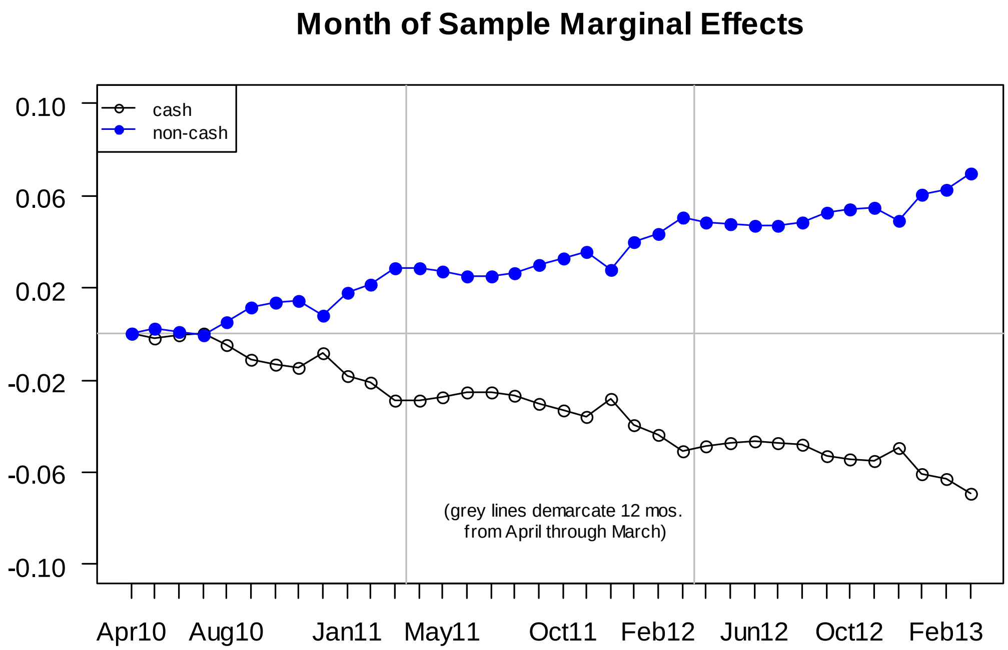
Month of Sample Marginal Effects
cash
non-cash
0.10
0.06
0.02
-0.02
-0.06
-0.10
(grey lines demarcate 12 mos.
from April through March)
Apr10
Aug10
Jan11
May11
Oct11
Feb12
Jun12
Oct12
Feb13

Figure B4.

Appendix C-D.

Robustness Checks: Transaction-level Regressions

As pointed out in the paper, considering our data set is so large, we do not work with the transaction-level data directly, instead aggregating it up to the fractions of transactions for each payment type on each day in each zip code. Moreover, in Section 4, we take an additional step to group our data by transaction size and estimate separate models for each group. In so doing, we directly incorporate the size of individual transactions into the analysis, and also allow all coefficient estimates to vary across transactions of different sizes. In terms of estimation, we use the fractional multinomial logit model (FMLogit), which specifically handles the fractional multinomial nature of our dependent variables.

While our approach has its advantages, it would be useful to compare our estimation results with transaction-level regressions for robustness checks. This is feasible by using subsamples. We conduct the following two sets of experiments, reported in the order of Appendix C and D. In the first one, we run the transaction-level regression using the multinomial logit model (MLogit) on a randomly selected subsample of 4.4 million transactions in our three-year data set. The results are shown to be comparable with our FMLogit findings in Section 3. In the second one, we make a direct comparison between the payment-share FMLogit model and the transaction-level MLogit model by using the exact same subsample, which includes all transactions with size of $6-$7 in March, 2013 (about 3.4 million transactions). The results are again very much consistent. The results are shown as follows.

Appendix C:

- Table C1. MLogit on a random sample: marginal effects
- Figure C1. Histograms of state effects: MLogit on a random sample
- Figure C2. Day of week marginal effects: MLogit on a random sample
- Figure C3. Day of month marginal effects: MLogit on a random sample
- Figure C4. Month of sample marginal effects: MLogit on a random sample

Appendix D:

- Table D1. Payment-share regression (FMLogit): marginal effects, $6-$7, March 2013
- Table D2. Transaction regression (MLogit): marginal effects, $6-$7, March 2013
- Figure D1. Histograms of state effects: FMLogit, $6-$7, March 2013
- Figure D2. Histograms of state effects: MLogit, $6-$7, March 2013
- Figure D3. Day of sample marginal effects: FMLogit, $6-$7, March 2013
- Figure D4. Day of sample marginal effects: MLogit, $6-$7, March 2013

Table C1. MLogit on a random sample: marginal effects

Variable	Cash	Debit	Credit	Check
Cash holding and payment choice				
Transaction amount	-0.006* (0.000)	0.004* (0.000)	0.001* (0.000)	0.000* (0.000)
Banks per capita	-0.216* (0.022)	0.140* (0.019)	0.076* (0.010)	0.000 (0.002)
Branches per capita	0.228* (0.022)	-0.149* (0.019)	-0.080* (0.010)	0.001 (0.002)
Robbery rate	-0.083* (0.006)	0.070* (0.005)	0.014* (0.003)	-0.001 (0.001)
Adoption of non-cash payments				
Median household income	-0.068* (0.004)	0.037* (0.003)	0.036* (0.002)	-0.005* (0.000)
Deposits per capita	-0.049* (0.011)	0.040* (0.010)	0.019* (0.005)	-0.010* (0.002)
Population density	-0.057* (0.010)	0.052* (0.009)	0.070* (0.005)	-0.065* (0.002)
Demographics				
Family households	-0.116* (0.006)	0.098* (0.006)	0.022* (0.003)	-0.003* (0.001)
Owner-occupied	0.020* (0.004)	-0.015* (0.004)	-0.010* (0.002)	0.004* (0.000)
Vacant housing	-0.015* (0.004)	0.002 (0.004)	0.010* (0.002)	0.003* (0.000)
Female	-0.124* (0.013)	0.137* (0.011)	0.000 (0.006)	-0.013* (0.001)
Age 15-34	-0.223* (0.014)	0.198* (0.012)	0.032* (0.006)	-0.007* (0.001)
35-54	-0.282* (0.020)	0.232* (0.018)	0.058* (0.009)	-0.008* (0.002)
55-69	-0.036 (0.014)	0.034* (0.013)	0.009 (0.007)	-0.007* (0.001)
$\geq$ 70	-0.057* (0.016)	0.015 (0.015)	0.038* (0.008)	0.004 (0.002)
Race black	0.073* (0.001)	-0.045* (0.001)	-0.022* (0.001)	-0.006* (0.000)
Hispanic	0.015* (0.002)	-0.013* (0.002)	0.001 (0.001)	-0.004* (0.000)
Native	0.126* (0.006)	-0.076* (0.005)	-0.045* (0.003)	-0.004* (0.000)
Asian	0.021 (0.010)	-0.021 (0.009)	0.012* (0.004)	-0.012* (0.002)
Pac-Islr	-0.573* (0.079)	0.703* (0.066)	-0.115 (0.047)	-0.014 (0.008)
other	0.045* (0.006)	-0.016* (0.005)	-0.029* (0.003)	0.000 (0.001)
multiple	0.040 (0.025)	0.011 (0.023)	-0.030* (0.011)	-0.021* (0.003)
Edu high school	-0.208* (0.006)	0.148* (0.006)	0.055* (0.003)	0.005* (0.001)
some college	-0.374* (0.006)	0.269* (0.006)	0.103* (0.003)	0.002* (0.001)
college	-0.211* (0.005)	0.136* (0.005)	0.071* (0.002)	0.004* (0.000)
Time & State	included	included	included	included
Pseudo R-squared	0.07	0.01	0.01	0.02
Transactions	4,403,595	4,403,595	4,403,595	4,403,595

Robust standard errors in parentheses. *Significant at 1%. Units of regression variables are defined in footnote 10 .

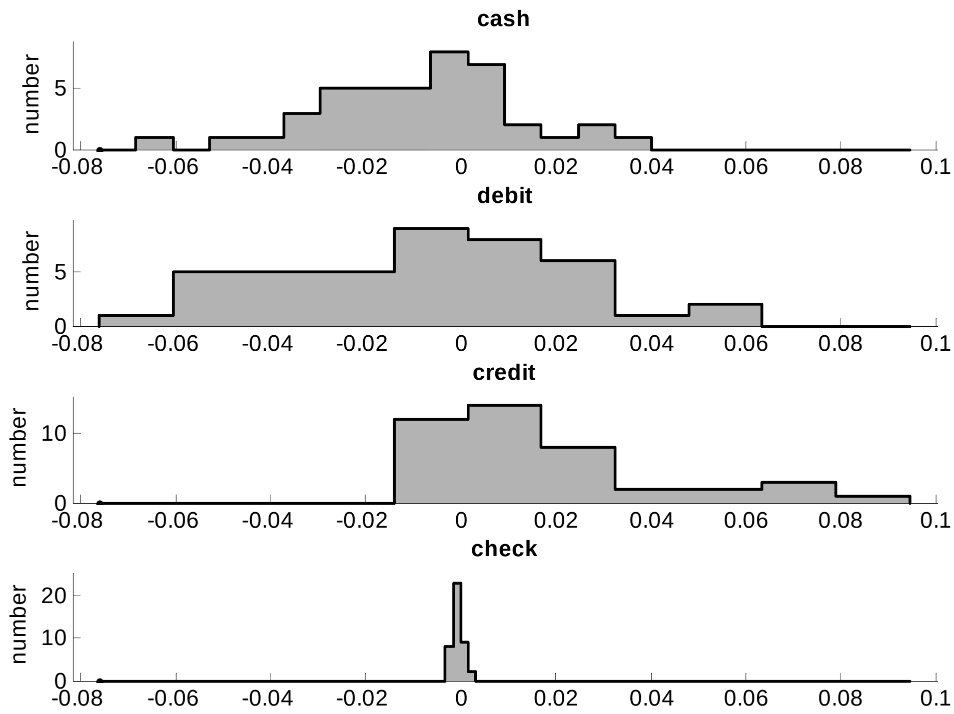

Figure C1. Histograms of state effects: MLogit on a random sample.

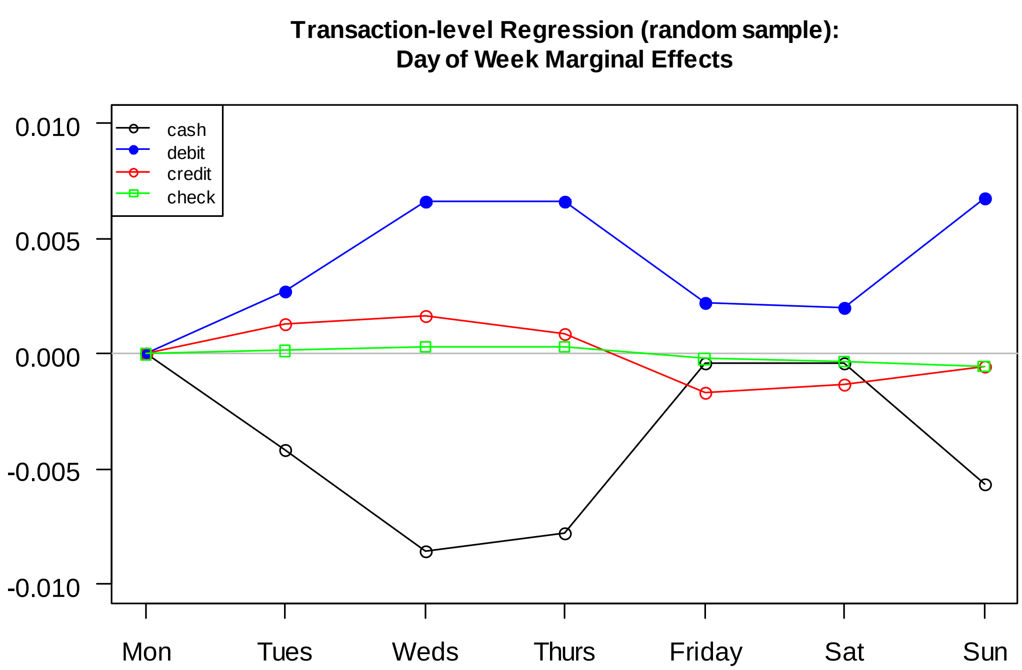

Figure C2.

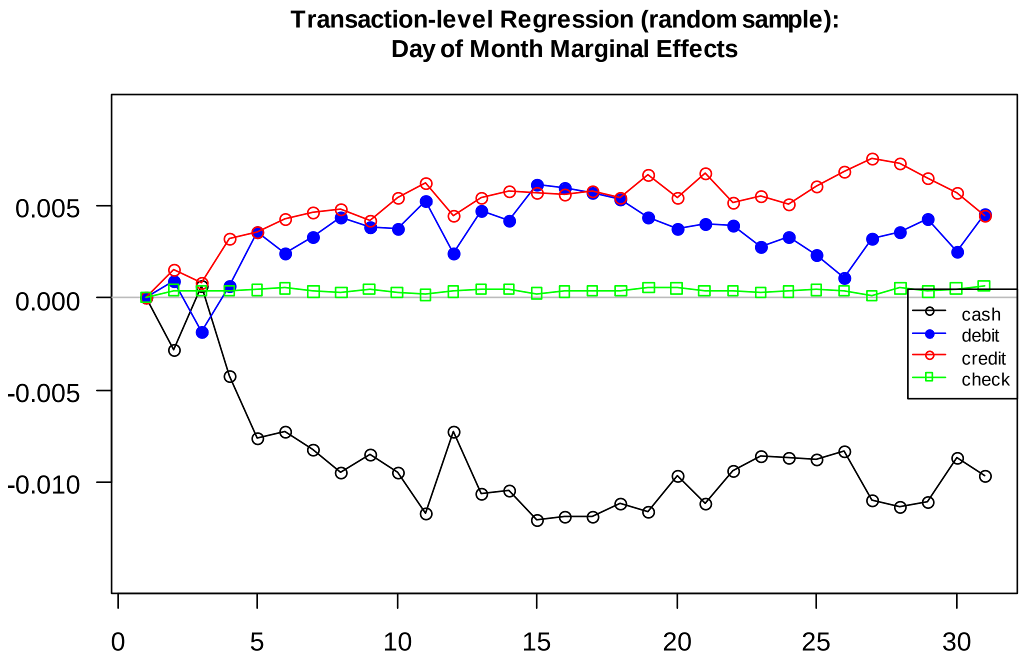

Figure C3.

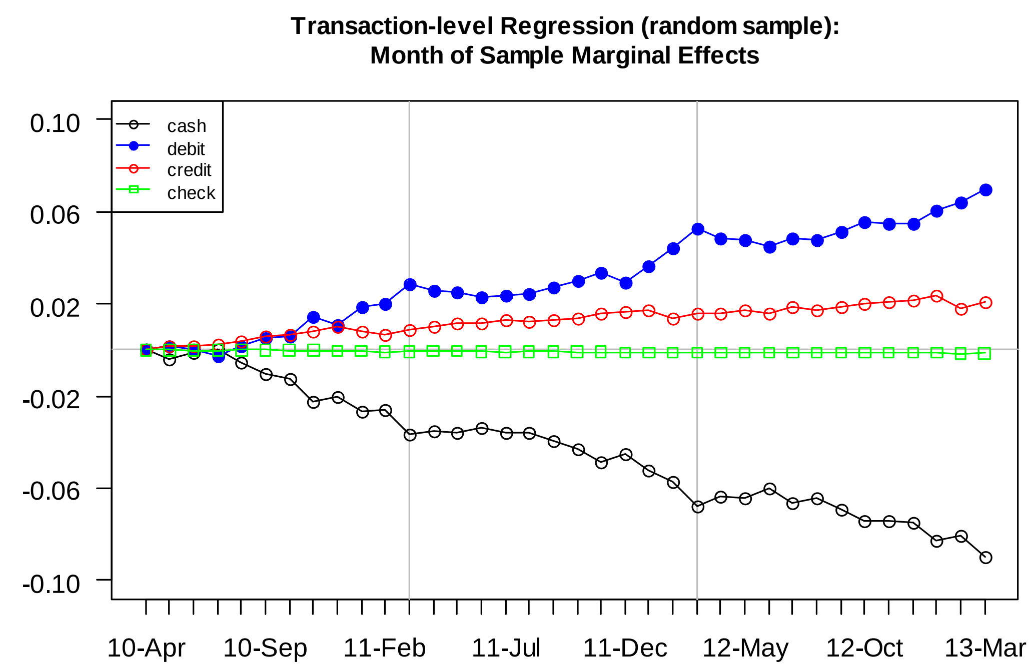

Figure C4.

Table D1. Payment-share regression (FMLogit): marginal effects
($6-$7, March 2013)

Variables	Cash	Debit	Credit	Check
Cash holding and payment choice				
Banks per capita	-0.212* (0.040)	0.084 (0.034)	0.131* (0.019)	-0.003 (0.002)
Branches per capita	0.218* (0.040)	-0.089* (0.035)	-0.133* (0.019)	0.004 (0.002)
Robbery rate	-0.022 (0.011)	0.053* (0.010)	-0.027* (0.006)	-0.004 (0.002)
Adoption of non-cash payments				
Median household income	-0.044* (0.007)	0.012 (0.006)	0.035* (0.003)	-0.003* (0.000)
Deposits per capita	-0.017 (0.018)	0.031 (0.015)	-0.005 (0.009)	-0.009* (0.002)
Population density	-0.124* (0.017)	0.081* (0.015)	0.079* (0.008)	-0.036* (0.003)
Demographics				
Family households	-0.083* (0.011)	0.084* (0.010)	0.002 (0.005)	-0.002 (0.001)
Owner-occupied	-0.003 (0.007)	-0.001 (0.006)	0.002 (0.004)	0.002* (0.001)
Vacant housing	0.000 (0.008)	-0.007 (0.007)	0.006 (0.004)	0.001 (0.001)
Female	-0.079* (0.021)	0.117* (0.019)	-0.029* (0.011)	-0.008* (0.002)
Age 15-34	-0.181* (0.024)	0.184* (0.022)	0.004 (0.012)	-0.008* (0.002)
35-54	-0.102* (0.033)	0.090* (0.030)	0.019 (0.017)	-0.007* (0.003)
55-69	0.058 (0.024)	-0.025 (0.022)	-0.027 (0.012)	-0.007* (0.002)
≥ 70	0.084* (0.028)	-0.120* (0.025)	0.036 (0.014)	0.001 (0.002)
Race black	0.067* (0.002)	-0.040* (0.002)	-0.025* (0.001)	-0.003* (0.000)
Hispanic	0.015* (0.004)	-0.014* (0.003)	0.001 (0.002)	-0.002* (0.000)
Native	0.138* (0.009)	-0.079* (0.007)	-0.057* (0.005)	-0.002* (0.000)
Asian	-0.020 (0.017)	-0.007 (0.015)	0.031* (0.008)	-0.003 (0.002)
Pac-Islr	-0.313 (0.134)	0.600* (0.110)	-0.274* (0.084)	-0.014 (0.011)
other	0.102* (0.010)	-0.053* (0.009)	-0.048* (0.006)	-0.001 (0.001)
multiple	-0.236* (0.042)	0.231* (0.039)	0.012 (0.020)	-0.006 (0.003)
Edu high school	-0.214* (0.011)	0.145* (0.010)	0.067* (0.005)	0.002* (0.001)
some college	-0.369* (0.011)	0.270* (0.010)	0.100* (0.006)	-0.001 (0.001)
college	-0.251* (0.009)	0.157* (0.008)	0.092* (0.004)	0.002* (0.001)
Time & state dummies	included	included	included	included
Pseudo R-squared	0.12	0.16	0.14	0.03
Zip code-days	137,082	137,082	137,082	137,082
Transactions	3,351,579	3,351,579	3,351,579	3,351,579

Robust standard errors in parentheses. *Significant at 1%. Units of variables are defined in footnote 10.

Table D2. Transaction regression (MLogit): marginal effects
($6-$7, March 2013)

Variables	Cash	Debit	Credit	Check
Cash holding and payment choice				
Banks per capita	-0.176* (0.026)	0.094* (0.024)	0.084* (0.012)	-0.002* (0.001)
Branches per capita	0.184* (0.026)	-0.101* (0.024)	-0.086* (0.012)	0.003* (0.001)
Robbery rate	-0.038* (0.007)	0.046* (0.007)	-0.007 (0.004)	-0.001 (0.001)
Adoption of non-cash payments				
Median household income	-0.057* (0.005)	0.029* (0.004)	0.030* (0.002)	-0.002* (0.000)
Deposits per capita	-0.042* (0.014)	0.043* (0.013)	0.004 (0.007)	-0.005* (0.001)
Population density	-0.083* (0.012)	0.046* (0.011)	0.056* (0.005)	-0.018* (0.001)
Demographics				
Family households	-0.067* (0.008)	0.067* (0.007)	0.001 (0.003)	-0.001 (0.000)
Owner-occupied	0.021* (0.005)	-0.016* (0.005)	-0.006 (0.002)	0.001* (0.000)
Vacant housing	0.022* (0.006)	-0.021* (0.005)	-0.001 (0.003)	0.000 (0.000)
Female	-0.138* (0.015)	0.174* (0.014)	-0.031* (0.007)	-0.004* (0.001)
Age 15-34	-0.159* (0.017)	0.172* (0.016)	-0.010 (0.008)	-0.003* (0.001)
35-54	-0.147* (0.024)	0.157* (0.022)	-0.007 (0.011)	-0.003* (0.001)
55-69	0.048* (0.017)	-0.028 (0.016)	-0.019 (0.008)	-0.002 (0.001)
≥ 70	0.102* (0.020)	-0.113* (0.019)	0.010 (0.009)	0.001 (0.001)
Race black	0.056* (0.002)	-0.032* (0.001)	-0.023* (0.001)	-0.002* (0.000)
Hispanic	0.012* (0.003)	-0.010* (0.002)	-0.001 (0.001)	-0.001* (0.000)
Native	0.118* (0.007)	-0.076* (0.006)	-0.041* (0.004)	-0.001* (0.000)
Asian	-0.027 (0.012)	0.004 (0.011)	0.024* (0.005)	-0.002 (0.001)
Pac-Islr	-0.487* (0.097)	0.615* (0.086)	-0.128 (0.054)	-0.001 (0.005)
other	0.064* (0.007)	-0.031* (0.006)	-0.033* (0.003)	0.000 (0.000)
multiple	-0.211* (0.031)	0.230* (0.029)	-0.014 (0.013)	-0.005* (0.001)
Edu high school	-0.208* (0.008)	0.147* (0.007)	0.059* (0.003)	0.002* (0.000)
some college	-0.396* (0.008)	0.296* (0.007)	0.101* (0.004)	0.000 (0.000)
college	-0.222* (0.006)	0.146* (0.006)	0.075* (0.003)	0.001* (0.000)
Time & state dummies	included	included	included	included
Pseudo R-squared	0.010	0.002	0.001	0.003
Transactions	3,351,579	3,351,579	3,351,579	3,351,579

Robust standard errors in parentheses. *Significant at 1%. Units of variables are defined in footnote 10.

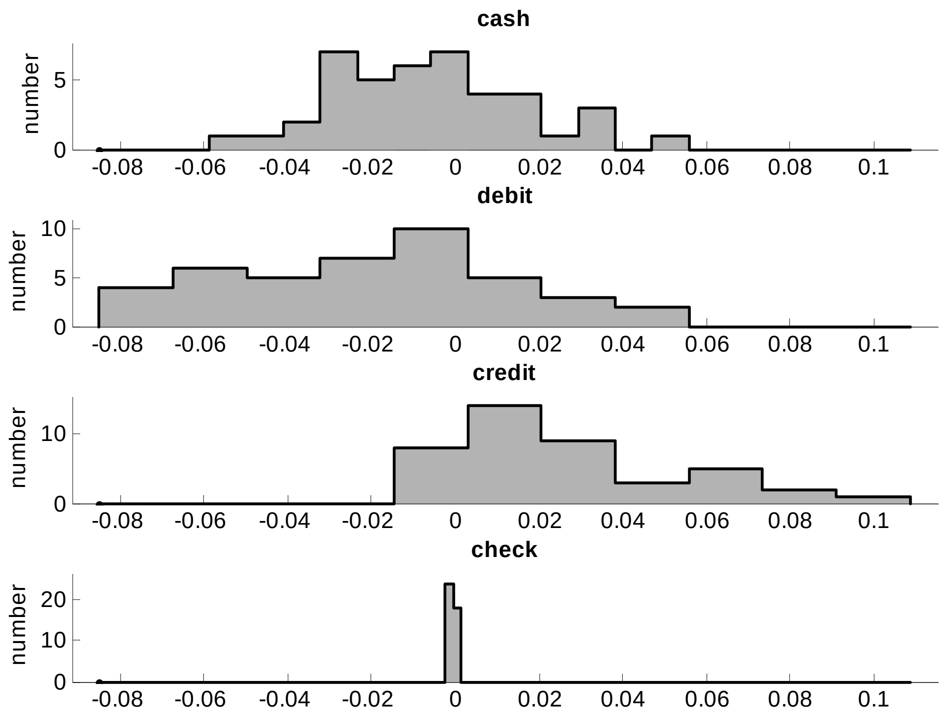

Figure D1. Histograms of state effects: FMLogit ($6-$7, March 2013).

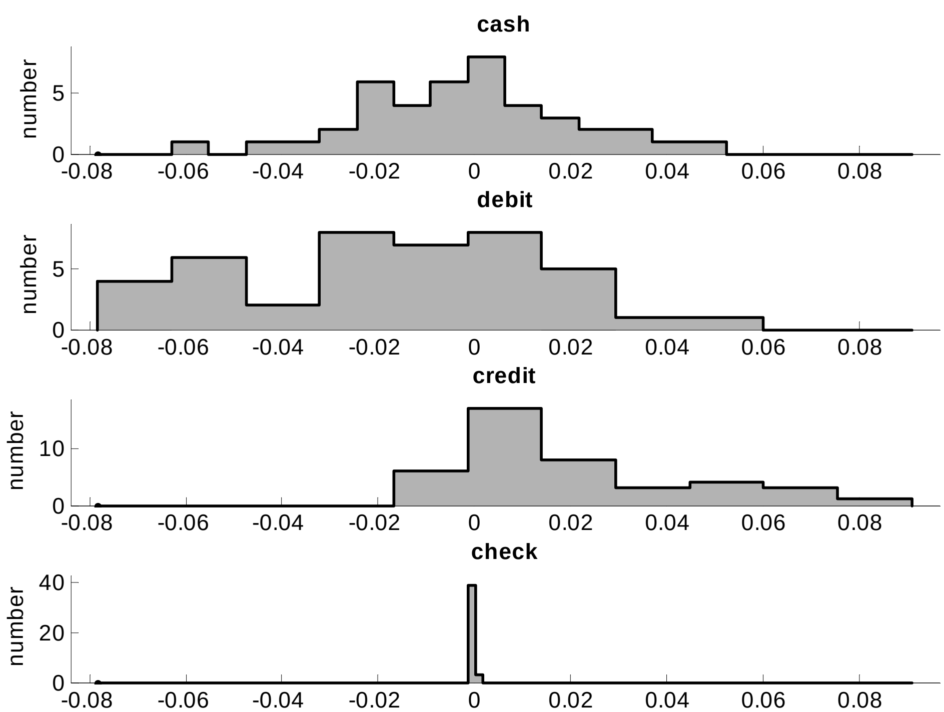

Figure D2. Histograms of state effects: MLogit ($6-$7, March 2013).

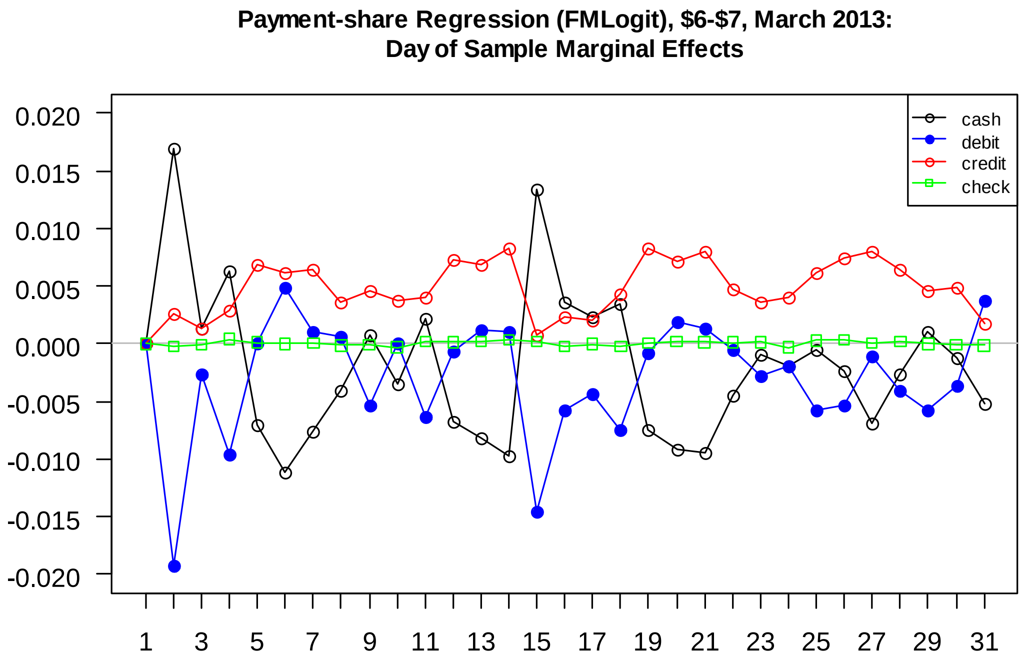

Figure D3.

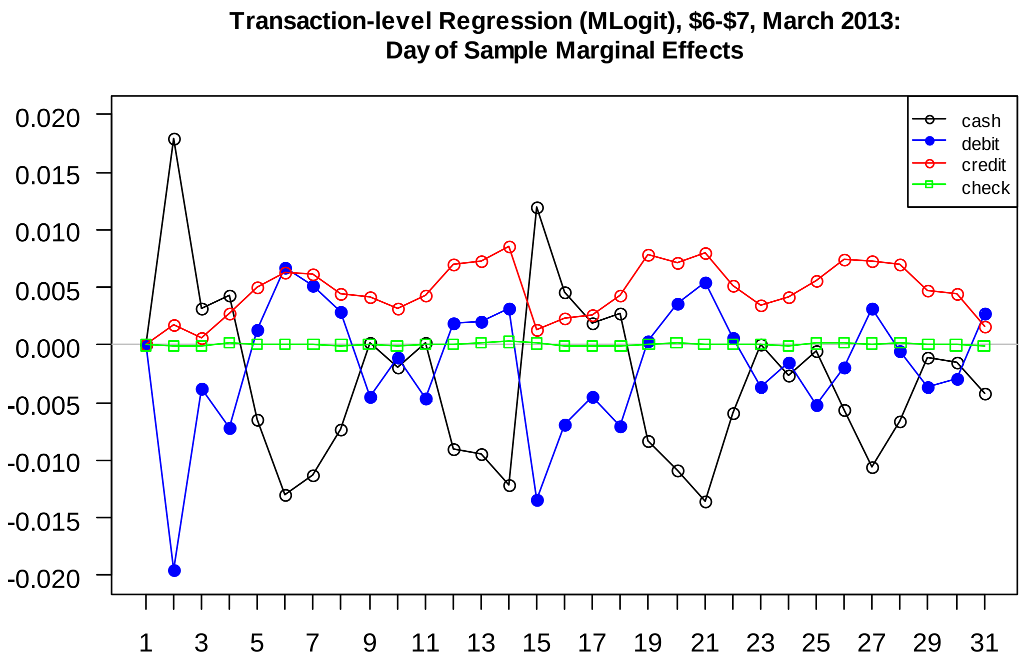

Figure D4.