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

Zhu Wang, Alexander L. Wolman*†

May, 2014

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 size of transaction, location, and time. 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 significant state-level variation that is not accounted for by the explanatory variables. We use the estimated model to forecast how the mix of consumer payments will evolve. Our estimates based on this large retailer, together with forecasts for the explanatory variables, lead to a benchmark prediction that the cash share of retail sales will decline by 2.54 percentage points per year over the next several years.

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 has provided outstanding research assistance, and Liz Marshall and Sabrina Pellerin provided us with valuable input.

†For helpful comments, we would like to thank Dave Beck, Marc Rysman, Mark Watson, and other colleagues.

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 cards 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 goal of understanding and evaluating payments system functioning. 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. As a result, this research has greatly deepened our understanding of how consumers choose to pay. On the other hand, 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, debit, and check use, based on a comprehensive dataset directly from 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 ten 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 the relationships between location-specific explanatory variables and payment choice. We also estimate time patterns of payment use associated with day of week, day of month, seasonal cycles and a trend. By combining these estimates with projections for the explanatory variables, we are able to project future use of currency in transactions, which can provide a benchmark for forecasting the future demand for currency.

A natural reference point for our work is Klee (2008), which also studied consumer payment choices at retail outlets. 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 versus 90 days. As a result, we are able to investigate the aforementioned time effects as well as state fixed effects that are not addressed in her study. We also assemble a larger set of demographic, banking, and other variables, which we identify with systematic variation in consumer payment choices. In addition, our analytical approach differs from Klee. Because our data set is so large (approximately 2 billion transactions), we do not work with the transaction data directly, instead aggregating it up to the fractions of transactions for each payment type on each day in each zip code. Moreover, we take an additional step and split our data to study payment choice

[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).

across different transaction sizes. In terms of estimation, we use the fractional multinomial logit model, which specifically handles the fractional multinomial nature of our dependent variables.

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).[2] While our data cannot 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 provides a relatively good fit to the data. For the zip-code-level variables, some of our main results are as follows. A large presence of bank branches and a population that is heavily black, Hispanic or Native American are associated with a high fraction of cash transactions and a low fraction of debit and check transactions. These effects tend to be larger for larger payment sizes; we refer to this property as amplification. On the other hand, zip codes with higher median income, more banks per capita, a higher robbery rate, and a relatively educated population are associated with a lower fraction of cash transactions, and again the effects tend to be amplified by transaction size. We also find significant residual state-level variation in the payment mix; 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 card use. Turning to the time effects, there are interesting patterns for day-of-week, day-of-month, and month-of-sample. For each of those frequencies we include dummy variables. Over the course of the week, cash and debit use are nearly mirror images of each other. Overall cash use peaks on Monday and Saturday, but the intra-week pattern differs markedly across transaction size, with cash use for the largest transactions peaking on Friday. Within the month, it is credit that comes closer to mirroring cash use, although the day-of-month dummies for credit and debit are highly correlated. Finally, our month-of-sample dummies 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. We also use our estimated coefficients to project the composition of payments beyond our sample. Taking into account the size distribution of payments as well

[2] According to *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.

as forecasts for the explanatory variables, we project that the cash fraction of transactions will decline by 2.54 percentage points per year. These results can also be used to assess whether the *level* of cash use in retail transactions will increase or decrease. Our answers depend 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 now 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 benchmark empirical model and estimated marginal effects. 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

Our data is from a large discount retailer with thousands of stores, covering a majority of U.S. states. The zip-code-level data discussed below and used in our empirical analysis covers most of those stores' zip codes. Our summary of the data in this section will refer to all stores for the time period of our sample, which is April 2010 through March 2013. The unit of observation is a transaction. For each transaction, we see the time, location, amount, and means of payment. Our study will focus on the four general-purpose means of payment: cash, debit cards, credit cards, and checks. We exclude special-purpose means of payment such as EBT, coupons and store return cards. The retailer also provides cash back services. We include only transactions that consist of a sale of goods, with one payment type used, where the payment type is cash, credit, debit, or check.[3] We do include transactions with cash back as long as they fit this description, treating the transaction size as the sale amount of goods involved in the transaction. All told, we cover more than 93% of the transactions in our sample period.

2.1 Transactions Data

Figure 1 summarizes the data at the daily level, displaying the fraction of payments 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 means of payment at this retailer, although its use is trending down, with debit trending up. There seems to be a weekly pattern in both cash and debit use, with the two moving in opposite directions. Credit displays a monthly pattern, rising over

[3] As in Klee (2008), the transactions we classify as credit card may include some signature debit card and prepaid 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.

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 across zip codes to focus on time variation. We turn now to the variation across zip codes. Figure 2 restricts attention to the last full month of the sample, March 2013, 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 nonstationarity 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 provides the motivation for including location-specific variables in our econometric model.

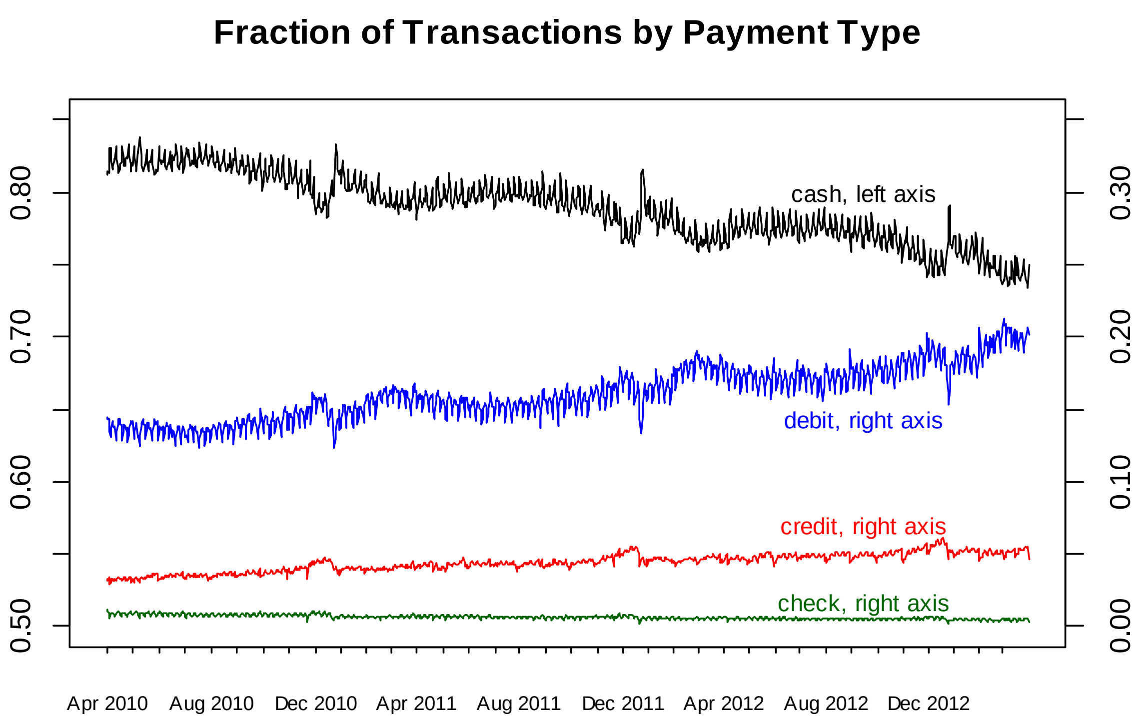

Figure 1.

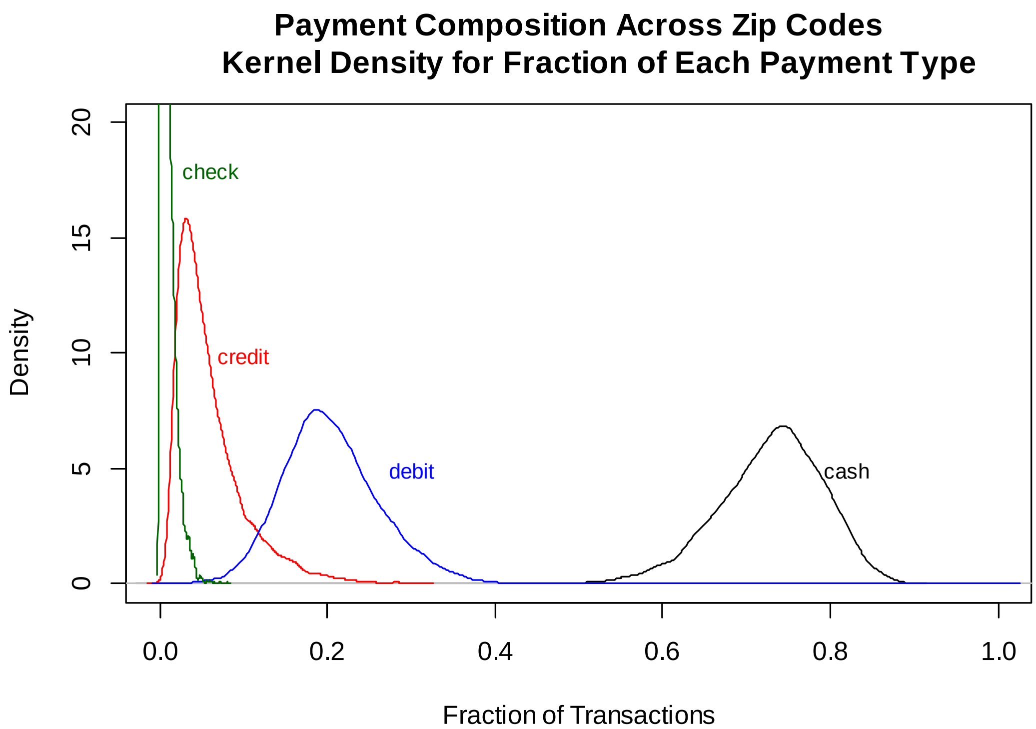

Figure 2.

Figures 3.a and 3.b provide information about the distribution of size of transactions, again restricting attention to March 2013. Figure 3.a 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 3.b plots the distribution of median transaction sizes across zip codes and days, also for March 2013 (representing 178,315 zip code days). Figure 3.b 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.

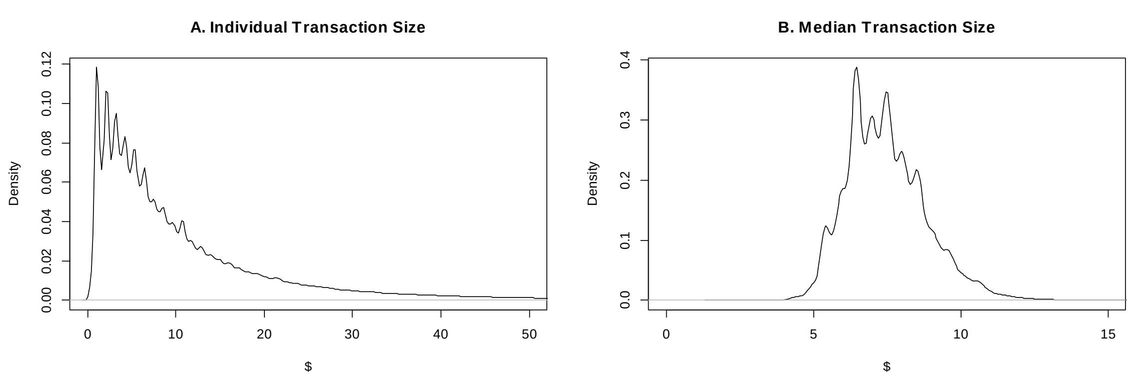

Figure 3. Kernel Densities of transaction size in March 2013.

Figure 4 displays information about the distribution of payment types across transaction sizes for March 2013. For each payment type, the solid line represents the median across zip codes of the fraction of payments in each size bin, and the dashed lines represent the 5th and 95th percentiles of the distribution. We use \$1 bins between \$1 and \$15, \$5 bins between \$15 and \$50, and then combine all transactions about \$50 into one bin. These categories were chosen to ensure a sufficient number of transactions in each bin. The top left panel shows that for transactions \$1 and below, the median zip code had 93 percent of payments made in cash, and, notably, even for transactions in the \$50 range the median zip code 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. Given the clear convenience benefits of using noncash payment forms for large transactions, the prevalence of cash in our data suggests that a significant fraction of this retailer's customer base may not have access to other means of payment – i.e. they may be unbanked. However, the prevalence of cash also renders the data especially revealing about the trend in cash use. A final feature of Figure 4 worth noting is the fanning out of the 5th and 95th percentiles. For large transaction sizes, the behavior of cardholders likely becomes increasingly different from the behavior of non-cardholders. Thus, variation in the fraction of cardholders across zip codes may be one factor contributing to the fanning out. For very large transaction sizes, fanning out may also be an artifact of a relatively small number of underlying transactions.

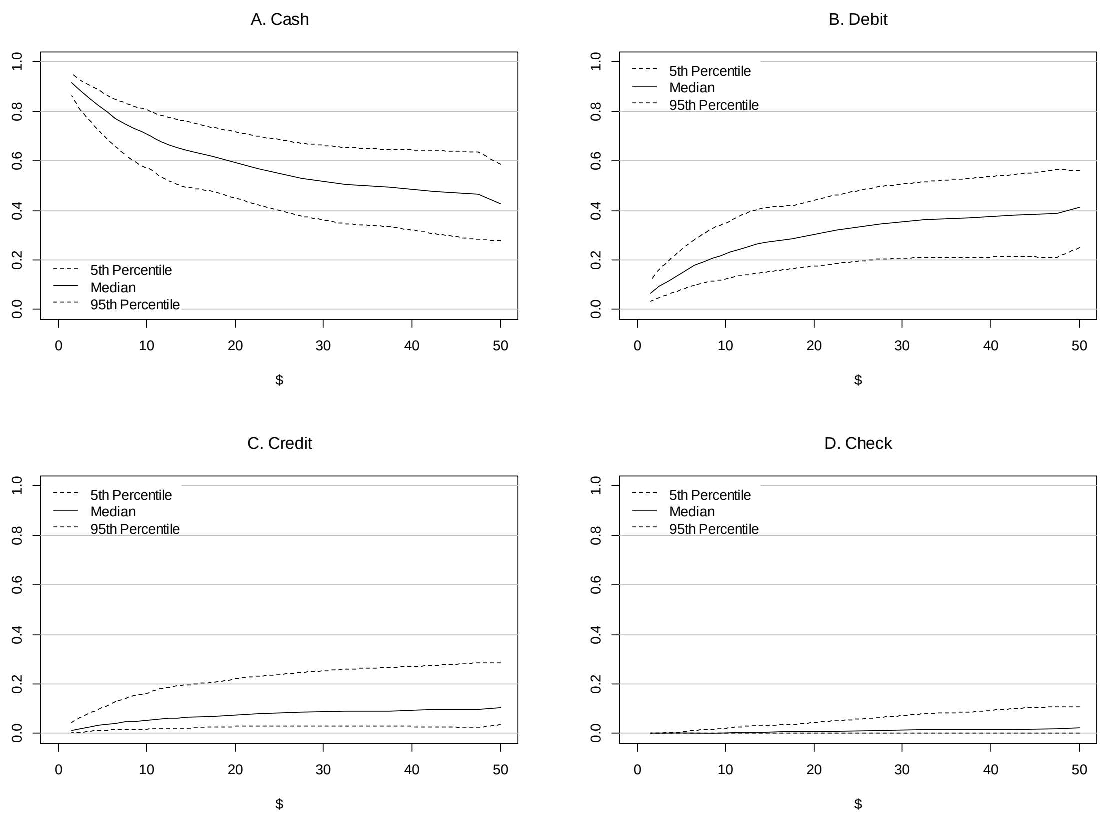

Figure 4. Payment mix by transaction size, across zip codes.

2.2 Zip-code-level Explanatory Variables

The large number of zip codes covered by our transaction data makes it feasible to include demographic and other location-specific variables in an econometric model of means of payment. And the figures above show that the transaction data exhibits a great deal of heterogeneity across zip codes, suggesting the value of including these variables. Table 1 lists the zip-code-level explanatory variables we will use (from 2011), and contrasts the distribution of those variables in our sample of zip codes to their distribution in the United States as a whole. Each variable except the robbery rate is measured at the zip code level (robberies are measured at the county level). For the most part the variable definitions are self-explanatory, and we defer discussion of their role in the empirical model until we present the results below. Here we simply contrast these variables' behavior in our sample and in the United States as a whole.[4]

From the first four rows in Table 1, the zip codes in our sample have lower banks, branches, income and deposits per capita than the United States as a whole. Figure 5 delves deeper into the difference in median income, plotting kernel smoothed density functions for median income in our sample of zip codes and in the United States.[5] 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. Returning to Table 1, population density is somewhat lower than the United States average, but there is also less variation in population density in our sample; the stores tend to be located in zip codes that are neither very sparsely nor very densely populated. Relative to the United States average, these zip codes also have a low percentage of owner-occupied dwellings, with little variation. The racial composition of these zip codes 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. Finally, there is a relatively low percentage of college graduates in our zip codes.

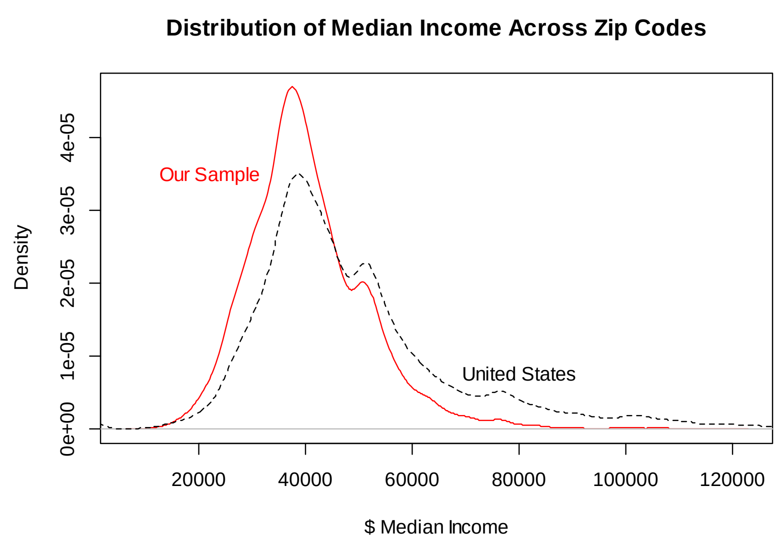

Figure 5.

[4] We discard some zip codes from our transactions data because of missing robbery data.

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

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%
Inventory behavior				
Branches per capita (%)	0.047 (0.214)	0.0045 - 0.1658	0.098 (1.084)	0.0047 - 0.72
Income and price				
Median Household income ($)	40,619 (11,390)	19,370 - 76,850	50,011 (21,475)	20,001 - 128,961
Deposits per capita ($)	2712 (20,158)	35.09 - 15,765	16,153 (1,205,581)	27.85 - 55,296
Banks per capita (%)	0.040 (0.213)	0.0041 - 0.1484	0.091 (0.98)	0.0044 - 0.69
Adoption/usage costs				
Population density (per mile^2)	1437 (2643)	4.2 - 12,021	1782 (5815)	1.8 - 21,159
Robbery rate ($1/10^5$)	13.88 (29.60)	0 - 179.15	14.12 (29.96)	0 - 179.15
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 Empirical Study: A Benchmark

In the preceding section we documented substantial variation in the composition of payments across time, location, and transaction size. We turn now to an empirical model aimed at explaining that variation. Our benchmark estimation presented in this section aggregates all transactions by zip code-day, and includes median payment size in a zip code on a day as an explanatory variable. The analysis is shown to provide a convenient summary of the data and provide much of the intuition. In the next section, we will take a more detailed approach by splitting the data into bins according to size of transaction before aggregating up to the zip-code day level, and run separate regressions for each bin. This allows the explanatory variables to have different direct effects depending on transaction size. As a result, we can better explain the sources and implications of variation in payment composition, both within and across transaction size classes.

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 (i.e. cash, debit card, credit card, and check) used in transactions at stores in one zip code on one day between April 1, 2010, and March 31, 2013.[6] The explanatory variables include the income, banking condition, 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 that we use 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.[7] The fmlogit model is a multivariate generalization of the method proposed by Papke and Wooldridge (1996) for handling univariate fractional response 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 conditions

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

[7] 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 logodds 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.

(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 purpose. 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.

3.2 Estimates for the Overall Payment Mix

The variables used in the regression are rescaled for an easy comparison of the coefficient estimates.[8] Table 2 reports the estimation results. The coefficient estimates are expressed in terms of marginal effects evaluated at the means of the explanatory variables.[9]

3.2.1 Inventory Behavior

Two of the variables we include reflect cash inventory considerations: daily median sale value in a zip code and bank branches per capita. Inventory models of money demand (e.g. Baumol, 1952, Tobin, 1956) suggest that consumers are more likely to use cash when the transaction sizes are small or when the cost of replenishing cash is low.[10] Our results confirm these predictions. As the value of the median sale increases, we find a higher fraction of noncash payments, particularly in debit cards, and to a lesser extent, in credit cards and checks. Evaluating at the mean of median sale value (at zip code level) of $6.86, the marginal effects indicate that a $1 increase in the median sale value reduces cash usage by 1.7 percentage points but raises debit card usage by 1.2 percentage points, credit card by 0.5 percentage points, and checks by 0.05 percentage points. Figure 6 shows that the effect of transaction size on payment mix is also amplified as the median sale value increases. That figure plots predicted payment mix, varying median sale and holding other variables fixed at their means. The range of median sale value in the figure is zero to $15 because that covers virtually all observations for median sale (see Figure 3.b). While the finding regarding median sale suggests that transaction size affects payment mix, we defer to Section 4 a more detailed analysis of that relationship.

On the other hand, we find that bank branches per capita has a positive effect on cash use. In principle, the number of banks per capita may determine competition in a local banking market, thereby determining the price of banking and payment services. However, conditioning on the number of banks per capita, more bank branches in a zip code may reduce consumers' costs of replenishing cash. Indeed we find that the fraction of cash usage increases with the number of branches per capita, mainly at the expenses of debit and credit cards: One additional bank branch per thousand residents increases cash usage by 2.4 percentage points but reduces debit card usage by 1.3 percentage points and credit card usage by 1.1 percentage points.

[8] 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.

[9] 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.

[10] While the basic Baumol-Tobin model does not address the choice of means of payment, the basic features of that model are suggestive of the reasoning in the text. Dotsey and Guerron-Quintana (2013) study an inventory-type model in which there is a nontrivial choice of payment type. See also Freeman and Kydland (2000).

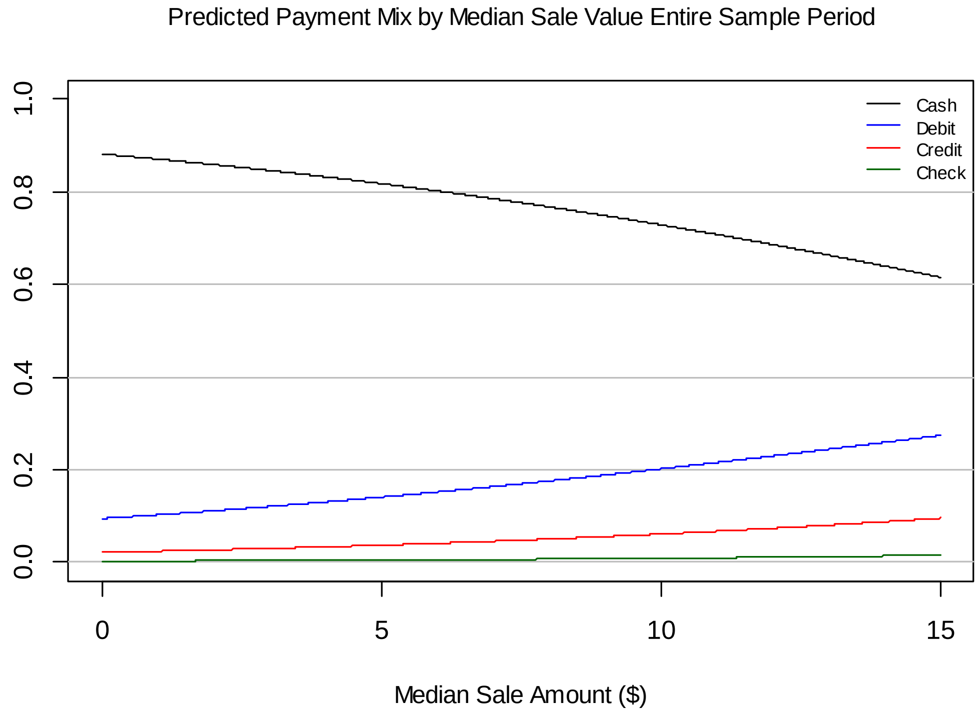

Figure 6.

3.2.2 Income and Price Effects

Income and price are also important determinants of payment choices. Our coefficient estimates show that the fractions of debit and credit card purchases increase with income while the fraction made with cash decreases. The magnitude of these effects suggests that for a $10,000 increase in median household income from its mean, cash use drops by 0.48 percentage points while credit and debit card use rise by 0.42 percentage points and 0.15 percentage points respectively. 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.

While prices associated with different payment instruments are not directly observed, it is possible to investigate the sensitivity of payment choices to factors that may be correlated with prices. In particular, we control for number of banks per capita in the zip code. Presumably, a higher number of banks means more competition and lower banking and payments prices. The findings confirm our hypothesis: more banks per capita reduces cash use, mainly replacing it with credit and debit card use. In terms of magnitude, one additional bank per thousand residents reduces cash usage by 2.3 percentage points, but raises debit card usage by 1.3 percentage points and credit card by 1.1 percentage points. We also investigate the effect of deposits per capita, which is a proxy measure of the banked population. The results are similar to what we found for banks per capita. A $10,000 increase of deposits per capita reduces cash usage by 3.6 percentage points, but it raises the fraction of debit card usage by 3.5 percentage points and credit card by 1.6 percentage points.

3.2.3 Adoption and Usage Costs

Adoption and usage costs also are important factors affecting consumer payment choices. We find that higher population density is associated with lower use of paper payments (i.e. cash and checks) and higher use of card payments. This may reflect the scale economies of adopting new payment instruments. 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 new payment instruments tend to be higher in areas with a high population density or more transaction activities. Quantitatively, an increase of 10,000 population per square mile reduces cash usage by 0.39 percentage points and check usage by 1.4 percentage points, but it raises debit card usage by 0.90 percentage points and credit card by 0.97 percentage points.

We also find that the robbery rate, which relates to the security cost of using cash relative to other payment means, significantly reduces consumer cash usage. In an area with a higher robbery risk, people tend to use debit cards more frequently in retail transactions. Our estimates show that a 0.1 percentage point increase of the robbery rate reduces cash usage by 0.46 percentage points but raises debit card usage by 0.63 percentage points.[11]

3.2.4 Demographics

Much 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 in terms of number of consumers. locations, and time.

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, cash usage is down by 0.093 percentage points and check usage is down by 0.008 percentage points, while debit card usage is up by 0.09 percentage points and credit card usage is up by 0.013 percentage points.

Comparing with renters, we find that a high percentage of homeowners is associated with greater use of credit cards and checks, but lower use of cash and debit cards. However, the magnitude is quite small: A one percentage point higher fraction of homeowners is only associated with the change of each payment type in the range of 0.1-0.9 basis points.

[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.

In terms of gender differences, we find that a high female population is associated with high debit card use 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 ratio reduces cash usage by 0.08 percentage points but raises debit card usage 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 noncash payments, or because families with children tend to use more cash or checks. In contrast, the age statistics show that the age profile with respect to cash and checks is nonmonotonic. A higher presence of the age group 55-69 is associated with a significantly higher fraction of cash usage, while a higher presence of people at age 70 and older is associated with a higher fraction of check usage. These finding suggest that the age variables may also be capturing cohort effects.

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

Turning to the education results, a more highly educated population (i.e. high school and above) is associated with a lower fraction of cash usage 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 usage of cash. 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
Inventory behavior				
Median sale value	-0.017* (0.000)	0.012* (0.000)	0.005* (0.000)	0.001* (0.000)
Branches per capita	0.243* (0.004)	-0.133* (0.003)	-0.113* (0.002)	0.003* (0.000)
Income and price				
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)
Banks per capita	-0.234* (0.004)	0.128* (0.003)	0.109* (0.002)	-0.002* (0.000)
Adoption/usage costs				
Population Density	-0.039* (0.001)	0.090* (0.001)	0.097* (0.001)	-0.148* (0.000)
Robbery rate	-0.046* (0.001)	0.063* (0.001)	-0.006* (0.000)	-0.011* (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)
$\geq$ 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 8.

3.2.5 State effects

The estimates for state dummies reveal marked variation in consumer payment choices across states. Figure 7 plots histograms of state dummies for each payment choice. Conditioning on the other variables in the regression, the cross-state variation appears largest in the fraction of debit card use, with a maximum difference of 14.8 percentage points. Credit cards rank the second with a maximum difference of 9.56 percentage points, and cash ranks the third with a maximum difference of 9.52 percentage points. The cross-state variation is smallest in check use with a maximum difference of merely 0.75 percentage points, reflecting the fact that check use only accounts for 2 percent of all transactions (cf. Figure 1).

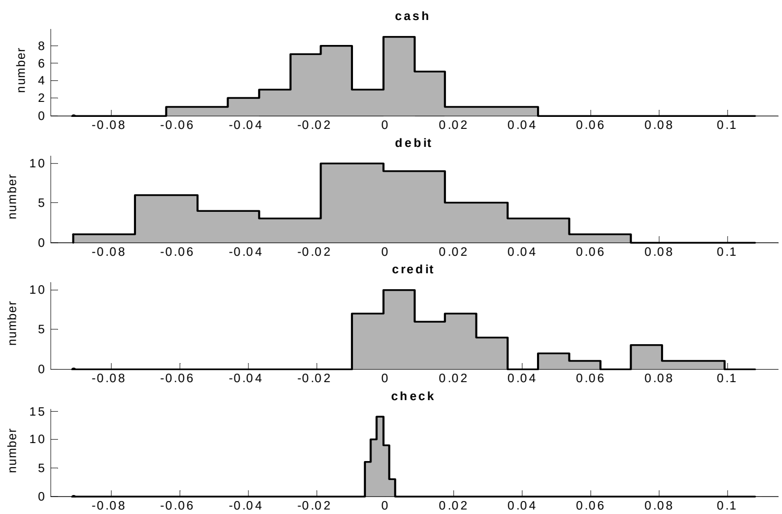

Figure 7. Histograms of state effects.

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 card use. The bottom states for debit card use, such as Minnesota, South Dakota, and North Dakota, appear as the top states for credit card and check use. New Jersey, which ranks the highest in terms of cash use, has the smallest fraction of credit card and check use. These patterns suggest that there may exist systematic variation in payments system usage and regulatory environments at the state level.

3.2.6 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 also 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.

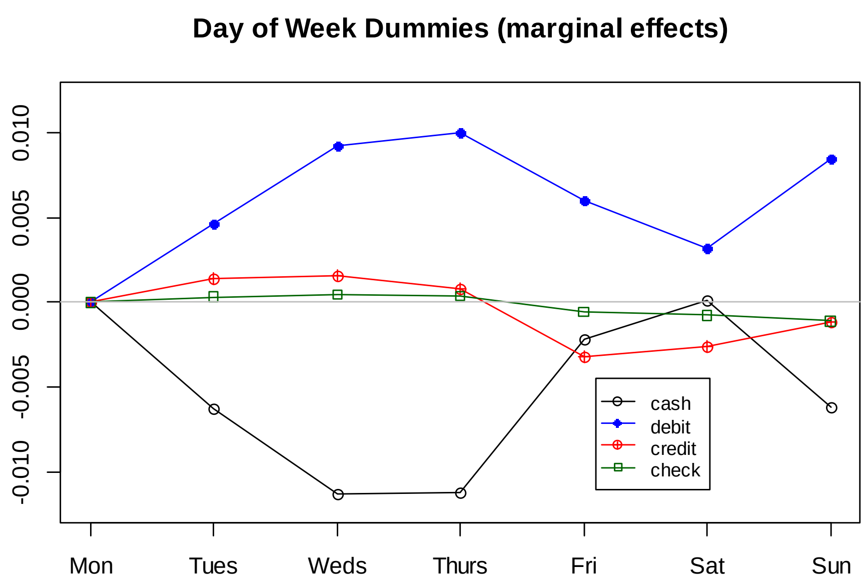

Figure 8.

Figure 8 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 use falls and debit use rises from Monday through Thursday, then cash use rises and debit use falls on Friday and Saturday, and the pattern reverses again on Sunday. Although the credit dummy 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 common patterns of cash replenishment: 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 use 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 9 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. Early in the month, cash use is at its highest and credit and debit use are at their lowest. Over the first 11 days of the month, cash use falls and credit use rises, and then cash and credit are relatively stable for the rest of the month. Debit has a similar pattern to credit, although the variation is smaller, and debit actually peaks before the middle of the month, declines (nonmonotonically) until the 26th, and then rises to the end of the month. Just as the weekly pattern seemed influenced by paychecks, it is also likely that the monthly pattern is 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. 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. Another interesting aspect of Figure 9 is the beginning-of-month volatility in cash, credit and debit. From the 1st through the 3rd, cash use falls and then rises before resuming its relatively smooth decrease. Credit is a mirror image – rising and then falling before resuming its smooth increase. And debit behaves like cash but with a one-day lag: It is flat from the 1st to the 2nd, falls on the 3rd and rises on the 4th, resuming its smooth increase.

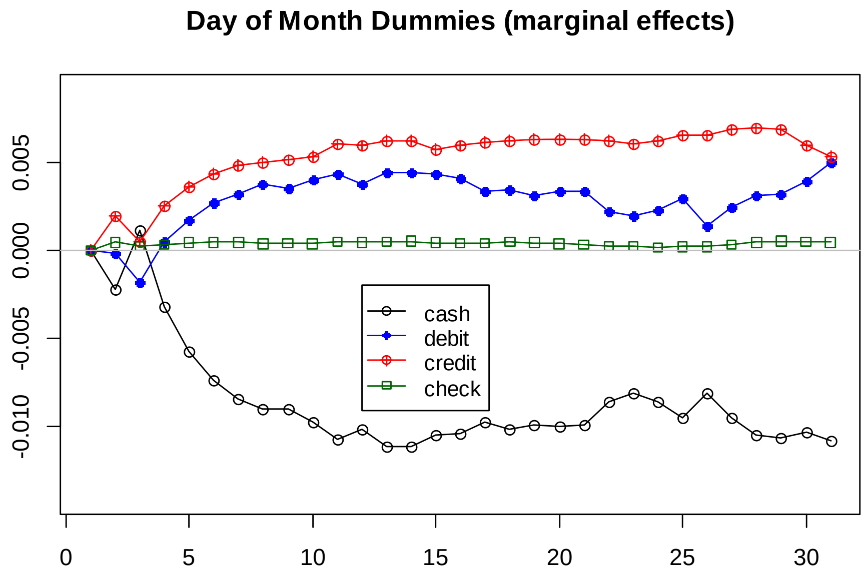

Figure 9.

Figure 10 plots the marginal effects for month-of-sample dummies. As mentioned earlier, these effects combine seasonality with a time trend and idiosyncratic monthly variaton. 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 10 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 11 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, 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) use 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 11 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 use was higher from April 2011 to March 2012 than in the other two years.

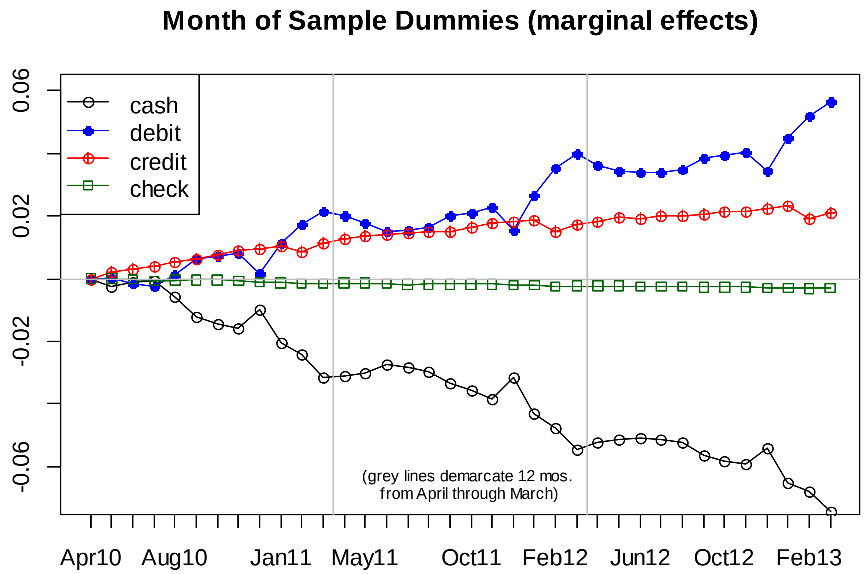

Figure 10.

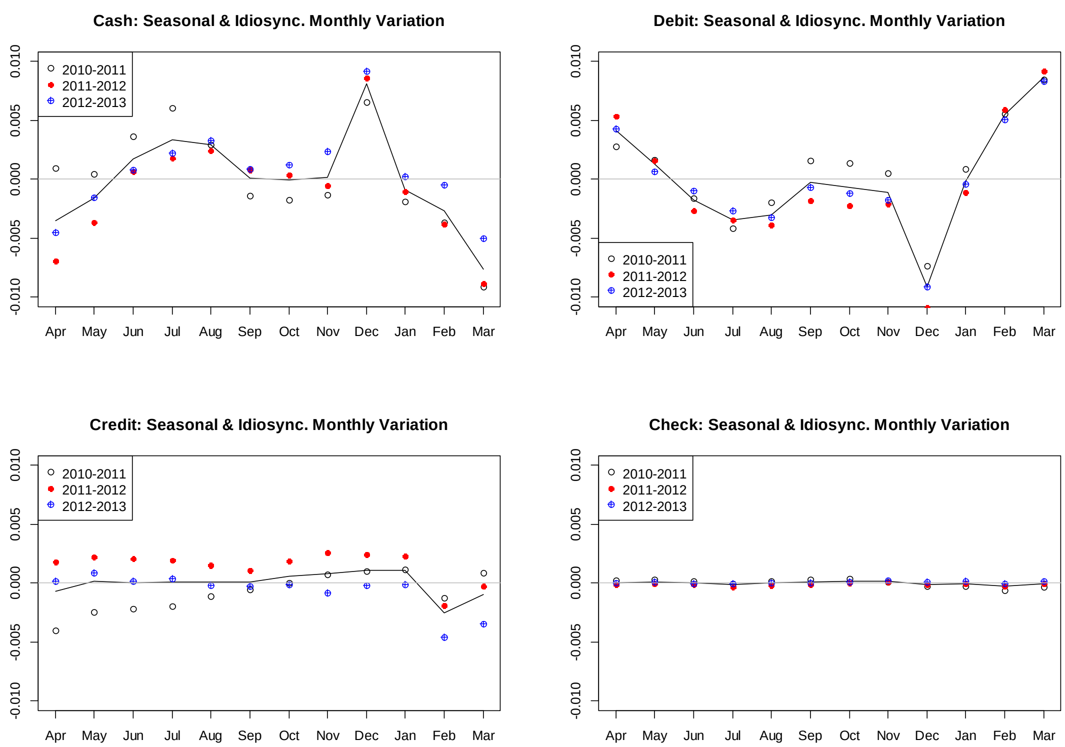

Figure 11.

3.2.7 Further Remarks

Figure 12 displays our model's predicted payment fractions along with the actual payment fractions for the entire sample used in the benchmark regression. The overall fit of our model is fairly good, with pseudo R-squared statistics (calculated as the square of the correlation between the model predicted values and the actual data) ranging around 0.57-0.59.

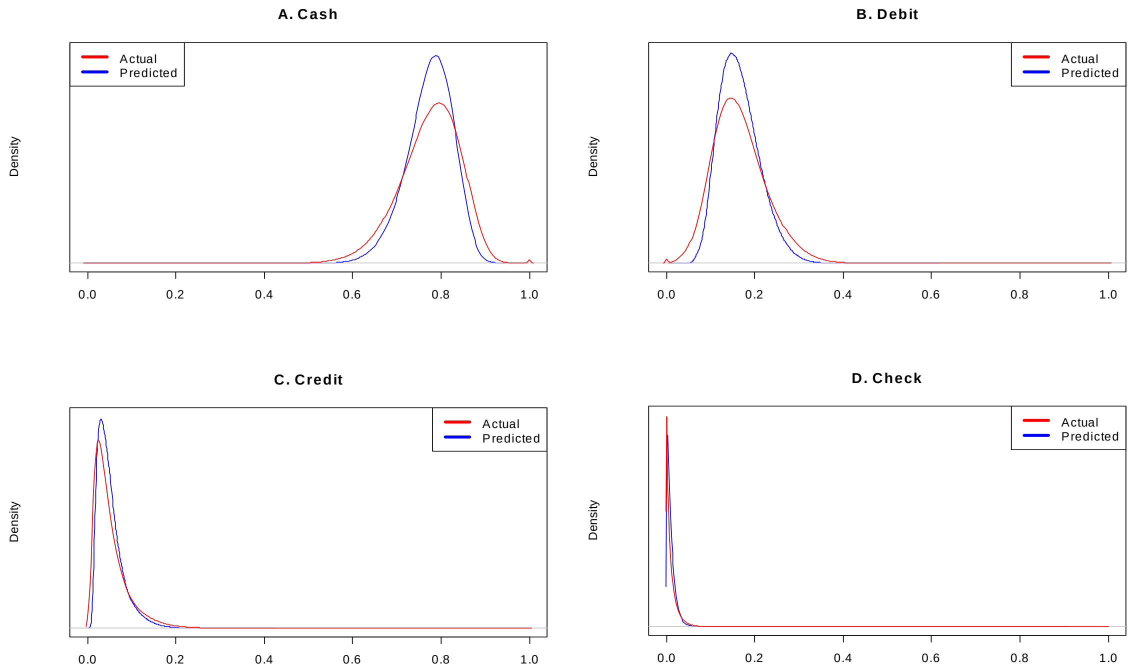

Figure 12. Actual and predicted payment fractions.

Our benchmark model provides a good first exploration of the data. Using the zip-code-level data, the specification allows us to analyze the full sample of observations (based on almost 2 billion transactions) with one regression model, which would be computationally infeasible if we instead used transaction-level data directly. On the other hand, the approach has its limitation: By aggregating transactions of different sizes and including median size as a regressor at the zip-code day level, we do not allow the size of individual transactions to play a role in explaining the payment choice for those transactions. This is in contrast with Klee (2008), who studies consumer payment behavior using transaction-level data and includes individual transaction size as a regressor. In what follows, we address this issue by grouping the data by transaction size and estimating separate models for each group. In so doing, we not only directly incorporate the size of individual transactions into our analysis, but also take a more flexible approach to estimation than Klee (2008) by allowing the coefficient estimates to vary across transactions of different sizes.

4 Estimation by Transaction Size

Two salient patterns in our data, as shown in Figure 4, are that (1) for a given location, the use of cash (non-cash) payments decreases (increases) in transaction size, and (2) the dispersion of payment mix across locations increases in transaction size. Our benchmark regression lumps together all transactions for each zip-code day and includes median transaction size as a regressor. Thus, it can only address these patterns to the extent that they reflect variation in median transaction size across zip-code days, rather than variation in individual transactions. In this section we explore an alternative approach that allows for a systematic relationship between means of payment and individual transaction size.

4.1 Empirical Specification

The key feature of the alternative approach is that 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 doing so, however, there are two possible regression specifications to consider. One is to stack the different transaction size classes into one regression and restrict the coefficient estimates to be the same across those size classes, apart from allowing for separate constant terms or including transaction size class as an explanatory variable. This is similar to the specification in Klee (2008), where the transaction size enters as an explanatory variable independent from others. Alternatively, we may take a fully flexible specification by running separate fmlogit regressions for each transaction size class. This allows the explanatory variables to have different effects in explaining payment mix for different transaction sizes. As we show below, the former specification imposes strong restrictions that may distort the findings.

Recall the fmlogit model that we introduced in Eqs (6) and (7). If we allow the coefficient estimates to vary by transaction size class, the estimated payment mix for transactions of size v conditional on location specific values of x is determined as

$$\hat{s}_{v,k} = \frac{\exp(x\hat{\beta}_{v,k})}{1+\sum_{m=1}^{M-1}\exp(x\hat{\beta}_{v,m})}, \qquad \text{where } k = debit, credit, check,$$

and

$$\hat{s}_{v,M} = \frac{1}{1+\sum_{m=1}^{M-1}\exp(x\hat{\beta}_{v,m})} \qquad \text{where } M = cash,$$

and where the coefficient estimates are indexed by transaction size v. Based on this specification, we can derive the estimated ratios of non-cash payments to cash payments as follows

$$\frac{\hat{s}_{v,k}}{\hat{s}_{v,cash}} = \exp(x\hat{\beta}_{v,k}) \Longrightarrow \ln\frac{\hat{s}_{v,k}}{\hat{s}_{v,cash}} = x\hat{\beta}_{v,k}.$$

Accordingly, the estimated level and dispersion of payment composition across locations, measured as

$$E(\ln \frac{\hat{s}_{v,k}}{\hat{s}_{v,cash}} \mid x) = E(x)\hat{\beta}_{v,k} \quad \text{and} \quad var(\ln \frac{\hat{s}_{v,k}}{\hat{s}_{v,cash}} \mid x) = var(x\hat{\beta}_{v,k}),$$

may vary with transaction sizes due to the variation of $\hat{\beta}_{v,k}$.

In contrast, if we take the alternative (simpler but more restrictive) specification, it is equivalent to assuming that

$$\ln \frac{\hat{s}_{v,k}}{\hat{s}_{v,cash}} = x\hat{\beta}_k + c\,(v,k)\,.$$

The coefficient estimates on zip-code-level variables take common values $\hat{\beta}_k$ across transaction size classes, and $c\,(v,k)$ represents either separate constant terms for each size class or the product of mean or median transaction size with its coefficient. This specification implies that

$$E(\ln \frac{\hat{s}_{v,k}}{\hat{s}_{v,cash}} \mid x) = E(x)\hat{\beta}_k + c\,(v,k) \quad \text{and} \quad var(\ln \frac{\hat{s}_{v,k}}{\hat{s}_{v,cash}} \mid x) = var(x\hat{\beta}_k);$$

only $c\,(v,k)$, not $E(x)\hat{\beta}_k$, can explain changes in the level of payment composition across transaction size, and the dispersion of payment composition as measured by the log share ratio cannot vary with transaction sizes!

To avoid imposing these restrictions, we take the fully flexible specification and run separate regressions for the different transaction size classes. In order to have roughly comparable numbers of transactions underlying each regression, we split the transactions data by sale value in one dollar increments for transactions below \$15 and in five dollar increments for transactions above \$15. For the sake of space, we report only the results for \$1-\$2, \$5-\$6, \$10-\$11, \$15-\$20, \$25-\$30, \$40-\$45 and above \$50 in Tables 4 and Tables A.1–A.3, but Figures A.1-A.4 plot the complete marginal effects for all zip-code level variables. We highlight the following findings from the estimates by transaction size.

4.2 Marginal Effects and Amplification

First, most non-dummy explanatory variables show a sign consistent with our benchmark estimates, but the marginal effects amplify significantly as transaction size increases. Comparing Table 2 and Tables 4 and A.1-A.3 shows that our benchmark 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 for this retailer). Therefore, the discussion of our marginal-effect estimates for the overall payment mix above also applies here for the appropriate size transactions. Moreover, as transaction size increases, the effect of each explanatory variable on the use of each payment form is increasing in absolute value. For example, comparing cash use between \$1-\$2 transactions and \$40-45 transactions, the effects of Pacific Islander, Native American, college education, median household income, deposit per capita and robbery rate rise by a factor of 5.5 to 9.8. The effects of age 35-54, branches per capita, banks per capita, 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 usage.[12]

Second, the marginal effects of state dummies show a consistent sign across transaction sizes, but the amplification effect is relatively mild compared with non-dummy variables. In Tables 5 and A.4-A.6, 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 the benchmark 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 mild amplification as transaction size increases. Taking cash as an example, Figure 16 shows that the maximum cross-state variation is 4 percentage points for $1-$2 transactions, and stays almost unchanged at 12-14 percentage points between $10-$11 and $40-$45 transactions. Similar patterns are found for debit, credit and check usage.

Third, the effects of time dummies also vary by size of payment, tending to be larger in absolute value the larger the transaction size. Comparing Figure 8 to Figure 14 reveals that while the benchmark cash day-of-week pattern is close to the ones estimated for $5-$6 and $10-$11 payments, it is different than the patterns estimated for very small and very large payment sizes. For the benchmark, which combines transactions of all sizes, cash use peaks on Monday and Saturday and troughs on Wednesday and Thursday. For the smallest transactions, the peak for cash is attained on Sunday as well as Monday and Saturday, and the trough shifts to Friday. For large transactions, the peaks for cash use are on Monday and Friday, and the troughs are on Wednesday and Sunday. The magnitudes of day of week effects are increasing in transaction size: For transactions in the $1 to $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 to $45 range. Finally, transaction size plays a role in how the payment instruments trade off against each other over the week. For small transactions, cash and debit marginal effects are virtually mirror images of each other over the week. In contrast, for the $25 to $30 and $40 to $45 regressions, the cash and debit marginal effects together are much higher on Friday than Monday, with the offset coming from relatively low credit use on Friday (Figure 15). For day-of-month dummies, 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 A.9); 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 to $6 range. Turning last to the month-of-sample dummies, these too exhibit interesting variation across the size-specific regressions (Figures 16 and A.10). 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.

[12] In very few cases, the marginal effects flip signs across transaction sizes. For example, the percentage of owner-occupied or vacant homes and percentage of population over 55 have positive effects on cash use in small-dollar transactions, but negative effects on cash use in higher-value transactions. In contrast, Asian and multiple-race populations have negative effects on cash use on small-dollar transactions but not on higher-value ones.

Table 4. Cash: marginal effects by transaction size

Variable	\$1-\$2	\$5-\$6	\$10-\$11	\$15-\$20	\$25-\$30	\$40-\$45	above \$50
Inventory behavior							
Branches per capita	0.032*	0.151*	0.300*	0.393*	0.491*	0.583*	0.597*
Income and price							
Median HH 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*
Banks per capita	-0.029*	-0.143*	-0.289*	-0.380*	-0.476*	-0.567*	-0.582*
Adoption/usage costs							
Population Density	-0.061*	-0.085*	-0.085*	-0.054*	-0.017*	0.000	0.035*
Robbery rate	-0.011*	-0.044*	-0.076*	-0.094*	-0.105*	-0.108*	-0.114*
Demographics							
Family HH	-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*
$\geq$ 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 8.

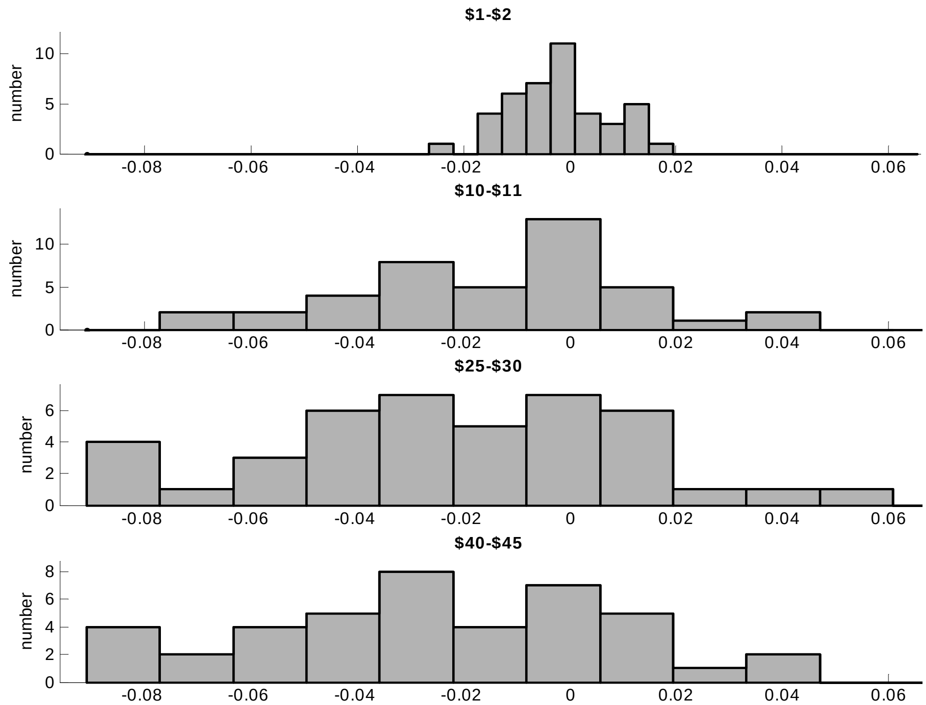

Figure 13. Cash: histogram of state effects.

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

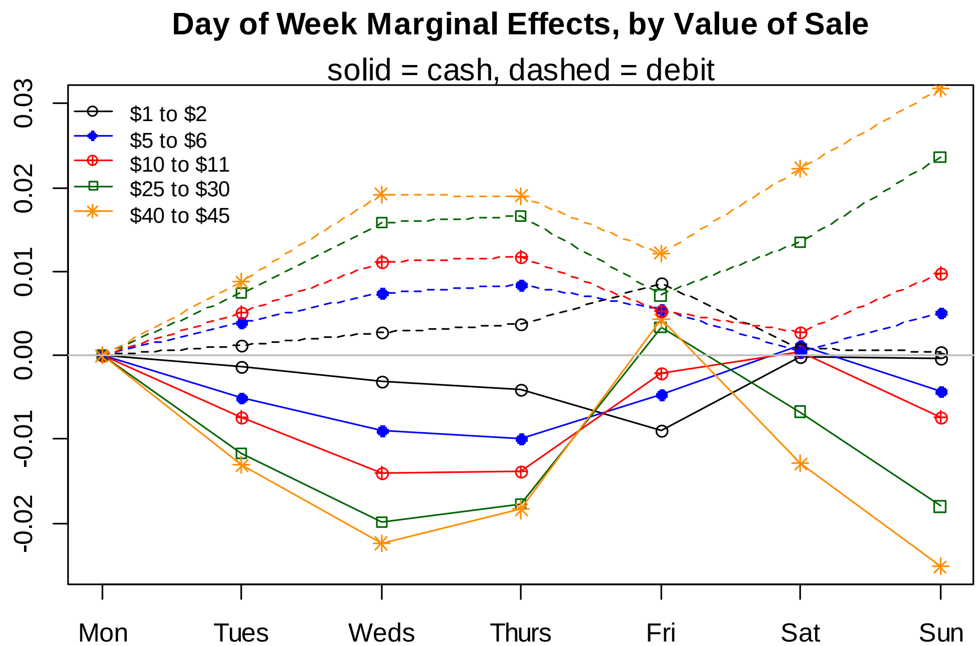

Figure 14.

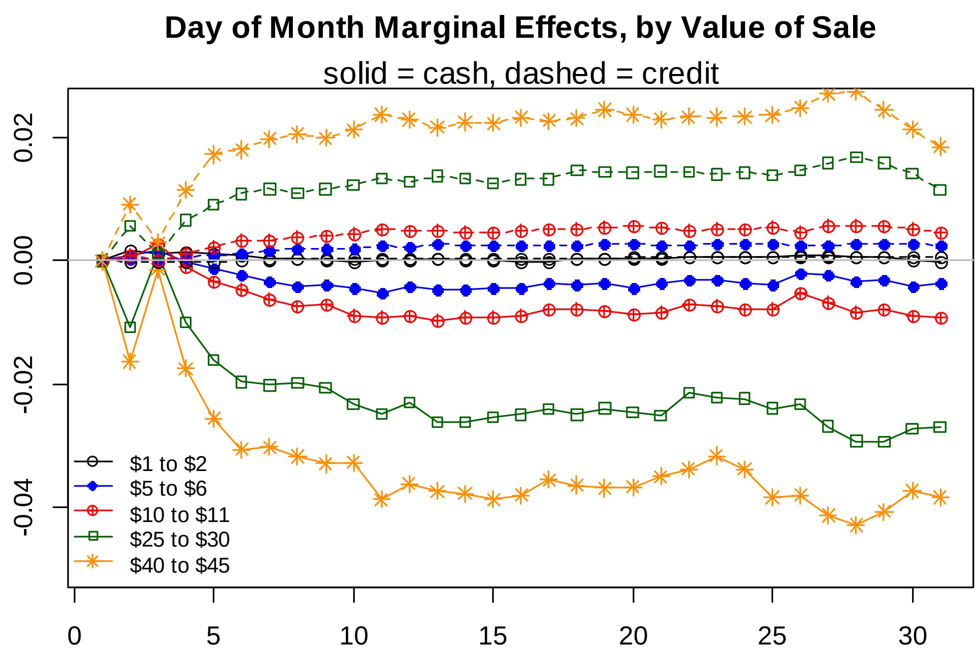

Figure 15.

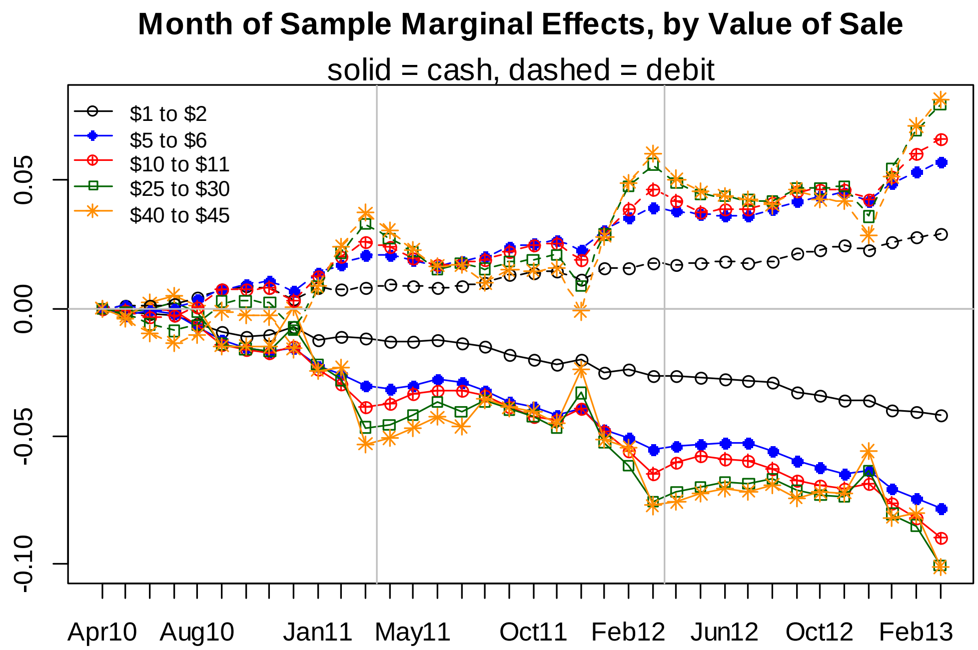

Figure 16.

Overall, the transaction size-class regressions suggest that payment mix varies across size of transactions not only because of a direct effect of size of transactions. Other variables that have independent explanatory power for payment mix exhibit different relationships to payment mix for different payment sizes. For most of the explanatory variables, the marginal effects on payment mix are increasing (in absolute value) in transaction size. This may explain the pattern we saw in Figure 4, in which the level and dispersion of the payment mix varied with transaction size. We now explore this further.

4.3 Payment Variation by Transaction Size

Figure 17 displays the estimated counterpart to the raw data of Figure 4. For each size class, we plot the median, 5th, and 95th percentiles of the distribution of predicted values for each payment fraction. 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 codes. We now discuss how those relationships are related to the amplification of marginal effects.

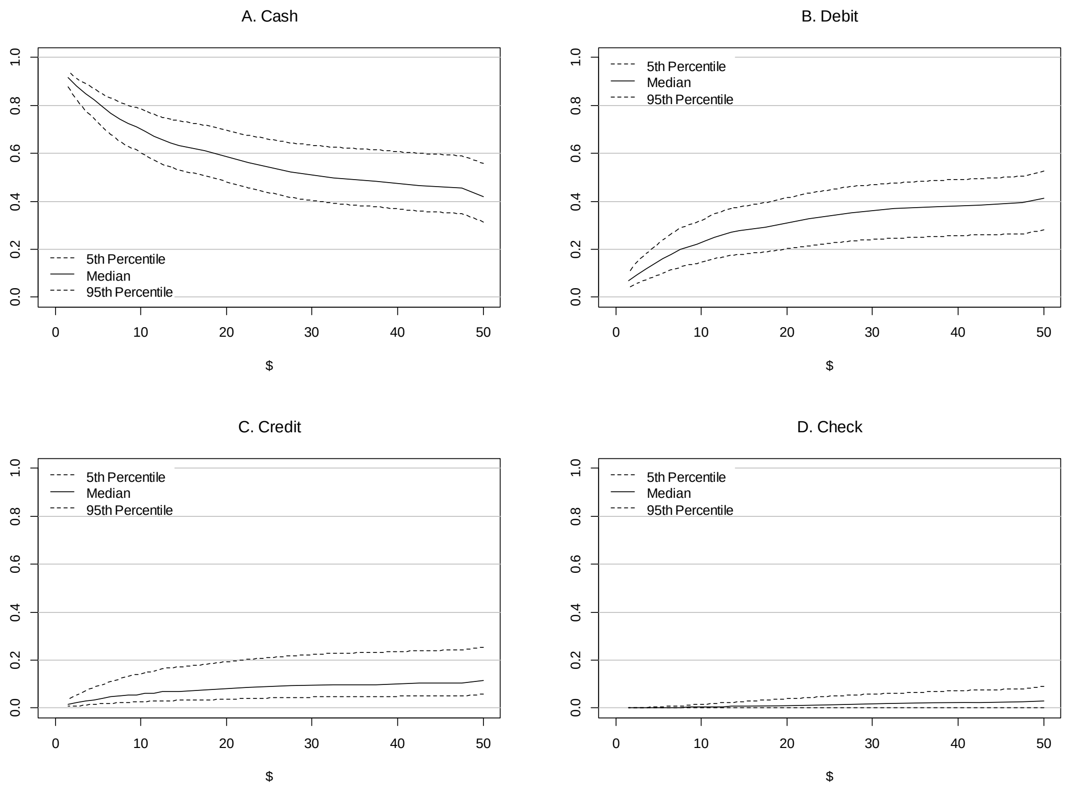

Figure 17. Predicted payment fractions by transaction size.

4.3.1 Level of Payment Composition

For the purposes of this section, we divide the explanatory variables into constant terms, which include the intercept and time and state-level fixed effects, and the rest, which are all zip-code-level variables. At first blush it may seem likely that the change in payment composition across size classes is associated with changes in the constant terms – that is, that a pure "size effect" is the dominant factor in determining the level of the payment mix. In the benchmark case, we did find evidence of a pure size effect, represented somewhat differently in that context by the coefficient on median sale value. This was illustrated in Figure 6. But in the benchmark regression, by construction the *only* way to explain changes in average payment mix across different median sale values was through a pure size effect. In contrast, the regressions for different size classes allow *all* coefficients to vary with the transaction size class. In principle then, the change in average (or median) payment mix across transaction sizes could be associated either with changes in the constant terms or with changes in the coefficients on zip-code-level variables.

In order to decompose the relationship between predicted (mean) payment composition and transaction size into components associated with the constant terms and the coefficients on zip-code-level variables, we alternate holding each set of estimated coefficients constant and varying the other. Specifically, we use the \$1-\$2 regression as our benchmark and hold all right-hand-side variables fixed at their means. 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. In each panel, the solid line represents the predicted values from each size-class regression, holding all variables fixed at their means. The lines marked with circles come from the first exercise described above – allowing only the constant terms to vary, and the lines marked with "x"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 of each payment type. In essence, amplification of marginal effects for zip-code-level variables accounts for the change in the average payment mix.

We move now from the mechanics of amplification of marginal effects to the economics. One dimension along which customers vary across zip codes is whether they possess payment cards, either credit or debit. The zip-code-level fraction of cardholders is not in our data, but the zip-code-level variables may be correlated with– and effectively proxying for– card ownership. Card ownership would be a relevant variable for any transaction size, but the inventory-theoretic money demand considerations introduced in section 3.2.1 suggest that card ownership ought to be more important for larger transactions.

Focusing on cash, where the fraction of payments is decreasing as transaction size increases, suppose that instead of the proxy variables we did have zip-code-level data on card ownership. The coefficient on card ownership for cash use would be negative and would be higher in absolute value for higher transaction sizes. Because the zip-code-level data is fixed across transaction sizes, the results in section 4.2 suggest that the amplification effect would lead to the kind of relationship we see in Figure 18: Changes in the coefficient on card ownership alone would lead to lower predicted cash fractions at higher transaction sizes.[13] Of course, we do not have data on card ownership, but this simple story provides useful intuition for our results.

[13] The derivations in Section 4.1 are only suggestive because they involve log share ratios and estimated coefficients, rather than shares and estimated marginal effects.

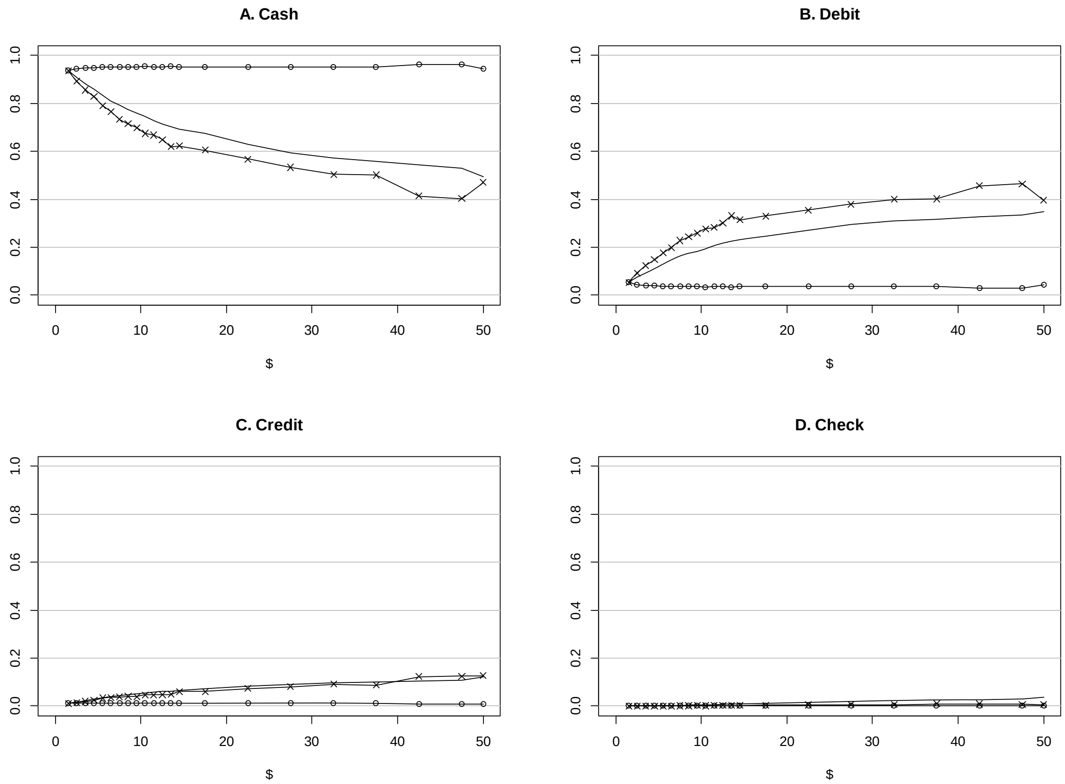

Figure 18. Decomposition of payment fractions by transaction size.
(Predicted (-) Fix zip-code-level coefficients (o) Fix constants(x))

Qualitatively, many of the results from the benchmark regression carry over to the regressions that separate transactions by size. However, the decomposition illustrated in Figure 18 reveals that those size-specific regressions are more than a robustness exercise. The sensitivity of the payment mix to zip-code-level variables changes systematically with transaction size, and is important for explaining the average payment mix for each size class.

4.3.2 Dispersion of Payment Composition

We turn now to the relationship between transaction size and dispersion of payment fractions. The relationship has a relatively straightforward connection to amplification of marginal effects. Both Figures 4 and 17 display a fanning out of the distribution of payment fractions as transaction size increases. Trivially, the fanning out in Figure 4 indicates that for higher transaction sizes, there is more variation in the data. Thus, if our empirical model is reasonably well-specified, we would expect amplification to occur. That is, if payment composition did not become more sensitive to at least some explanatory variables as transaction size increased, our model could not be consistent with the fanning out property. Given that amplification occurs for most explanatory variables, it seems likely that the empirical model is consistent with fanning out.

And Figure 17 confirms that the model is reasonably specified according to this criterion. As transaction size varies then, so does the relationship between payment mix and zip-code-level variables, in a way that is important for explaining both the level and dispersion of the payment mix.

4.4 Time Trends and Long-run Perspectives

In Section 4.2 we reviewed the time dummy effects from the transaction size 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 17 displayed the predicted payment mix for each transaction size, evaluated at the means of the explanatory variables. From the month-of-sample effects (Figures 16 and A.10) we know that the predicted payment mix varies across the sample period. Figure 19 thus compares the predicted payment mix 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.

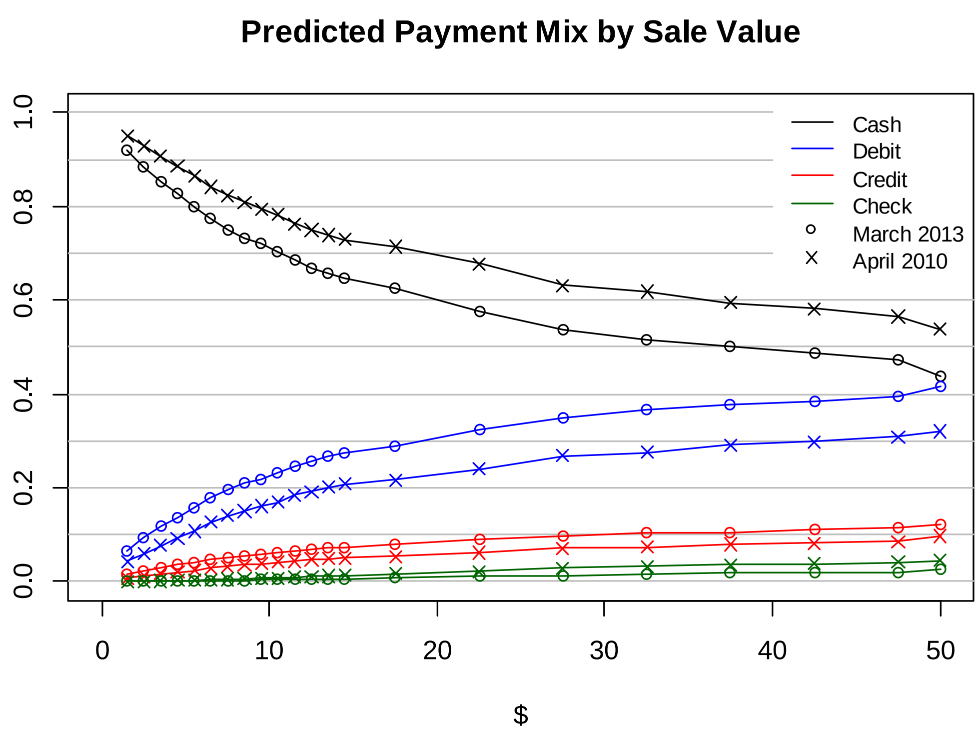

Figure 19.

As with the benchmark regression, we estimated linear time trends for each payment size within each payment type. The resulting linear trends are plotted as annual percentage point growth rates 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 growth of less than 1 percentage point per year for \$1-\$2 transactions to growth of 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. Credit growth 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.

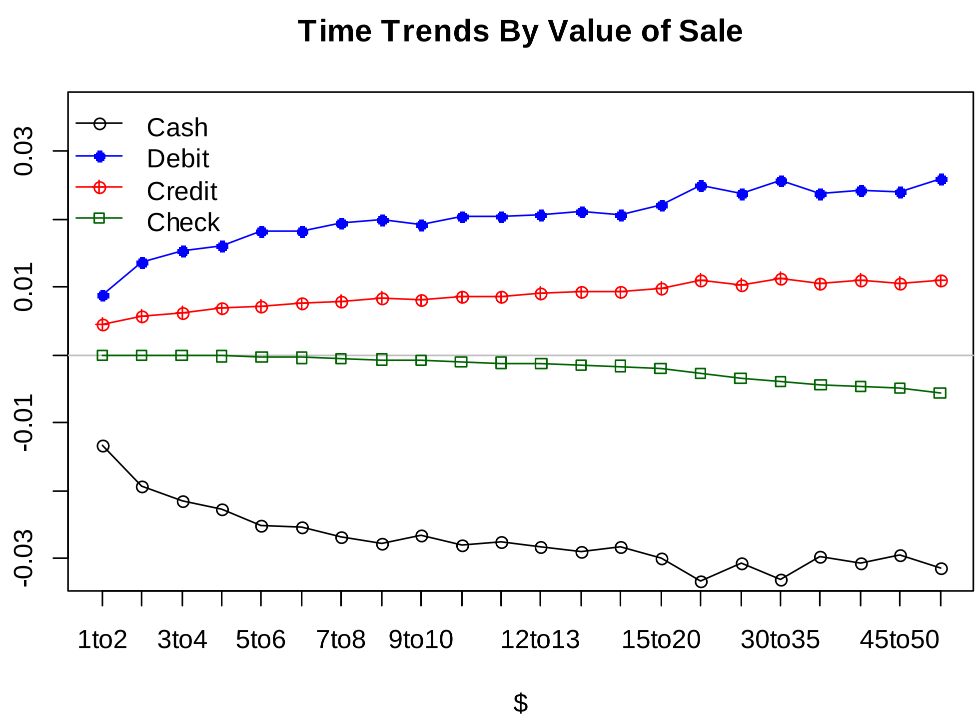

Figure 20.

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, evaluated at the means of the zip-code-level explanatory variables. We then need to forecast each of those variables, as well as the month of sample dummies. 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.[14] We forecast median 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. 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 2013, this gives us benchmark forecasts for the payment mix that do not take into account any time trend. Note that there is a separate forecast associated with each of the payment size regressions.

In Figure 21, the solid red line without symbols represents the predicted cash fractions, by transaction size, for March 2013. The two black lines with open and closed circles represent the forecasts for 2015 and 2020, respectively, based on the demographic forecasts only. For all transaction sizes, the demographic forecasts imply only a slight reduction in the percentage of cash transactions: Less than 1 percentage point from 2013 to 2015, and less than 2 percentage points from 2013 to 2020.

Next we 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 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. For cash, the time trend ranges from a decline of 1.3 percentage points per year for \$1 to \$2 transactions to a decline of 3.3 percentage points per year for \$20 to \$25 transactions. 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 size of setup cost; 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.

[14] 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.

Adding the time-trend effect to the pure demographic forecasts yields our prediction for the retailer's payment mix by transaction size in 2015 and 2020. The red lines with open and closed squares in Figure 21 display the predicted cash fractions for 2015 and 2020, incorporating both the demographic projections and the time trends. For the smallest transactions ($1 to $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 82.4 percent in 2020. For transactions in the $5 to $6 range, cash accounted for 80.0 percent of transactions in March 2013, and we predict that cash will fall to 74.9 percent in 2015 and 61.9 percent in 2020. And for transactions in the $40 to $45 range, the predicted decline in cash is from 48.6 percent of transactions in 2013 to 41.9 percent in 2015 and 26.1 percent in 2020.

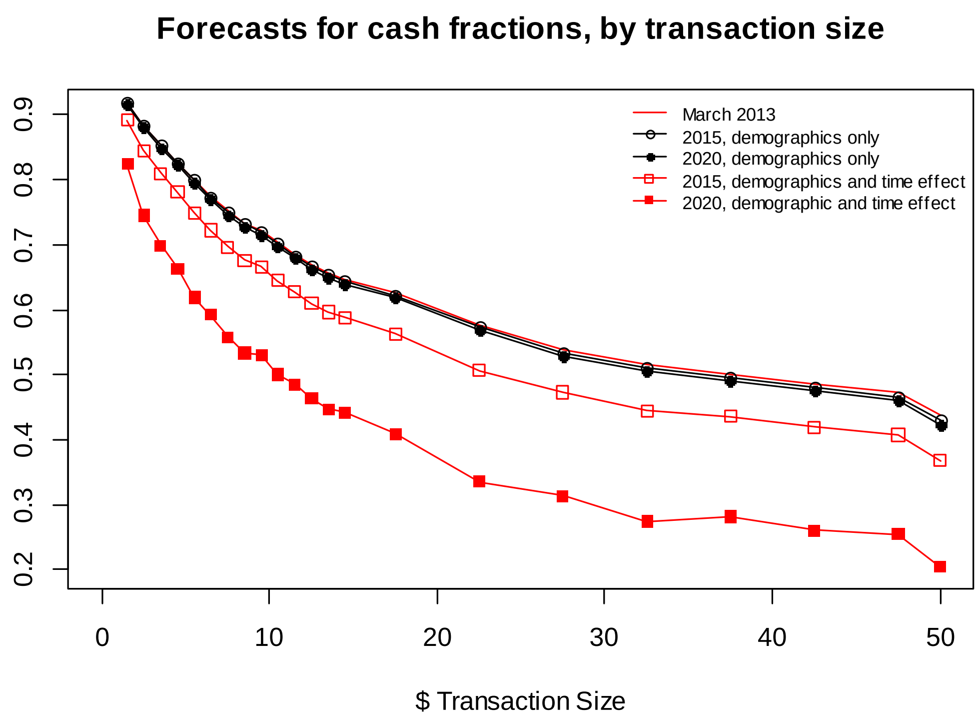

Figure 21.

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 69.9 percent of transactions, and for 2020 cash will account for 57.2 percent of transactions. From 2013 to 2020 then, we forecast that the cash share of transactions will decline by 2.54 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 CDF of payment size shifts to the right exactly one bin, so that, for example, the fraction of transactions less than \$2 in 2020 is identical to the fraction of transactions less than \$1 in 2012. Under this additional assumption, instead of forecasting a 57.2 percent cash share in 2020 we would forecast a 53.5 percent cash share, representing a decrease of 3.1 percentage points per year.

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 United States 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.[15] 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 nominal retail sales, and on the other hand, the combination of a falling in-person share of retail sales and 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:

$$\begin{gathered} C_1/C_0 < 1 \ \Rightarrow \frac{(s-\delta)\,(1+\mu)}{s} < 1 \\ \Rightarrow s < \frac{\delta\,(1+\mu)}{\mu} \\ \text{or, in terms of } \delta, \\ \delta > \frac{s\mu}{1+\mu}. \end{gathered} \tag{9}$$

[15] 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/.

Assuming that $\mu = 0.025$ (the growth rate of in-person retail in 2012), and given our estimated $\delta = 0.0254$, it follows from (9) that the *level* of cash use must be falling regardless of the overall cash share. 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 cutoff 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 cutoff 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. Finally, there is the question of the overall cash share of in-person retail (transactions, not value), as of 2013. We cannot answer that question here, but 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 size of transaction, location, and time. 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 and monthly cycles in payments, as well as time trends and significant state-level variation that is not accounted for by the explanatory variables. We use the estimated model to forecast how the mix of consumer payments will evolve and to forecast future demand for currency. 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. When we combine that time trend with forecasts for the explanatory variables, and with information about the size distribution of payments, we project that the cash share of transactions will decline at 2.54 percentage points per year, from its current level of 75 percent.

Although the retailer we study represents a small fraction of the value of United States retail sales, in absolute terms it has a large number of cash transactions – approximately half a billion per year. As such, our projections for the retailer 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 we plan 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.

Although our data is extremely rich, it does have limitations, and in order to interpret some of our results it would be necessary to complement our data with information on the behavior of individual consumers. In particular, to better understand the patterns in our time dummies, we need information about the behavior of 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 customers paying, as opposed to time variation in the payment choices of a fixed set of customers.

References

[1] 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.

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

[3] 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.

[4] 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.

[5] 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.

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

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

[8] Dotsey, M. and P. A. Guerron-Quintana (2013). "Interest Rates and Prices in an Inventory Model with Money and Credit." Federal Reserve Bank of Philadelphia Working Paper 13-5.

[9] Federal Reserve System (2011). The 2010 Federal Reserve Payments Study. www.frbservices.org/files/communications/pdf/press/2010_payments_study.pdf

[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] 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.

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

[17] 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.

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

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

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

Table A1. Debit: marginal effects by transaction size

Variable	$1-$2	$5-$6	$10-$11	$15-$20	$25-$30	$40-$45	above $50
Inventory behavior							
Branches per capita	-0.007*	-0.072*	-0.178*	-0.231*	-0.292*	-0.345*	-0.350*
Income and price							
Median HH 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*
Banks per capita	0.004	0.067*	0.171*	0.223*	0.282*	0.335*	0.340*
Adoption/usage costs							
Population Density	0.044*	0.058*	0.092*	0.156*	0.270*	0.363*	0.480*
Robbery rate	0.015*	0.051*	0.087*	0.113*	0.142*	0.145*	0.174*
Demographics							
Family HH	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 8.

Table A2. Credit: marginal effects by transaction size

Variable	\$1-\$2	\$5-\$6	\$10-\$11	\$15-\$20	\$25-\$30	\$40-\$45	above \$50
Inventory behavior							
Branches per capita	-0.025*	-0.083*	-0.128*	-0.170*	-0.202*	-0.227*	-0.232*
Income and price							
Median HH 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*
Banks per capita	0.024*	0.079*	0.123*	0.164*	0.194*	0.218*	0.222*
Adoption/usage costs							
Population Density	0.019*	0.069*	0.126*	0.177*	0.239*	0.272*	0.331*
Robbery rate	-0.004*	-0.003*	0.000	0.001	-0.002	0.003	-0.001
Demographics							
Family HH	-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*
$\geq$ 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 8.

Table A3. Check: marginal effects by transaction size

Variable	\$1-\$2	\$5-\$6	\$10-\$11	\$15-\$20	\$25-\$30	\$40-\$45	above \$50
Inventory behavior							
Branches per capita	0.000*	0.003*	0.006*	0.008*	0.003	-0.011	-0.014*
Income and price							
Median HH 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*
Banks per capita	-0.000*	-0.003*	-0.005*	-0.006*	0.000	0.014*	0.020*
Adoption/usage costs							
Population Density	-0.002*	-0.041*	-0.134*	-0.279*	-0.491*	-0.635*	-0.846*
Robbery rate	-0.000*	-0.004*	-0.012*	-0.020*	-0.034*	-0.041*	-0.060*
Demographics							
Family HH	-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*
$\geq$ 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 8.

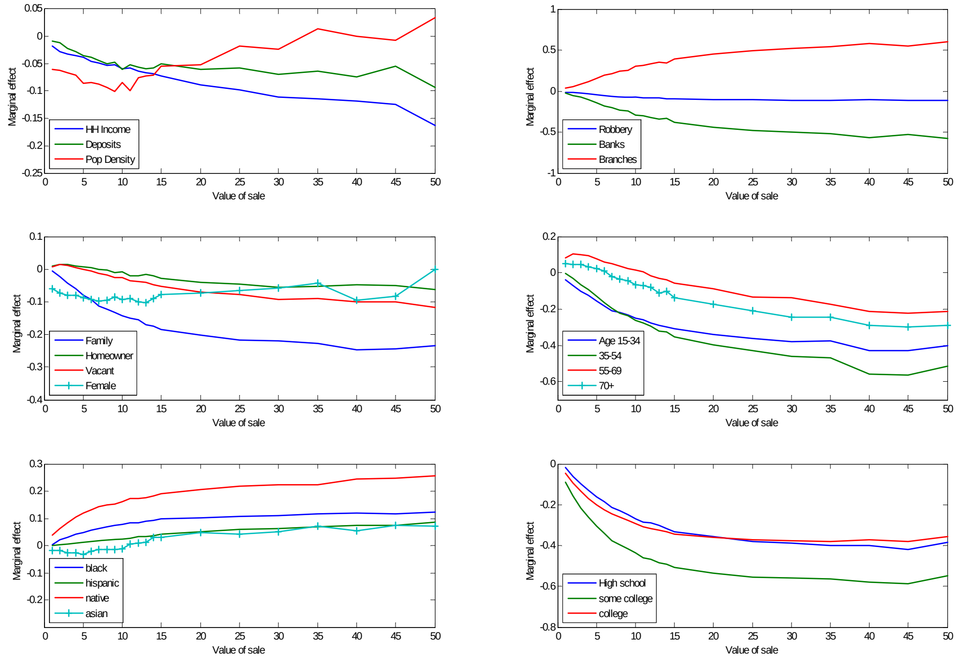

Figure A1. Cash marginal effects by transaction size.

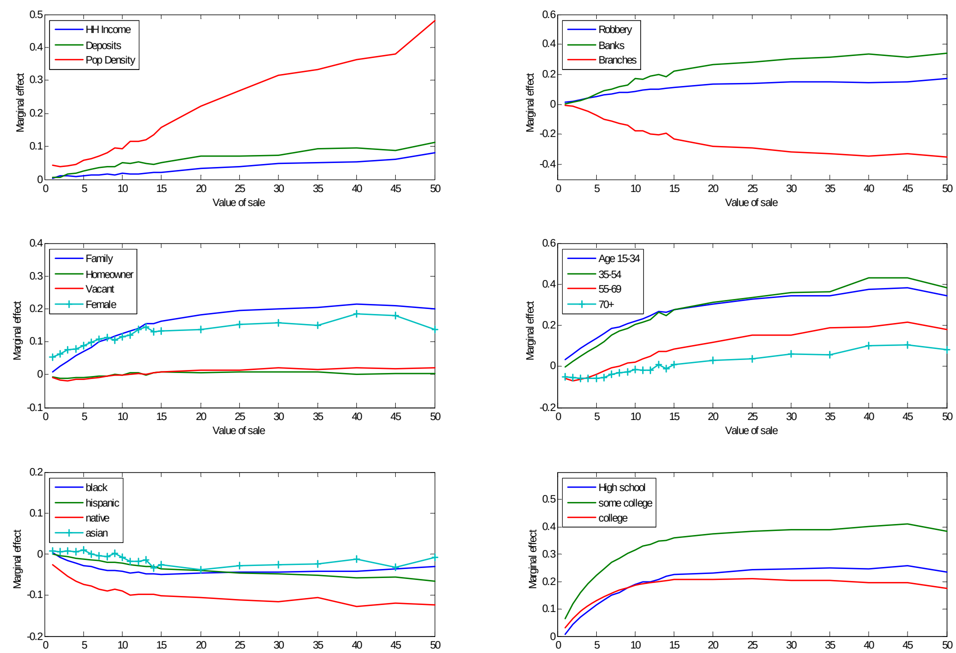

Figure A2. Debit marginal effects by transaction size.

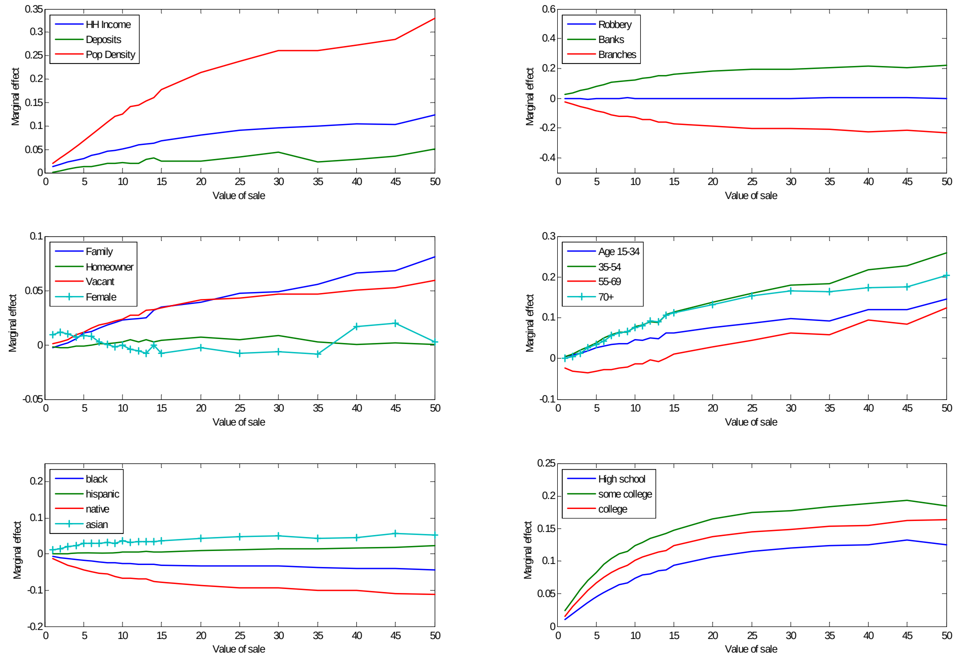

Figure A3. Credit marginal effects by transaction size.

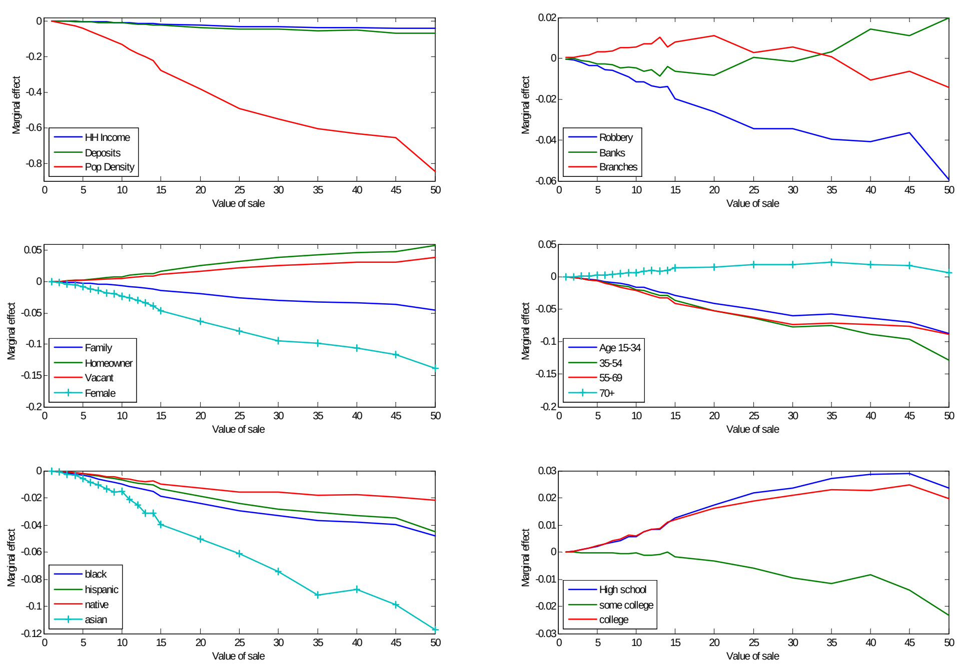

Figure A4. Check marginal effects by transaction size.

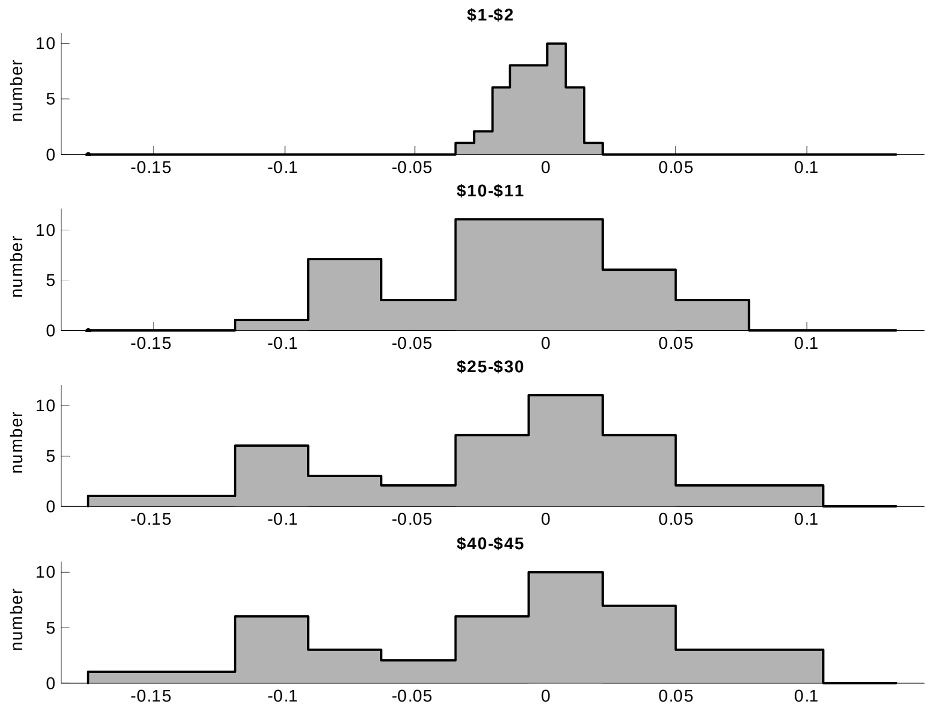

Figure A5. Debit: histograms of state effects by transaction size.

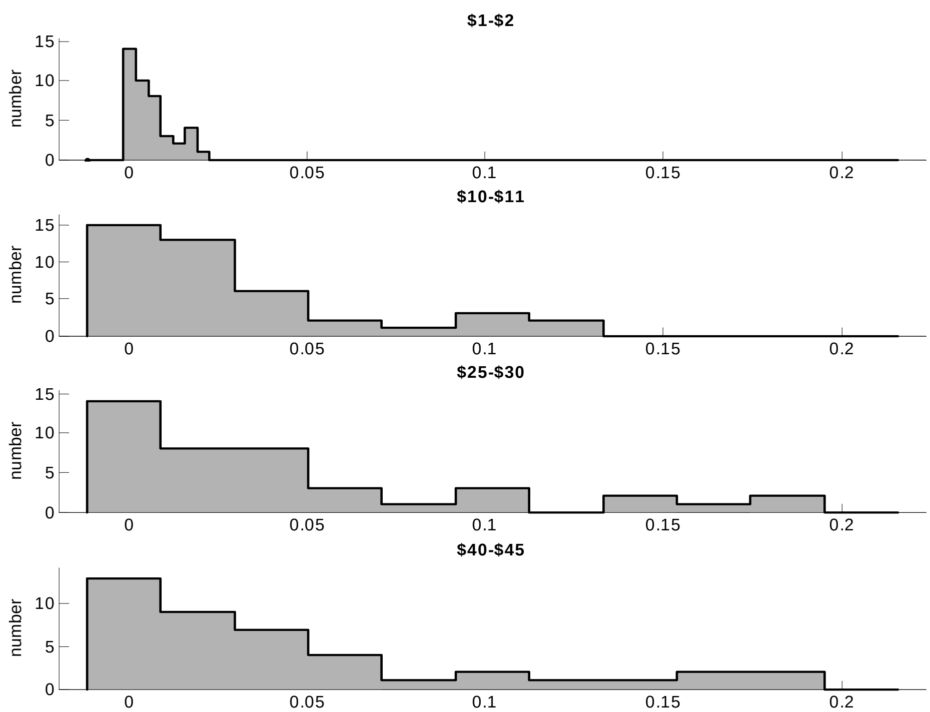

Figure A6. Credit: histograms of state effects by transaction size.

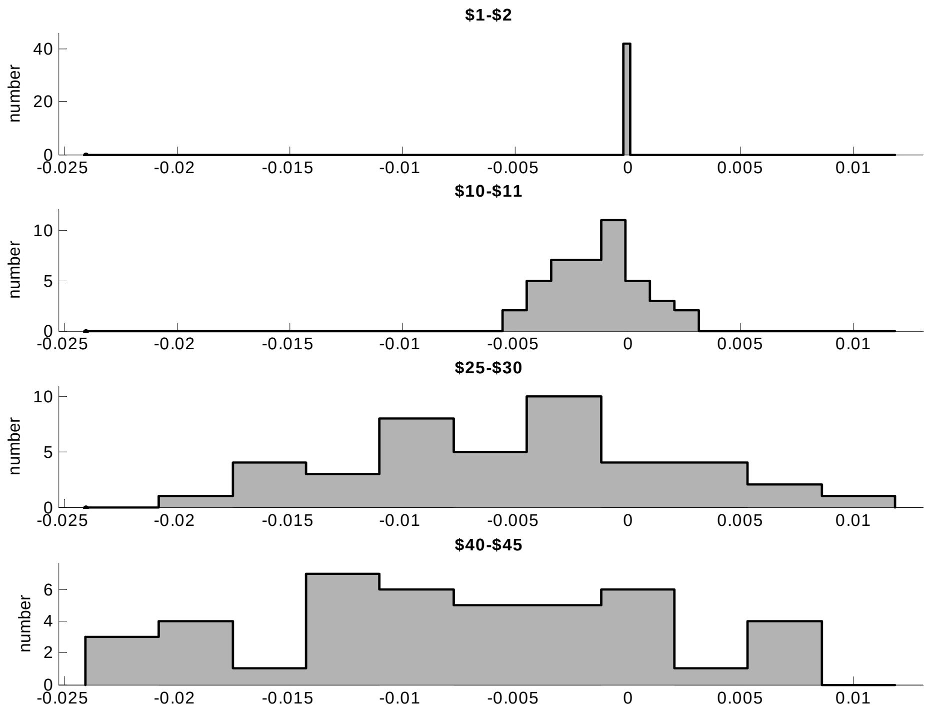

Figure A7. Check: histograms of state effects by transaction size.

Table A4. Debit card: 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 card: 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

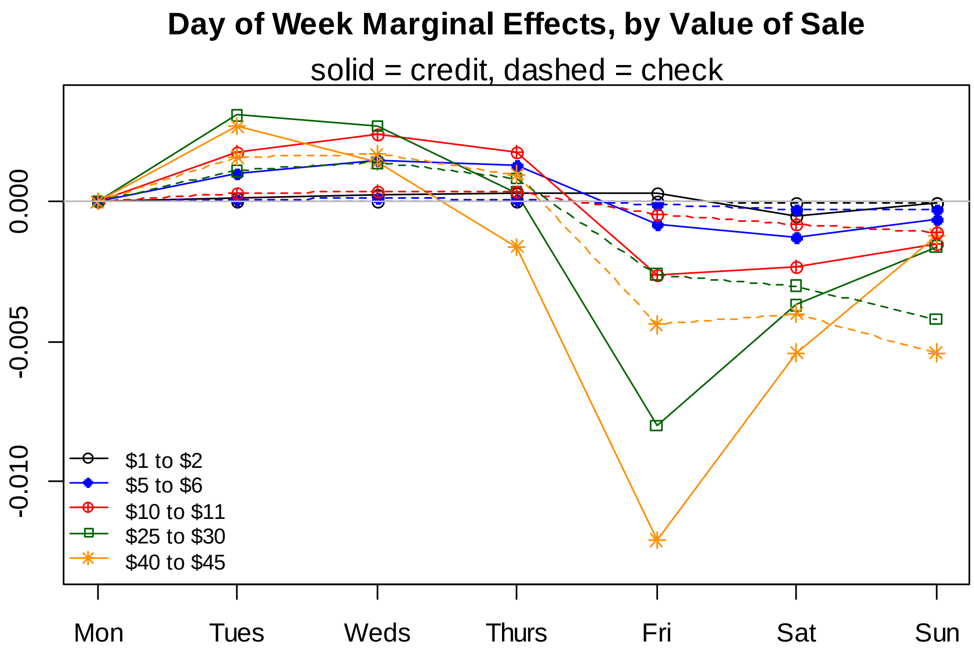

Figure A8.

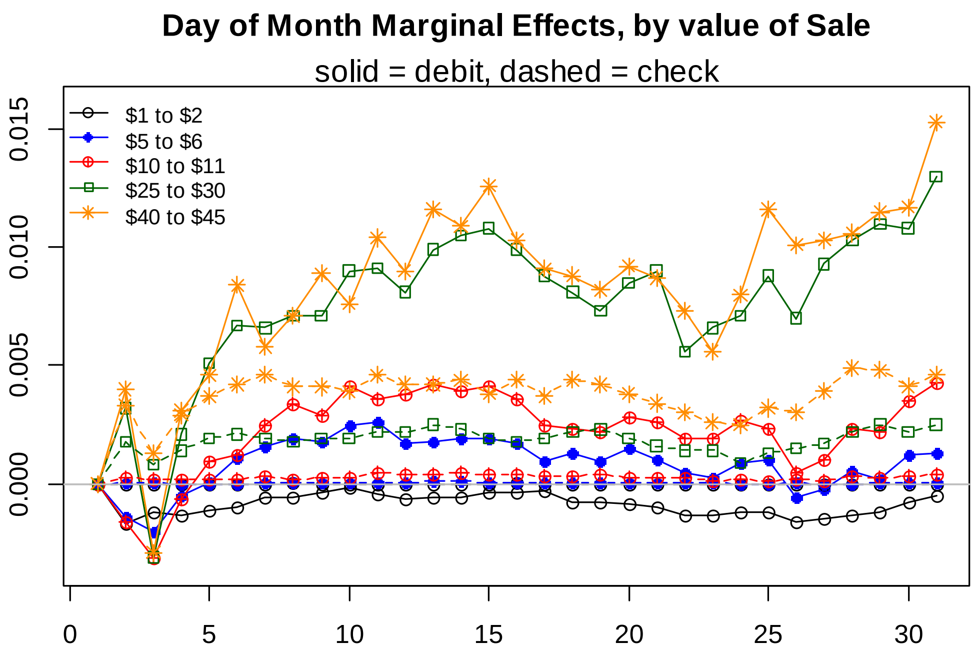

Figure A9.

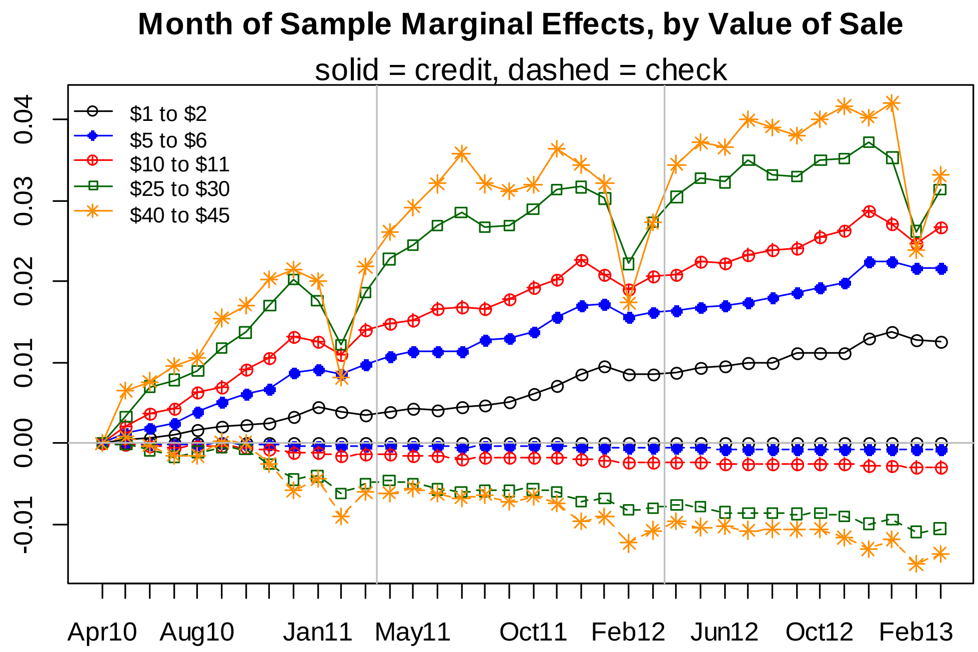
Month of Sample Marginal Effects, by Value of Sale
solid = credit, dashed = check
$1 to $2
$5 to $6
$10 to $11
$25 to $30
$40 to $45
0.04
0.03
0.02
0.01
0.00
-0.01
Apr10
Aug10
Jan11
May11
Oct11
Feb12
Jun12
Oct12
Feb13

Figure A10.