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

Zhu Wang, Alexander L. Wolman*†

March, 2014

Abstract

This paper uses transaction level data from a large discount 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, and to forecast future demand for currency. 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 PSG 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

According to the 2010 Federal Reserve System Payments Study, there were approximately 90 billion transactions undertaken in 2009 with credit cards, debit cards and checks. The Payments System Study and other work have documented in detail developments in non-cash payments over time, but little is known about the behavior of cash transactions. Our study helps to fill that gap. We report new evidence on cash use in retail transactions, as well as credit, debit and check use, based on comprehensive data covering 36 months from a discount retailer with several thousand stores in the United States. 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, 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 provides 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 vs. grocery store, and a more recent time period – 2010-2013 vs 2001. Second, compared with Klee's data, we see richer geographic variation – several thousand zip codes vs. 99 census tracts, and richer time variation – more than 1000 vs. 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 two 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 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, the grocery-store data that Klee uses 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 non-trivial 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).[1] While our data cannot address the underground economy or overseas cash holding, it has the desirable properties for studying cash use that (i)

[1] 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.

the 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 the future of 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 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 are sensitive to assumptions about the current share of cash in overall transactions, and the growth rate of in-person retail sales. However, under a plausible scenario the level of cash use is declining now and will continue 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 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. Although we have data for all stores and all transactions, our zip-code level data – discussed below – only covers a subset 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. The number of payment types is large; in addition to cash, check, credit and debit the retailer accepts electronic benefits (EBT), coupons and some other forms of payment. 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.[2] We do include transactions with cash back, as long as they fit this description. Although the absolute number of transactions that we omit is large because of the scope of the data, we include the overwhelming majority of transactions.

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 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 non-stationarity 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.

[2] As in Klee (2008), the transactions we classify as "credit" may include some signature debit transactions.

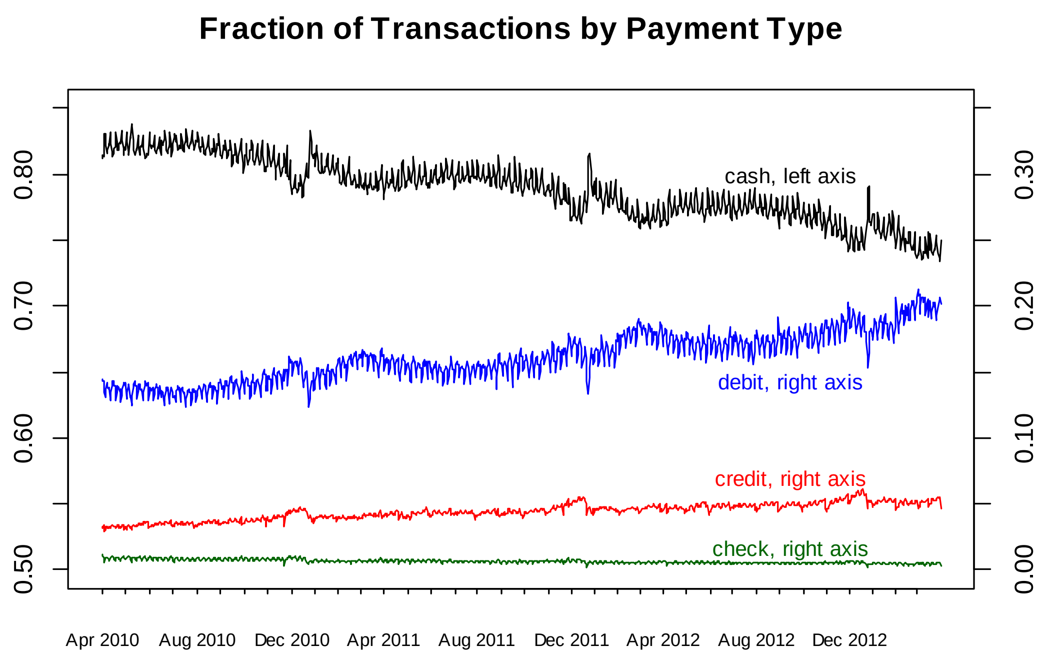

Figure 1.

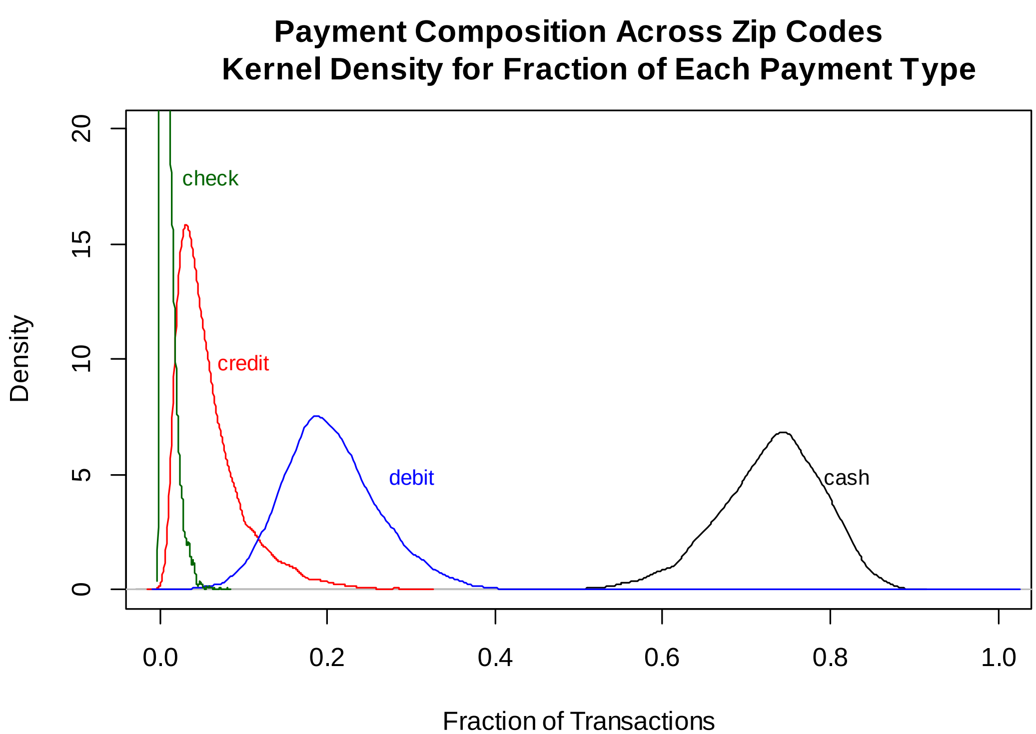

Figure 2.

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.

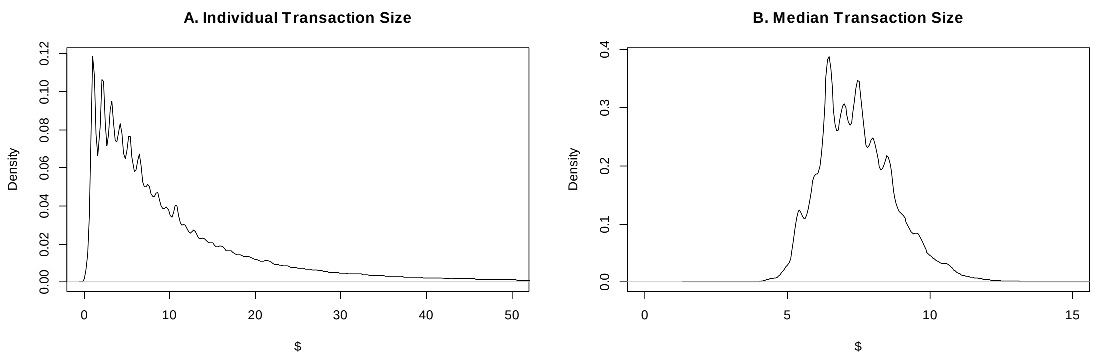

Figure 3. Kernel Densities of Transactions in March 2013

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 $1 size bin, and the dashed lines represent the 5th and 95th percentiles of the distribution. The top left panel shows that for transactions $1 and below, the median zip code had 93% 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 non-cash 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.

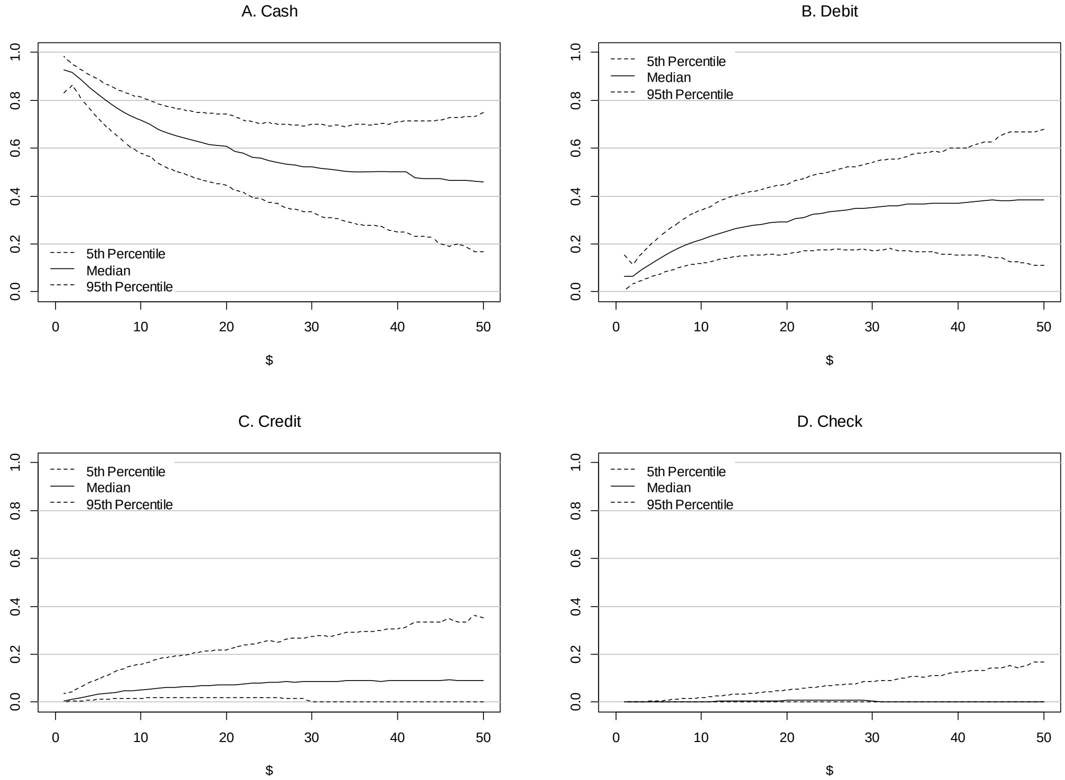

Figure 4. Payment mix by size of payment, across zip codes.

2.2 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 U.S. as a whole.[3]

From the first four rows in Table 1, the zip codes in our sample have fewer banks and branches per capita and lower income and deposits per capita than the U.S. as a whole. Figure 5 delves deeper into the difference in median income, plotting kernel smoothed density functions for median income across zip codes in our sample and in the United States.[4] Although the mode is not much different for the two densities, there is much less mass above the mode in the zip codes where our retail outlets are located. Returning

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

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

to Table 1, population density is somewhat lower than the U.S. 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 U.S. average, these zip codes also have a low percentage of owner-occupied dwellings, and little variation in that percentage. 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.

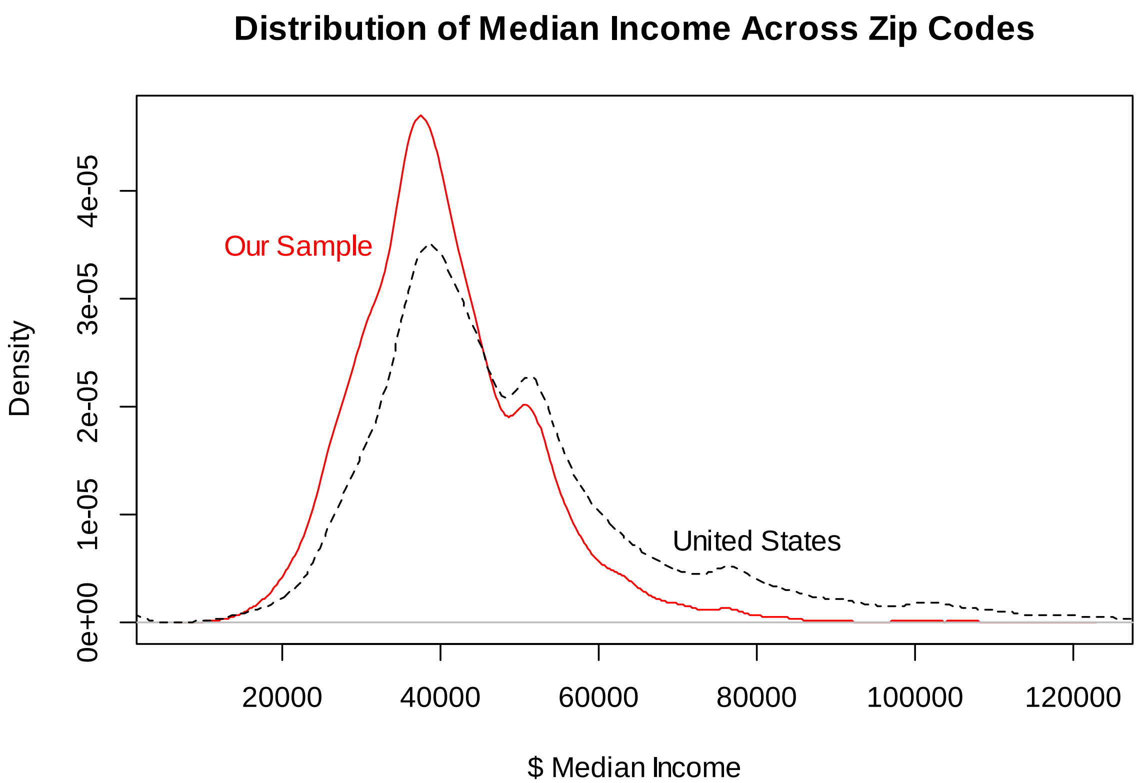

Figure 5.

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

	Our sample		Entire U.S.	
Variable (unit)	Mean (S.D.)	1% - 99%	Mean (S.D.)	1% - 99%
Inventory behavior				
Branches per 100 residents	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 100 residents	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 per 100,000 residents	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
Owner-occupied housing	56.67 (12.62)	19.34 - 80.18	60.29 (15.68)	9.86 - 87.28
Vacant housing	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 15-34	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
$\geq$ 70	10.03 (3.78)	3.25 - 21.42	10.64 (4.36)	2.27 - 23.93
Race 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
Education: 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 analysis 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 source 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.[5] 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.[6] 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 (NBER 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 the

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

[6] 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.

following conditions (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 accommodate the 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

As a benchmark, we estimate the fmlogit model of the payment mix for all transactions together. The variables used in the regression are rescaled for an easy comparison of the coefficient estimates.[7] Table 2 reports the estimation results. The coefficient estimates are expressed in terms of marginal effects evaluated at the means of the explanatory variables.[8]

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.[9] Our results confirm these predictions. As the value of the median sale increases, we find a higher fraction of non-cash 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 about $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 the cash usage by 2.4 percentage points but reduces debit card usage by 1.3 percentage points and credit card by 1.1 percentage points.

[7] 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.

[8] 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.

[9] 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 non-trivial choice of payment type.

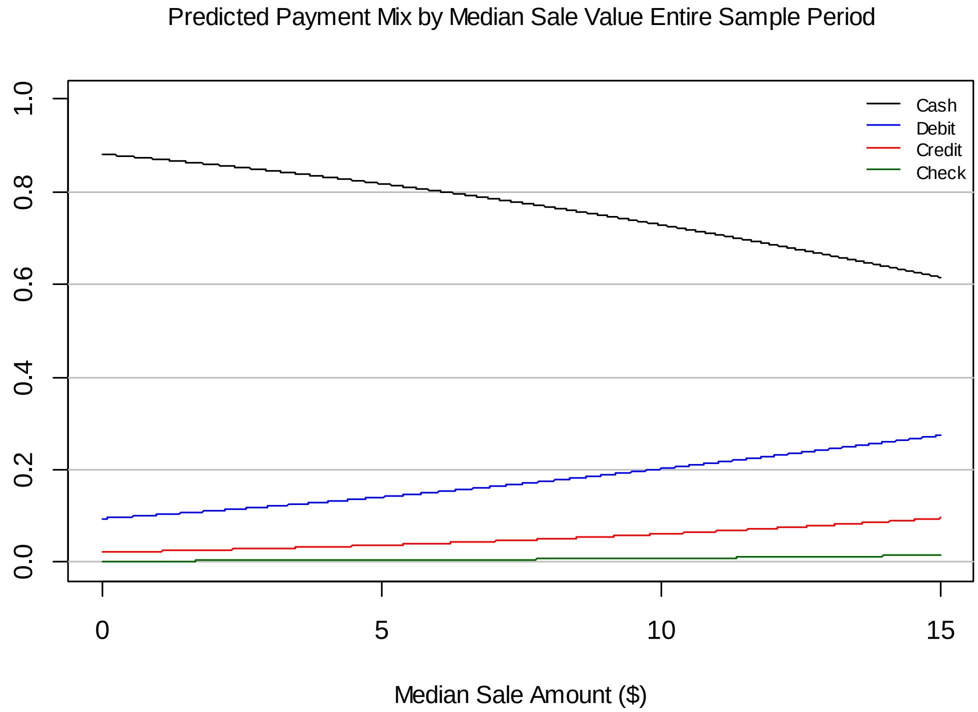

Figure 6.

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. Quantitatively, a $10,000 increase of deposits per capita reduces cash usage by 3.6 percentage points, but 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 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 a 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.

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). However, those studies are typically based on a small sample (hundreds or perhaps thousands of consumers). We evaluate these effects in a sample that has much broader coverage of consumers and locations. In fact, if we assume that our data covers a fixed set of consumers, then the number of consumers can be determined by making an assumption about the frequency of transactions for each consumer. With weekly transactions our data would cover more than ten million consumers, and with daily transactions it would cover more than one million consumers.

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

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%, 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 (e.g. which may relate to our earlier discussion of robbery) or male's preference for anonymity on certain consumption goods (e.g. Klee (2008) argues that certain types of items-beer and cigarettes in particular - are more likely to be purchased with cash than with other forms of payments).

Age statistics also are related to the prevalence of different payment types. A higher presence of older age groups is associated with greater use of payment cards relative to the baseline age group, under 15. This might be simply because minors do not have access to non-cash payments, or because families with children tend to use more cash or checks. In contrast, the age statistics show that the age profile with respect to cash and checks is non-monotonic. A higher presence of the age group 55-69 is associated with a significantly higher 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 (i.e. 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 to the differences from the below-high-school group.

Table 2. Marginal effects for full sample

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 HH 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 HH	-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 7.

3.2.5 State effects

The estimates for state dummies reveal remarkable 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% of all transactions (cf. Figure 1).

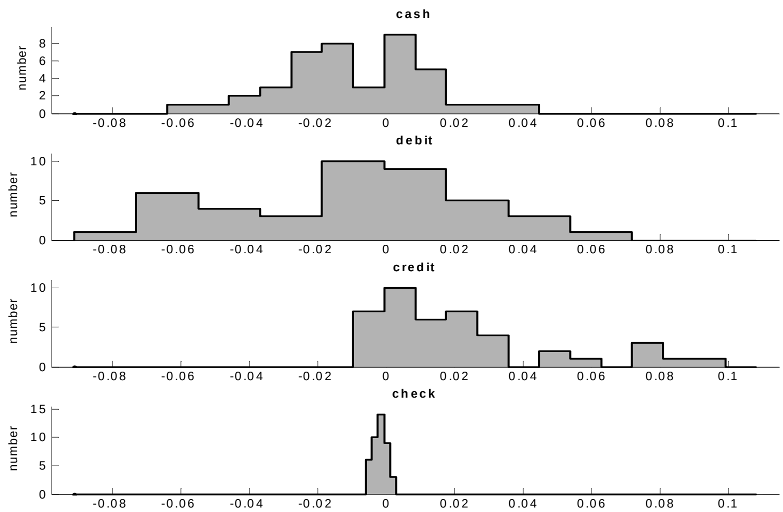

Figure 7. Histograms of State Dummies of Payment Choices

Table 3. 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 to be 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 environment 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 set of customers.

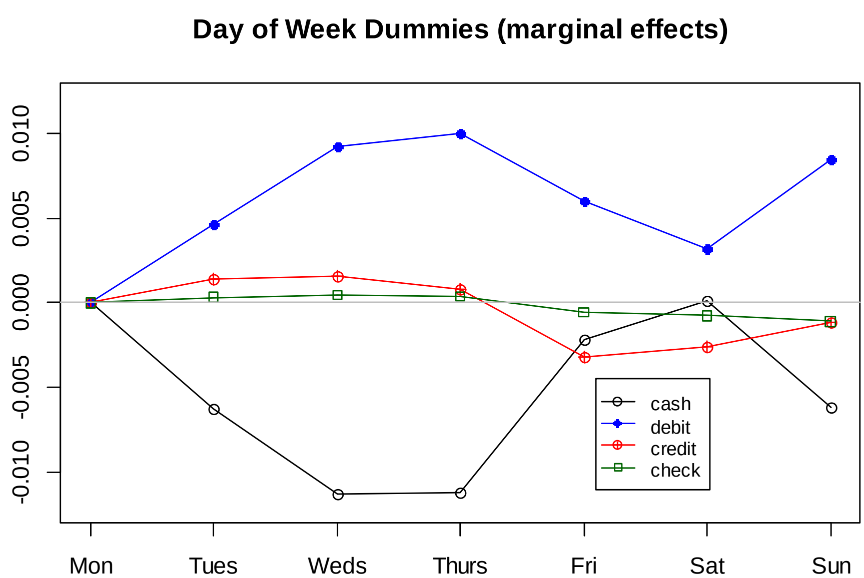

Figure 8.

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 that 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 eleven 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, althought 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. In addition, preliminary evidence on cash-back, discussed below, suggests that ATM withdrawals peak at the beginning of the month, consistent with high cash use early in the month. 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 first to the second, falls on the third and rises on the fourth, resuming its smooth increase. Although we will not provide a complete explanation for the beginning-of-month pattern, the payment-size regressions below will help shed some light on it.

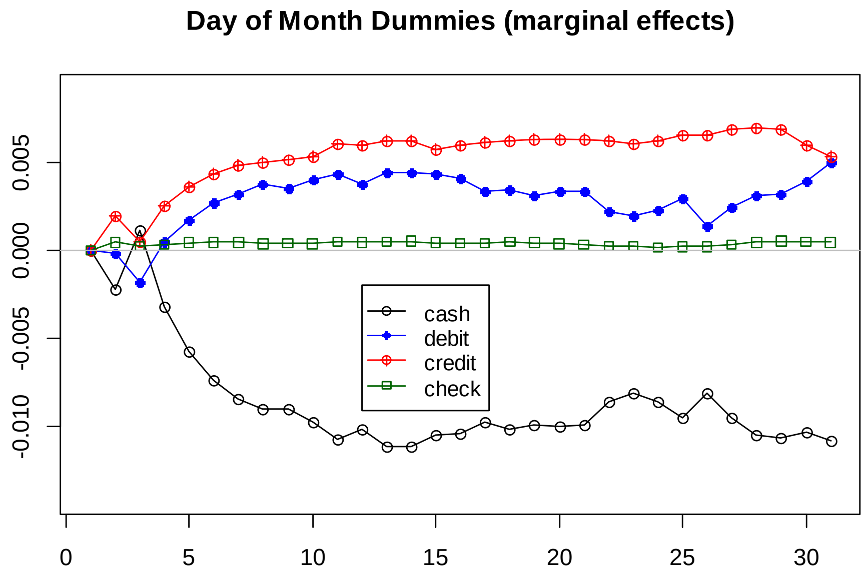

Figure 9.

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 twelve-month periods. Comparing these periods, both the seasonal and trend are striking, but 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 2011-2012, and blue for 2012-2013. 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 one 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.

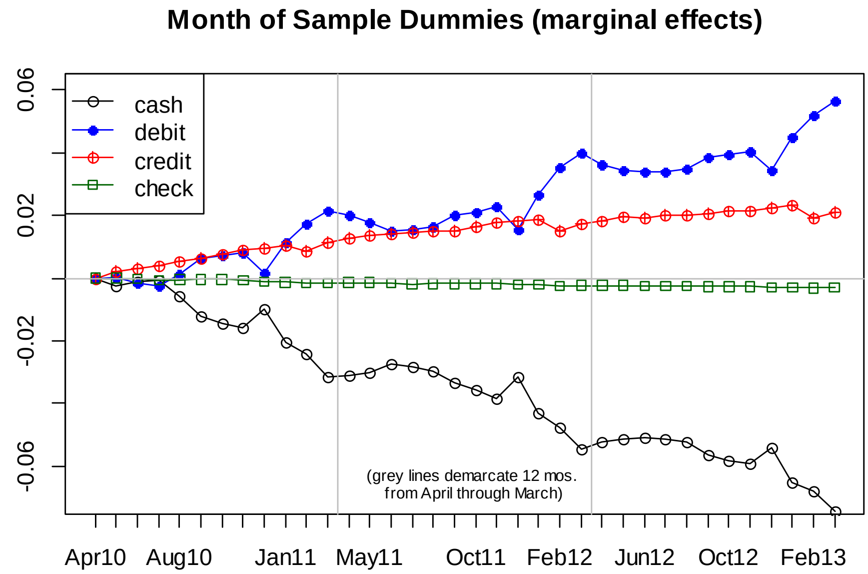

Figure 10.

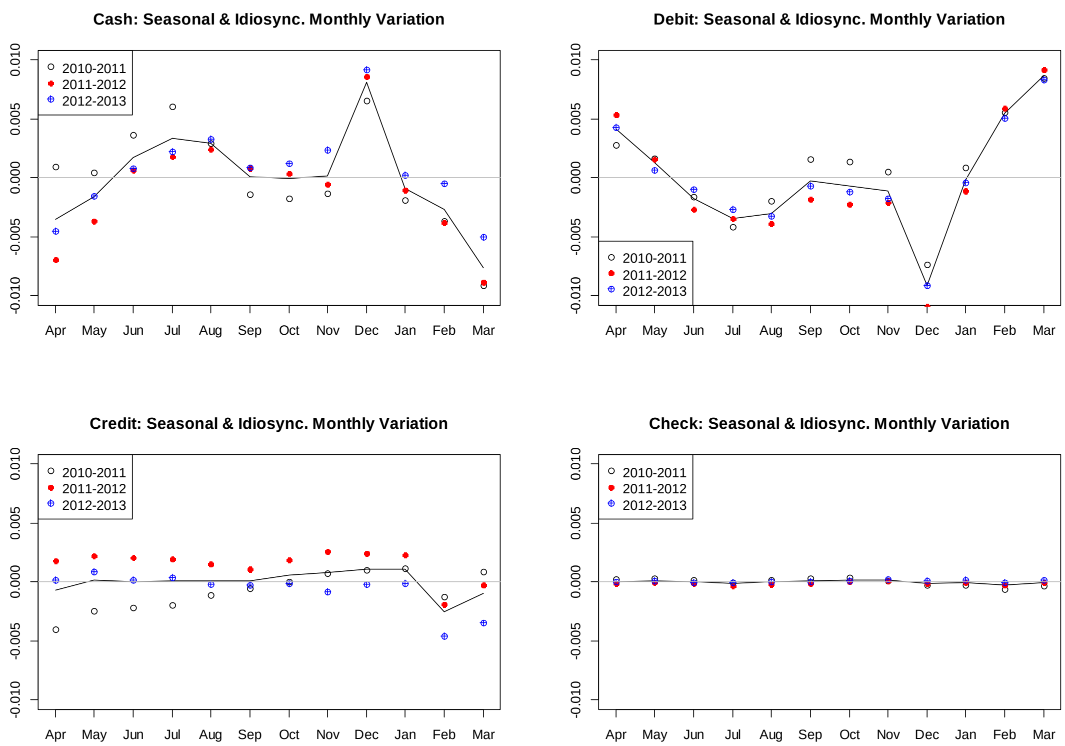

Figure 11

3.2.7 Further Remarks

The overall fit of our model is fairly good, as shown in Figure 12. The pseudo R-squared statistics, calculated as the square of the correlation between the model predicted values and the actual data, range around 0.57-0.59.

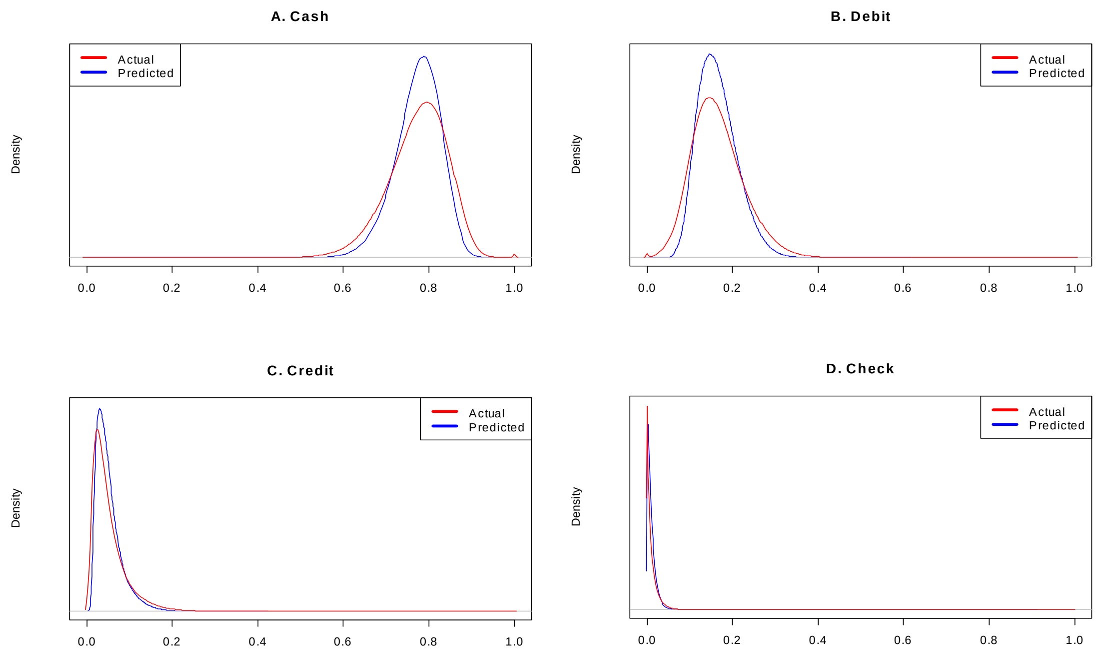

Figure 12. Actual and Predicted Payment Fractions

Our above estimation provides a good first exploration of the data. Using the zip-code-level data, the model specification allows us to analyze the full sample of observations (based on almost two billion transactions) with one regression model, which would be computationally infeasible if we instead estimated a model using 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 size of transaction 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

The estimates presented in this section use a specification that is identical to the benchmark in all but two respects. We omit median sale value from the list of explanatory variables, and we subdivide the sample by sale values before aggregating to the day and zip code level. In contrast to the benchmark regression, this approach directly allows for the explanatory variables to have different effects in explaining payment mix for different transaction sizes. 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 (all other results are available on request). However, Figures A.1-A.4 plot the marginal effects for all zip-code level variables. We highlight the following findings from the estimates by transaction size.

4.1 Marginal Effects and Amplification

First, most non-dummy explanatory variables show a sign consistent with our full-sample 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 full-sample marginal-effect estimates fall between our subsample 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.[10]

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 the marginal effect of using each payment type across transaction sizes. The ranking of states is generally consistent across transaction sizes, 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, Fig 16 shows that the maximum cross-state variation is 4 percentage points for $1-$2 transactions, and stay almost unchanged at 12-14 percentage points between

[10] 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.

$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 is the transaction size. Comparing Figure 8 to Figure 14 reveals that while the full-sample cash day-of-week pattern is close to the one estimated for $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 magnitude 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 one 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 four 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.

Overall, the findings of our subsample 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 potentially explain the pattern we saw in Figure 4: (1) for a given location, the use of cash (non-cash) payments decreases (increases) in transaction size, and (2) the variation in payment mix across locations amplifies as payment size increases. We explore this further in the discussion below.

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*
≥ 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 Table 1.

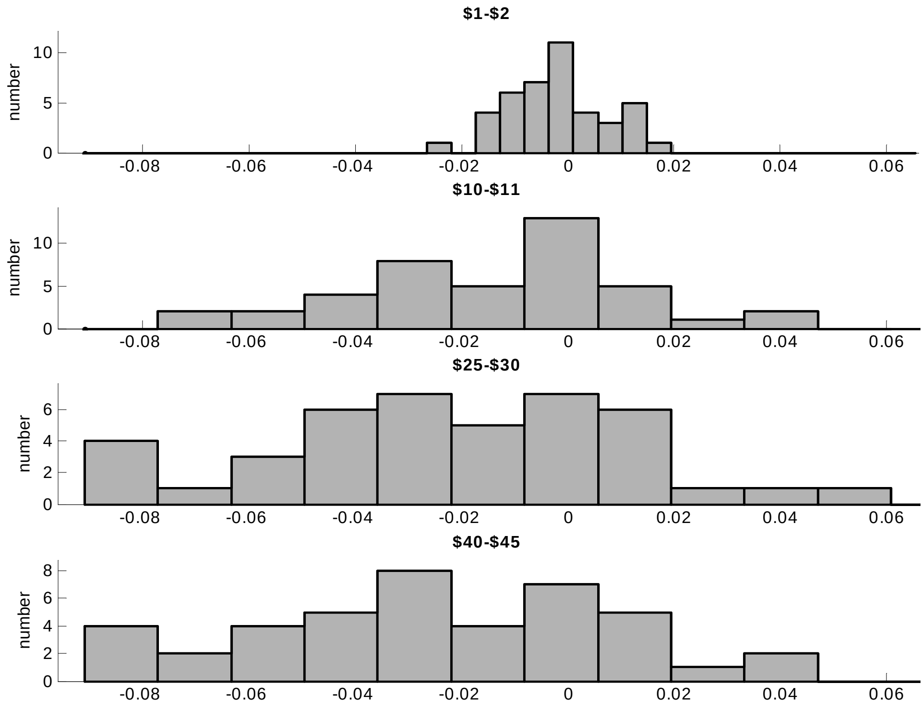

Figure 13

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

Table 5.

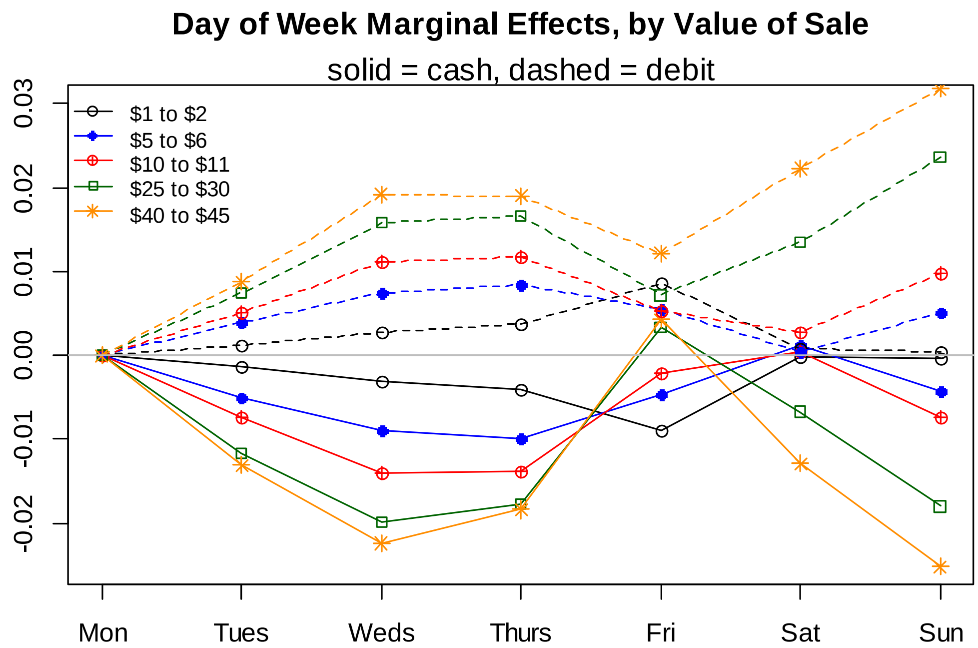

Figure 14.

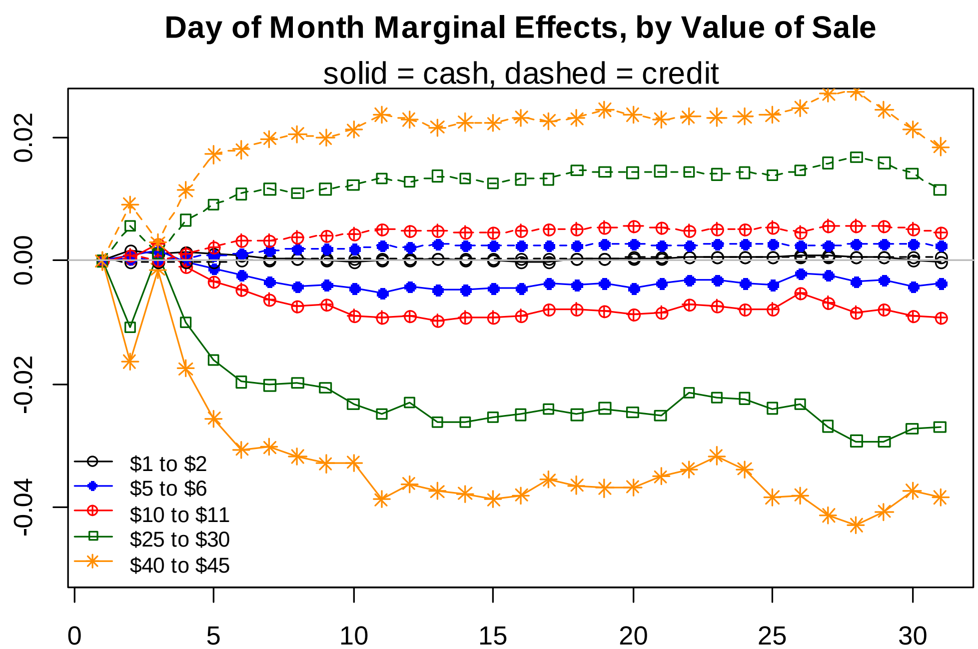

Figure 15.

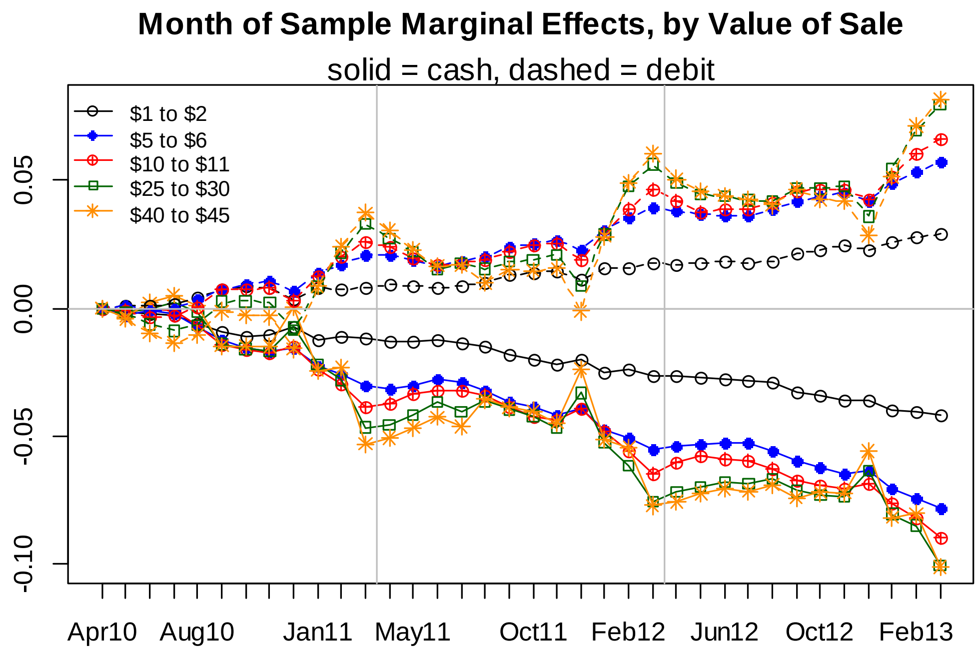

Figure 16.

4.2 Payment Variation by Transaction Size

We have seen that the absolute value of marginal effects generally increases as transaction size increases. We discuss here why amplification occurs, and how amplification is important for understanding the differences in payment composition across transaction sizes. We will relate amplification to the variation in payment composition, both within and across transaction size classes. Within each size class there is significant dispersion in payment composition across dates and locations. As transaction size increases that dispersion tends to increase, but the average levels of each payment type also change.

4.2.1 Dispersion of payment composition

The roots of amplification can be seen in the raw data. Figure 4 displays information about the distribution of payment mix across zip-code days in March 2013. As we mentioned earlier, the panels in Figure 4 display a fanning out of the distribution of payment fractions as transaction size increases. Trivially, the fanning out indicates that for higher transaction sizes, there is more variation in the data to be explained. 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 virtually all explanatory variables, it seems likely that the empirical model is consistent with fanning out. Figure 17 displays an approximate 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. This is an approximate counterpart to Figure 4 because above $15 our regressions are for $5 intervals. Fanning out occurs for the distribution of predicted values for all payment types, although we do not see as much dispersion for very large transaction sizes as appeared in Figure 4. This discrepancy may be due to the fact that the number of underlying transactions is quite small for the largest transaction sizes, leading to greater variance in the sample of payment fractions. In a loose sense, there is more variation in the zip-code level explanatory variables than there is in the day of week, day of month, and state dummy variables. It is thus natural to presume that the zip-code level variables account for the majority of the increase in dispersion for larger transaction sizes – even though the state dummies and time dummies also exhibit amplification. Indeed the amplification associated with the state and time (day of week and month) dummies seems to be small relative to the change in dispersion displayed in Figure 17

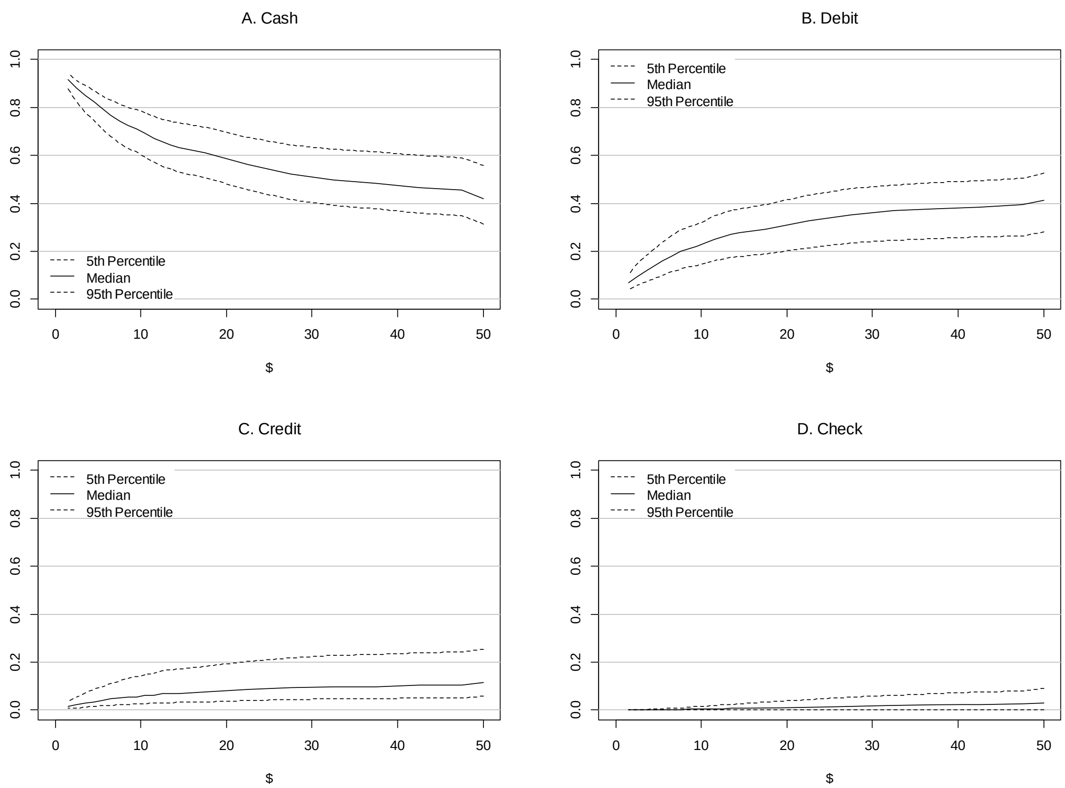

Figure 17. Predicted Payment Fractions

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 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. To the extent that the zip-code level variables are in fact correlated with card-holding, amplification of marginal effects may represent these inventory considerations. That is, at higher transaction sizes, the fraction of cash payments (for example) becomes increasingly sensitive to the fraction of card-holders, and we pick this up with increased sensitivity to our explanatory variables.

4.2.2 Level of payment composition

If amplification of marginal effects has a clear connection to the dispersion of payment fractions within and across size classes, is it also related to the variation in the level (mean or median) of payment fractions across size classes? For the purposes of this discussion, we divide the explanatory variables into time and state-level fixed effects, which we collectively refer to as "constant terms" 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. This pure "size effect" was indeed found in the full sample regression and is illustrated in Figure 6. But in the full sample regression, the *only* way to explain changes in average payment mix across different median sale values was through a pure size effect. We included median sale value as a regressor and held fixed all other coefficients. In contrast, the regressions for different size classes allow *all* coefficients to vary with the transaction size class. In principle then, the change in payment mix across transaction sizes could be associated with either changes in the constant terms or 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 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 constant 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 the constant terms fixed 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, and 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 is 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 not only for the increased dispersion of payment mix within each size class, but also for the change in the average payment mix. Focusing on cash, where the fraction of payments is decreasing as transaction size increases, we can relate this result to the idea that the zip-code level variables in part are proxying for card-ownership. Suppose that instead of these proxy variables, we had zip-code-level data on card ownership. As described above, the coefficient on card ownership for cash use would be negative, and would be higher in absolute value for higher transaction sizes. In a linear probability model, because the zip-code-level *data* is fixed across regressions, the amplification effect would lead to the kind of relationship we see in the top left pane of Figure 18: changes in the coefficient on card ownership alone would lead to lower predicted cash fractions at higher transaction sizes. Of course, we do not have data on card ownership, and the model is nonlinear, but we think this simple story provides useful intuition for our results.

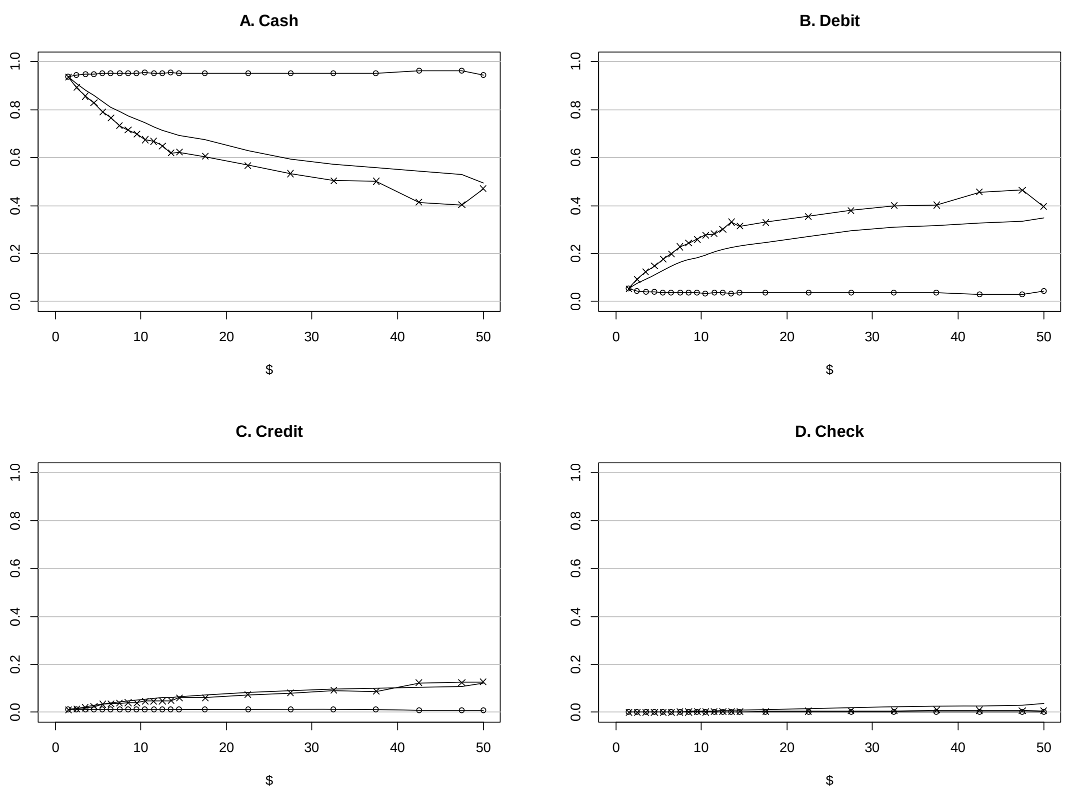

Figure 18. Decomposition of Payment Fractions

(Predicted (-) Fix zip-code level coefficients (o) Fix constants(x))

Qualitatively, many of the results from the full sample regression carry over to the regressions which 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 not only for explaining dispersion within a transaction size class, but also the average payment mix for each size class. These properties are missing in models that impose identical coefficients on zip code level variables across transaction sizes, such as our full sample benchmark model, or Klee (2008). As a result, those models may be less successful in matching the quantitative variation in payments composition across transaction sizes.

4.2.3 Time Trends and Long-run Perspectives

In Section 4.1 we reviewed the time dummy effects from the transaction size regressions. Here we discuss how the predicted payment mix varies from the beginning to end of the sample period, as well as the time trends implied by the estimated time effects.

Figure 18 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.

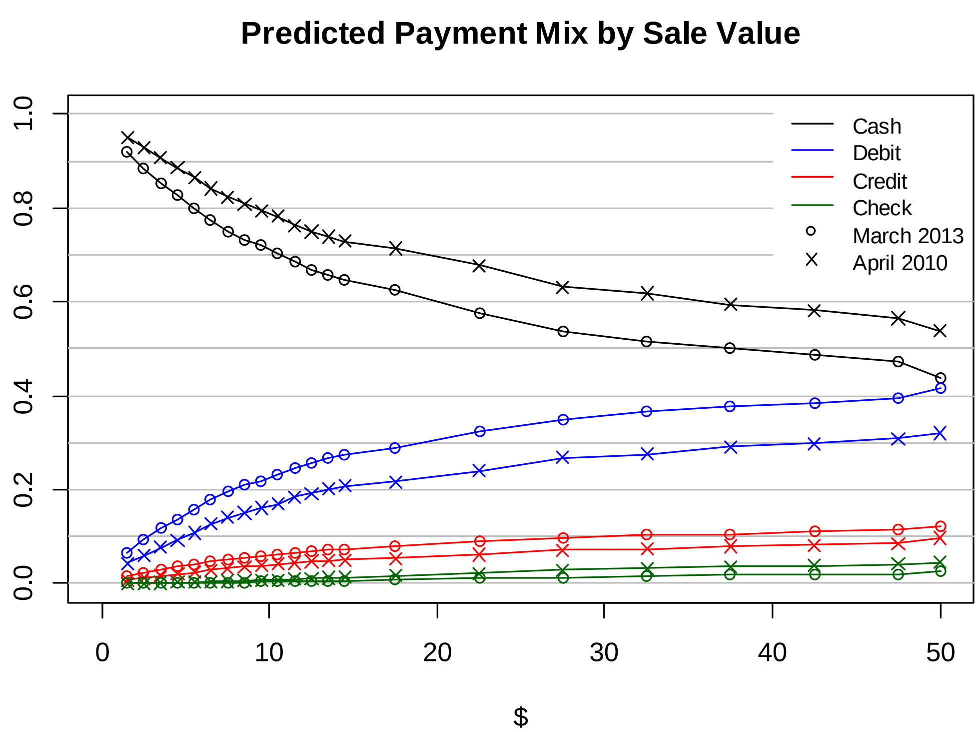

Figure 19.

As with the full sample 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 one 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.

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 the demand for currency. We turn to this topic in the next section.

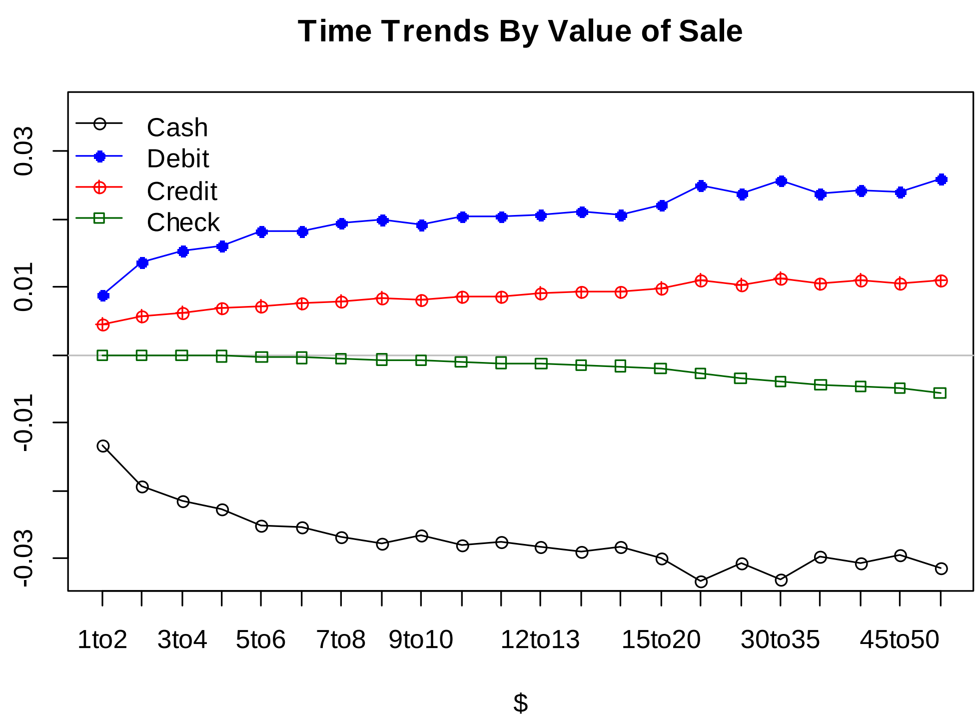

Figure 20.

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 U.S. Census Department's projections, adjusted for the level differences between the means of our sample and the national averages.[11] We forecast median household income to grow at a 2.5% 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% in 2011, and we forecast that it will reach 29.04% in 2015 and 32.04% in 2020. Bank branches per capita are forecast to increase at 1% per year. The housing vacancy rate is forecasted to decline from 13.16% in 2011 to 12.25% in 2015 and 11.75% 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 one percentage point from 2013 to 2015, and less than two 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 21, 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.

Adding the time-trend effect to the pure demographic forecasts yields our prediction for the retailer's

[11] 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.

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% of the total in March 2013, and we predict that cash will fall to 89.2% in 2015 and 82.4% in 2020. For transactions in the $5 to $6 range, cash accounted for 80.0% of transactions in March 2013, and we predict that cash will fall to 74.9% in 2015 and 61.9% in 2020. And for transactions in the $40 to $45 range the predicted decline in cash is from 48.6% of transactions in 2013 to 41.9% in 2015 and 26.1% in 2020.

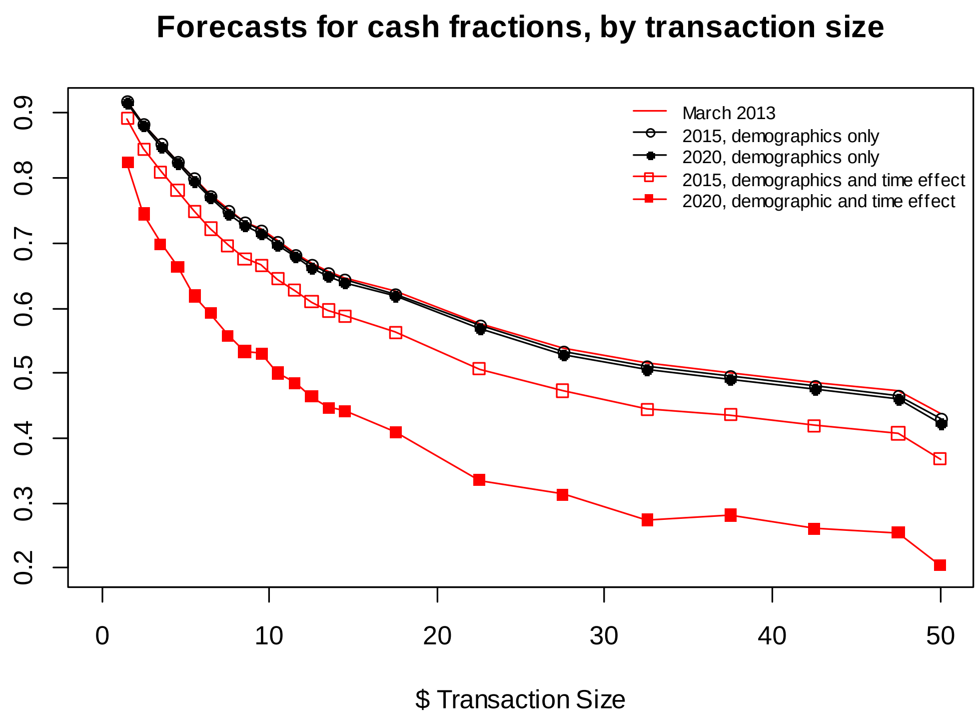

Figure 21.

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 2012 this yields cash transactions as 75.0% of the total. The forecast for 2015 is that cash will account for 69.9% of transactions, and for 2020 cash will account for 57.2% 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% nominal income growth, then the cash fraction of transactions would 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% cash share in 2020 we would forecast a 53.5% cash share, representing a decrease of 3.1 percentage points per year.

5.2 The Future of Currency

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 7 years. Whether the trend will continue is of course uncertain, but the presence of that trend in our data is quite clear. We argued in the introduction that the uniquely cash-intensive nature of our data, while rendering it unrepresentative of the U.S. economy, made it particularly well-suited to studying the behavior of cash. As such, we can use our forecasts to think about the future of currency use more broadly.

Nominal retail sales in the U.S. grew at a 3.7% rate in 2013.[12] However, currency is a feasible payment instrument only for in-person sales, and the in-person component of retail sales grew only 2.5% 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:

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

Assuming that $\mu = 0.025$ (the growth rate of in-person retail in 2012), and given our estimated $\delta = 0.0254$,

[12] 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/.

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 0.77, 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 two 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%.

Although the retailer we study represents a small fraction of the value of U.S. retail sales, in absolute terms it has a large number of cash transactions – 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] Baumol, W.J. (1952). "The transactions demand for cash: an inventory theoretic approach." *Quarterly Journal of Economics* 66 (4), 545–556.

[2] Borzekowski, R., Kiser, E.K., Ahmed, S. (2008). "Consumers' use of debit cards: patterns, preferences, and price response. *Journal of Money, Credit, and Banking*," 40 (1), 149-172.

[3] Borzekowski, R., Kiser, E.K. (2008). "The choice at the checkout: Quantifying demand across payment instruments. *International Journal of Industrial Organization*, 26(4), 889–902.

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

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

[6] Dotsey, Michael and Pablo 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.

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

[8] Fung, B., Huynh, K., & Sabetti, L. (2011). "How do you pay? The role of incentives at the point-of-sale." Working Paper 2011-23, Bank of Canada.

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

[10] Judson, Ruth A. and Richard D. Porter (2004)."Currency demand by federal reserve cash office: what do we know?" *Journal of Economics and Business*, 56 (4), 273-285.

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

[12] 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

[13] McAndrews, James and Zhu Wang (2012), "The Economics of Two-Sided Payment Card Markets: Pricing, Adoption and Usage." Research Working Paper 12-06, Federal Reserve Bank of Richmond.

[14] Mullahy, John (2010). "Multivariate fractional regression estimation of econometric share models," NBER Working Papers 16354.

[15] Papke, Leslie E and Jeffrey 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.

[16] Rochet, J. C. and Tirole, J. (2006). "Two-sided markets: A progress report." *RAND Journal of Economics*, 37, 645-667.

[17] Rochet, J. C. and Tirole, J. (2002), "Cooperation among Competitors: Some Economics of Payment Card Associations," *RAND Journal of Economics* 33: 549-570.

[18] Rysman, M. (2009). "The economics of two-sided markets." *Journal of Economic Perspectives*, 23, 125-144.

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

[20] Telyukova, I.A. and R. Wright (2008). "A Model of Money and Credit, with Application to the Credit Card Debt Puzzle." *Review of Economic Studies*, 75(2), 629–647.

[21] Tobin, J. (1956). "The interest elasticity of the transactions demand for cash." *The Review of Economics and Statistics* 38 (3), 241–247.

[22] 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 Table 2.

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 Table 2.

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 Table 2.

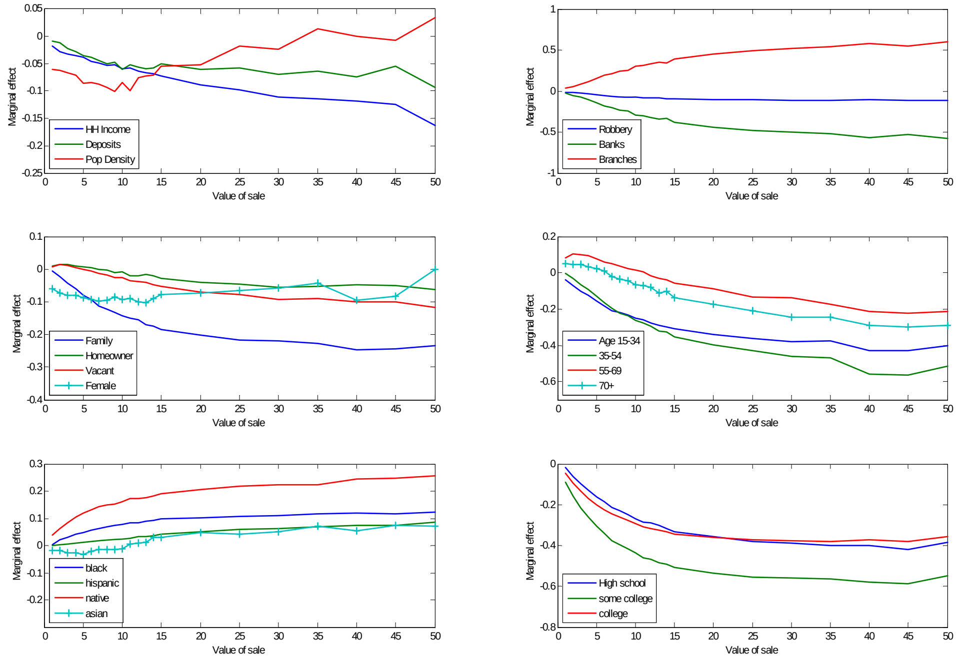

Figure A1. Cash marginal effects by transaction size.

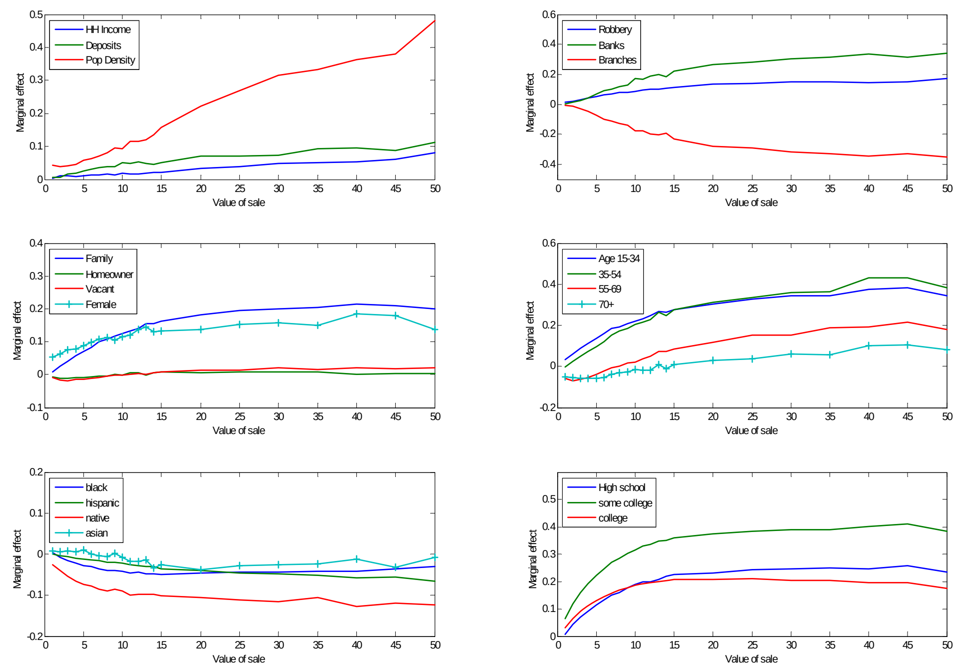

Figure A2. Debit marginal effects by transaction size.

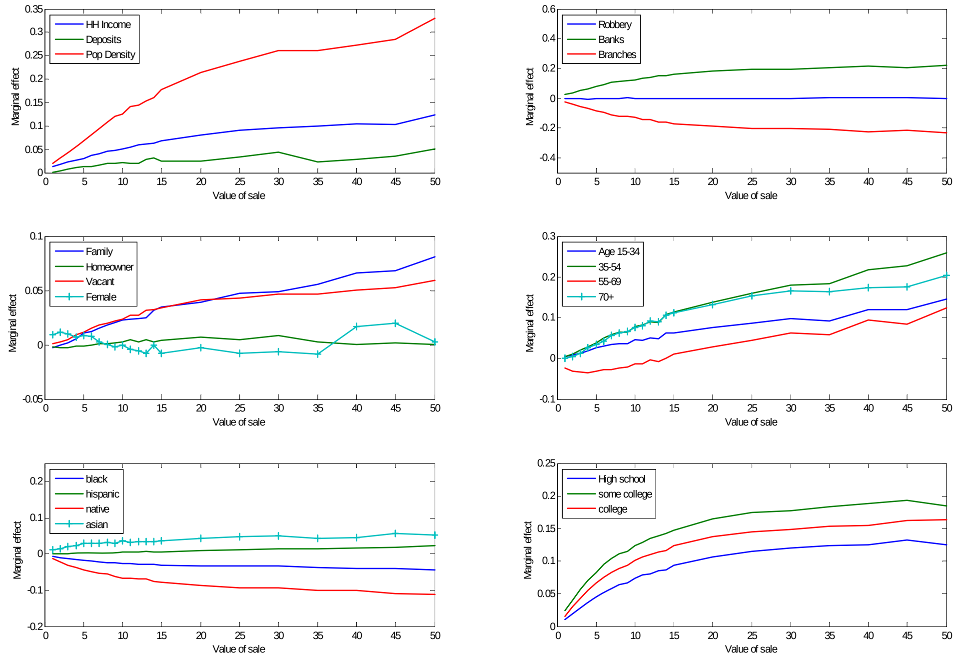

Figure A3. Credit marginal effects by transaction size.

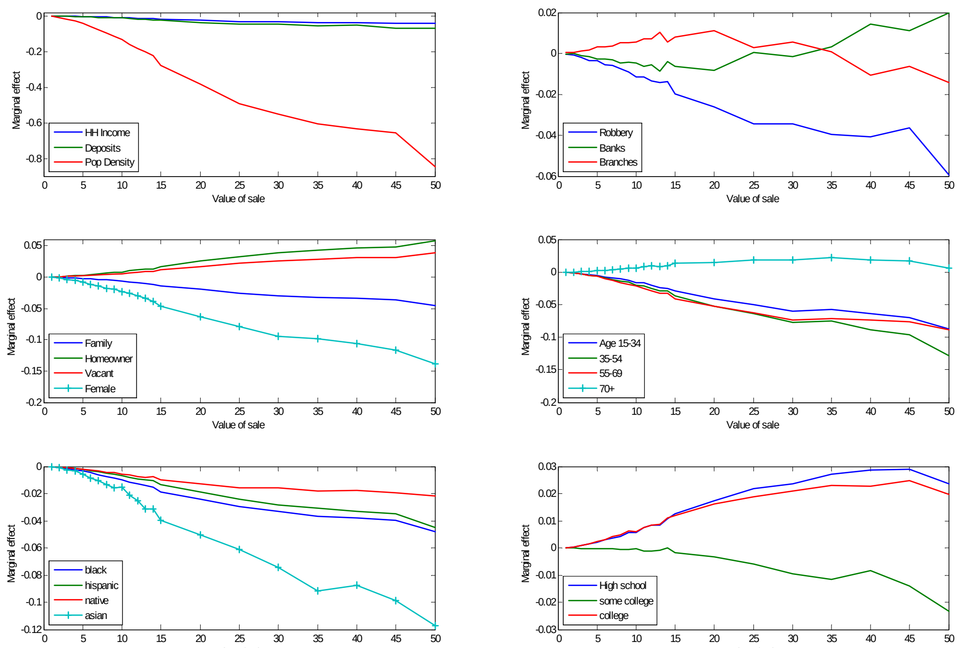

Figure A4. Check marginal effects by transaction size.

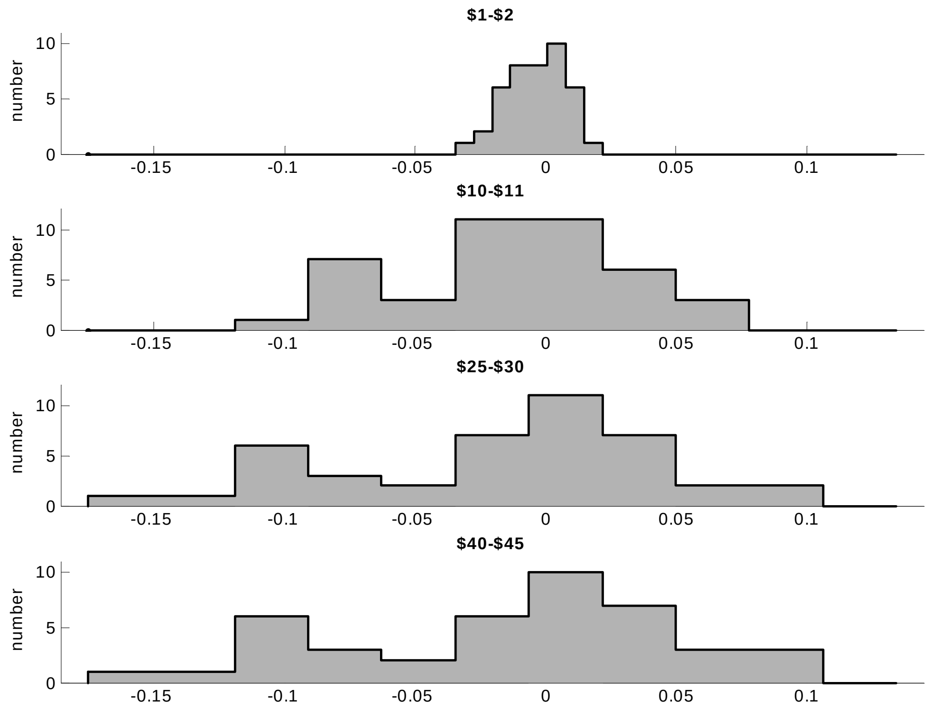

Figure A5. Debit: state effects by transaction size.

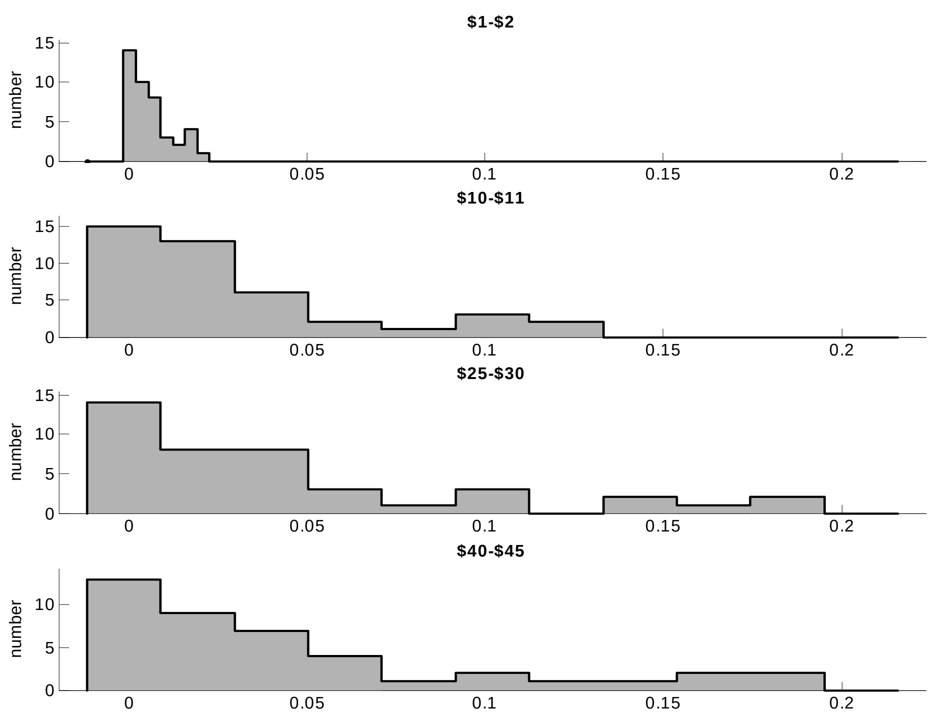

Figure A6. Credit: state effects by transaction size.

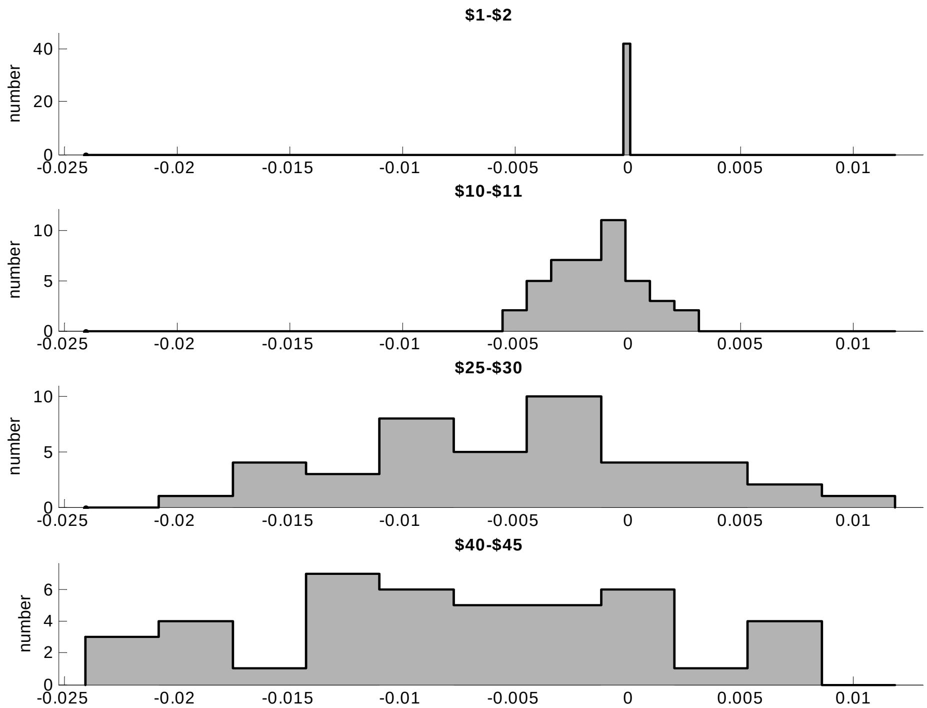

Figure A7. Check: state effects by transaction size.

Table A4. Debit card: 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: 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: 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

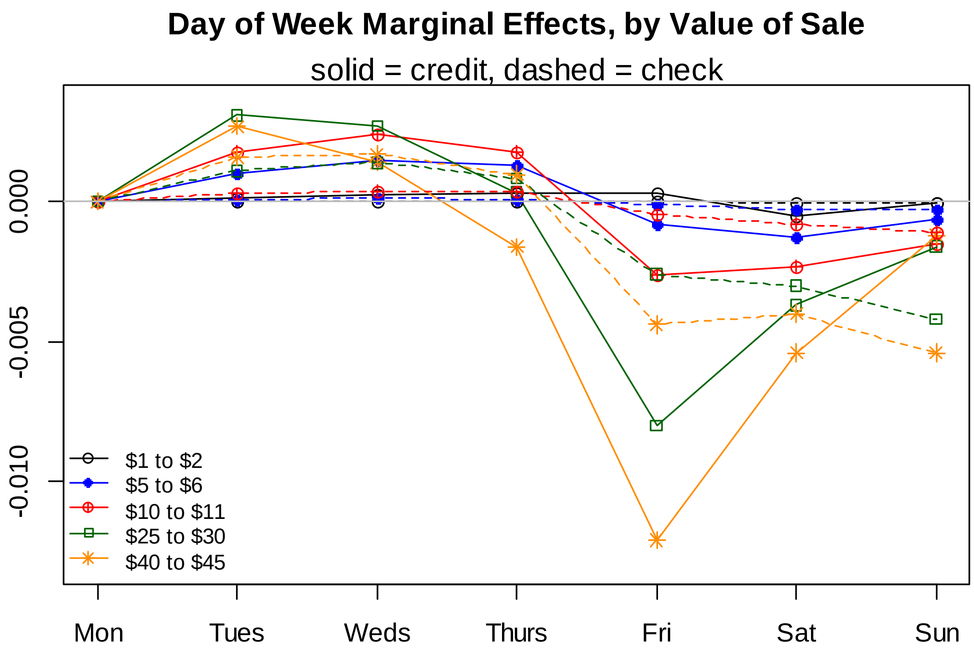

Figure A8.

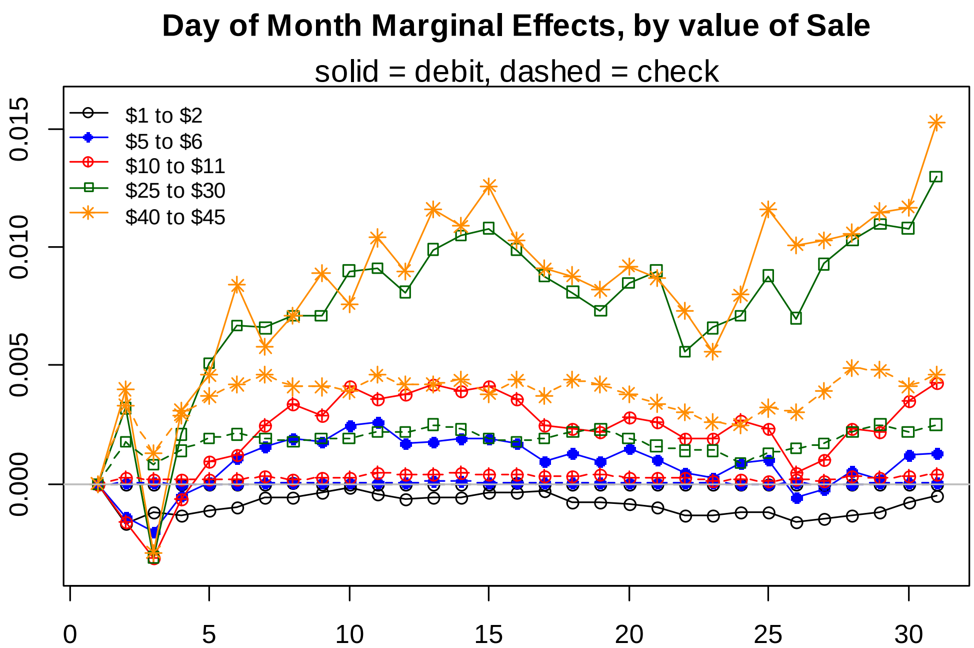

Figure A9.

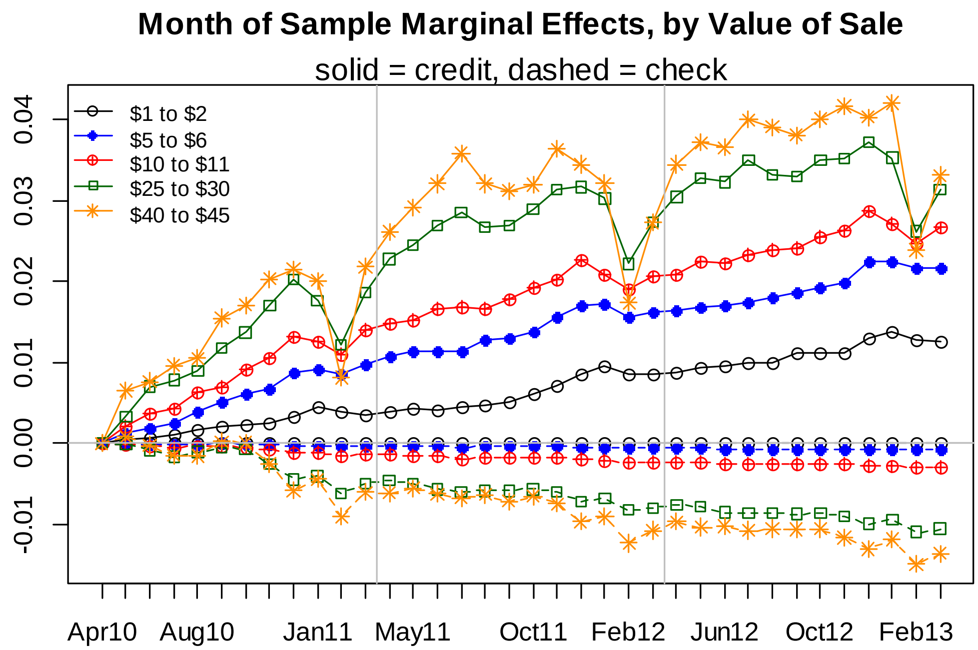
Month of Sample Marginal Effects, by Value of Sale
solid = credit, dashed = check
$1 to $2
$5 to $6
$10 to $11
$25 to $30
$40 to $45
-0.01
0.00
0.01
0.02
0.03
0.04
Apr10
Aug10
Jan11
May11
Oct11
Feb12
Jun12
Oct12
Feb13

Figure A10.