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

Zhu Wang and Alexander L. Wolman

Federal Reserve Bank of Richmond

June 4, 2015

Getting the Balance Right

ECB and Bank of Finland joint conference, Helsinki

Introduction

- How do consumers choose to pay at the point of sale?
  - Key to understanding transaction demand for money, evolution of payments system
  - Rich data on non-cash payments from bank surveys
  - Cash? Mainly from small-sample consumer surveys
- We use merchant transaction data, as in Klee (2008).
  - 1 large discount chain, thousands of locations in U.S.
  - 2 billion retail transactions ($\approx$ millions of consumers)
  - 3 full years, from April 1, 2010 to March 31, 2013

Introduction

- We estimate the relationships between location-specific explanatory variables and payment choice.
- We also estimate payment patterns associated with day of week, day of month, seasonal cycles and a time trend.
- Our estimates have implications for theories of money demand and payment choice (e.g., Baumol 1952, Tobin 1956, Prescott 1987, Freeman and Kydland 2000, Lucas and Nicolini 2013, Alvarez and Lippi 2009, 2014).
- We also project future use of currency in this sector: cash still dominates discount retail, but its share is falling at approximately 2.5 pps per year.

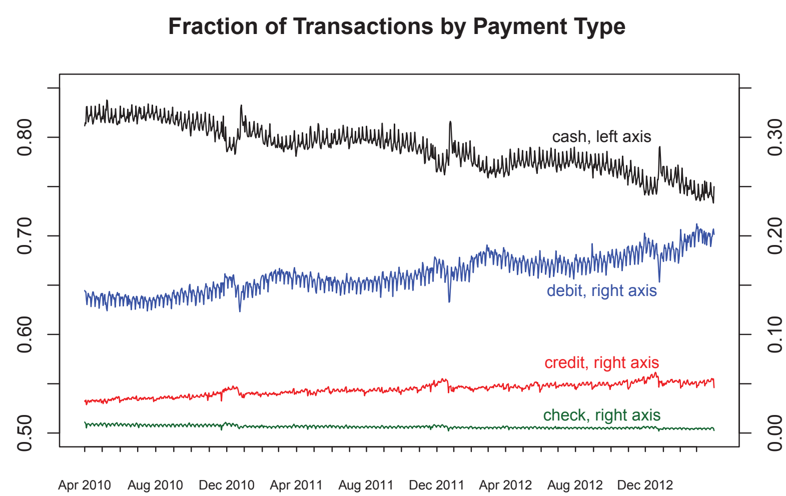
Outline

1. Data: transactions, zip-code level explanatory variables.
2. Econometric model: FMLogit for payment shares.
3. Results for estimating overall payment mix: data aggregated by payment type to zip-code day.
4. Results for separate models by transaction size.
5. Projections of future cash use.
6. Conclusion.

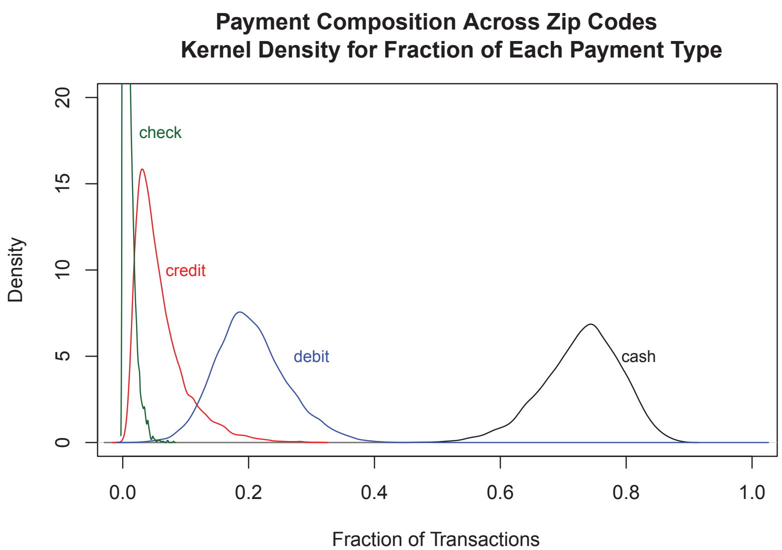
Transactions data

- Discount retailer, several ’000 stores, dozens of states across most of the U.S.
- Data covers April 1, 2010 - March 31, 2013.
- We restrict to cash, debit, credit, check (the four general purpose payment types).
- More than 1.75 million transactions per day.
- Median transaction size $\approx$ \$7.

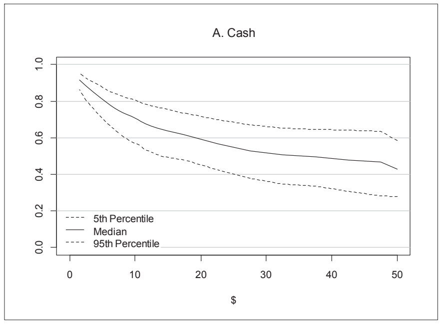
Payment variation across time

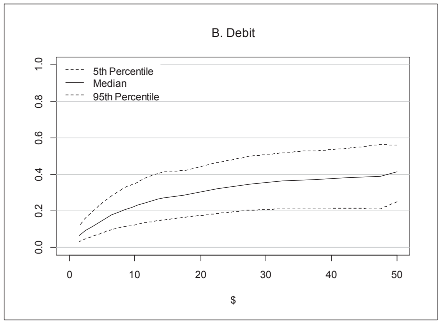


Payment variation across locations, March 2013



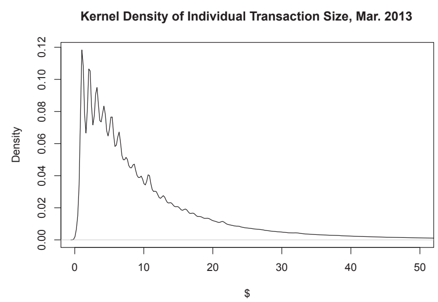
Payment variation across transaction sizes: level and dispersion, March 2013

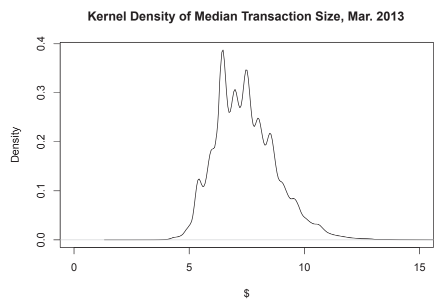




Transaction size distribution, March 2013

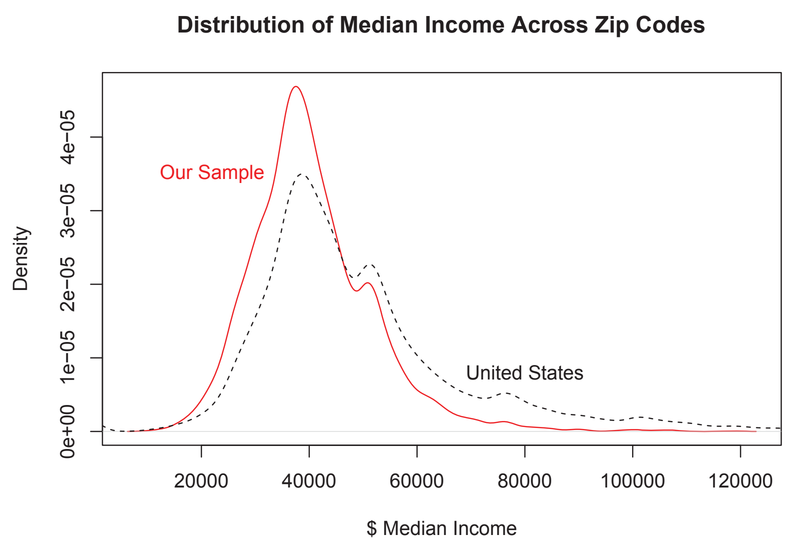
Transactions concentrated below $15





Explanatory variables

- Zip-code level variables, fixed across time.
  - Cash use in inventory-theoretic models
    - bank concentration (HHI), bank branches per capita
    - robbery rate
  - Adoption of non-cash payment means
    - median household income, deposits per capita
    - population density
  - Demographic variables: age, sex, race, education, housing status, family status
- State dummies, fixed across time.
- Time dummies: day-of-week, day-of-month, month-of-sample.


Distribution of Median Income Across Zip Codes
Our Sample
United States
Density
0e+00
1e−05
2e−05
3e−05
4e−05
20000
40000
60000
80000
100000
120000
$ Median Income

Empirical model (FMLogit, Mullahy 2010)

- *Reduced-form* model of $s_{i,k}$ = share of payment type k in zip-code day i.
- Shares sum to one, can be zero or one $\Rightarrow$ FMLogit:

$$E[s_k \mid x] = G_k(x;\beta) = \frac{\exp(x\beta_k)}{\sum_{m=1}^{4} \exp(x\beta_m)}.$$

Normalize $\beta_{\text{cash}} = 0$ for identification:

$$\underset{k=1,2,3}{G_k} = \frac{\exp(x\beta_k)}{1 + \sum_{m=1}^{3} \exp(x\beta_m)}, \quad G_{\text{cash}} = \frac{1}{1 + \sum_{m=1}^{3} \exp(x\beta_m)}.$$

- x are zip-code level explanatory vars., state/time dummies.

Estimating overall payment mix

- Payment shares based on all transactions for a zip-code day, 4.5 million observations (zip-code days).
- Include median transaction size as an explanatory variable.
- For continuous x variables, report marginal effects evaluated at the mean.
- For dummies, report “discrete effects” evaluated at mean.

Findings: zip-code level variables (1)

Economic Variables	Cash	Debit	Credit	Check
Median transaction size	-0.018*	0.012*	0.005*	0.001*
HHI (bank concentration)	0.030*	-0.023*	-0.010*	0.003*
HHI*metro	-0.050*	0.032*	0.024*	-0.005*
Branches per capita	0.007*	-0.005*	-0.004*	0.001*
Robbery rate	-0.054*	0.063*	0.000	-0.010*
Median household income	-0.033*	0.005*	0.036*	-0.008*
Deposits per capita	-0.006*	0.016*	0.000	-0.010*
Population density	-0.038*	0.079*	0.091*	-0.131*

* significant at 1%.

Branches per capita = number of bank branches per 100 residents in a zip code. Median household income in $100,000 per household. HHI measured at county or MSA level, transformed to lie between 0 and 1. Deposits per capita in $10,000 deposits per resident in a zip code. Population density is measured in 100,000 residents per square mile in a zip code. Robbery rate = number of robberies per 100 residents in a county.

Findings: zip-code level variables (2)

Demographics		Cash	Debit	Credit	Check
Family households		-0.098*	0.089*	0.016*	-0.006*
Female		-0.052*	0.080*	-0.005*	-0.023*
Age share:	15-34	-0.184*	0.163*	0.034*	-0.013*
	35-54	-0.152*	0.115*	0.053*	-0.016*
	55-69	0.031*	-0.000	-0.013*	-0.018*
	$\geq$ 70	-0.024*	-0.038*	0.054*	0.007*

* significant at 1%.

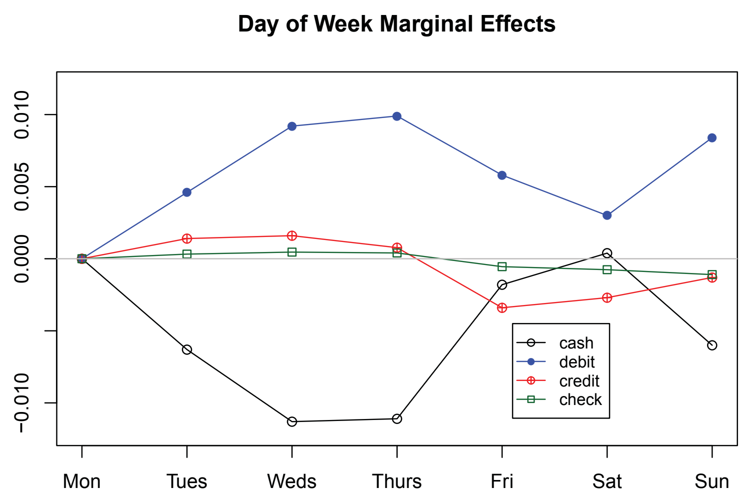
Findings: zip-code level variables (3)

Demographics	Cash	Debit	Credit	Check
Race: black	0.055*	-0.025*	-0.020*	-0.011*
hispanic	0.024*	-0.019*	0.003*	-0.007*
native	0.133*	-0.074*	-0.052*	-0.007*
asian	-0.018*	0.001	0.032*	-0.022*
Educ: high school	-0.202*	0.138*	0.057*	0.007*
some college	-0.322*	0.233*	0.088*	0.001*
college	-0.225*	0.140*	0.079*	0.007*

* significant at 1%.

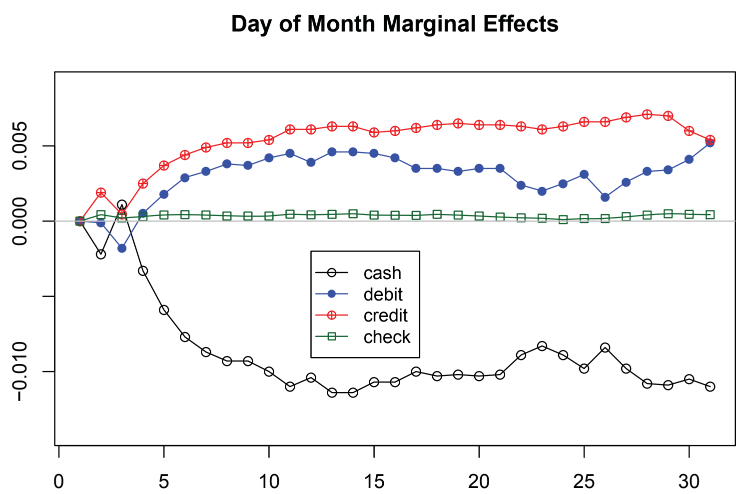
Day-of-week effects

Interesting patterns, but small magnitudes.



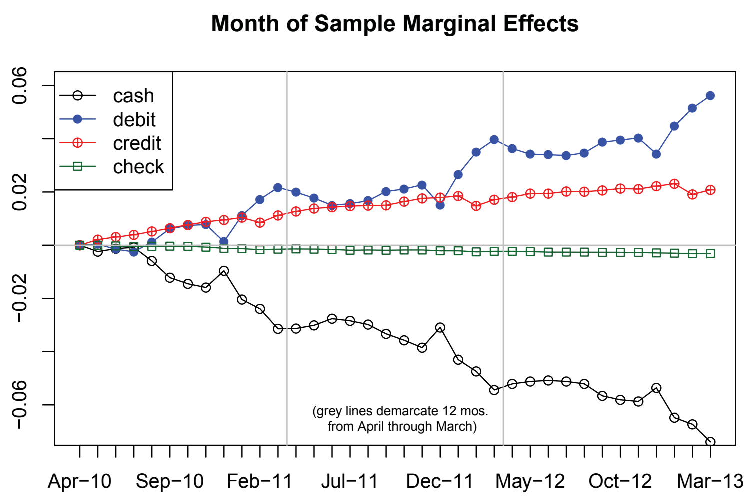
Day-of-month effects

Interesting patterns, but small magnitudes.



Month-of-sample effects

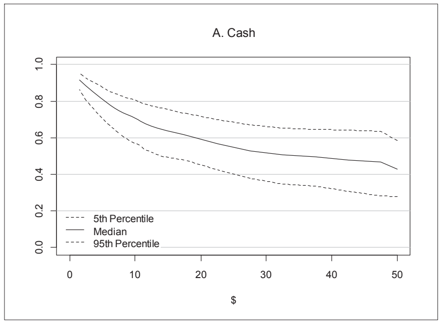
Interesting patterns and large magnitudes!

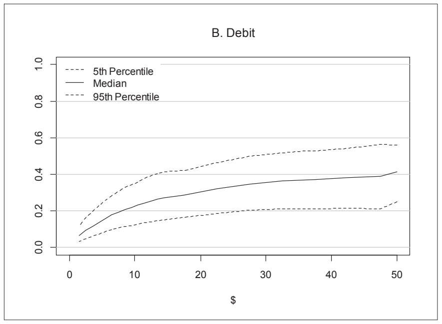


Separate models by transaction size

- The benchmark model provides a useful summary of consumer payment mix across locations and dates.
- For a given zip code in a given day, the overall payment mix depends on **consumer payment choice at each transaction size** combined with **the transaction size distribution**, possibly through consumer characteristics, location/time fixed effects, and median transaction size.
- To better understand **consumer payment choice at each transaction size**, we take a step further, run separate models by transaction size, which
  - take into account individual transaction sizes
  - allow transaction size to affect both coeffs and constants

Payment variation across transaction sizes: level and dispersion, March 2013



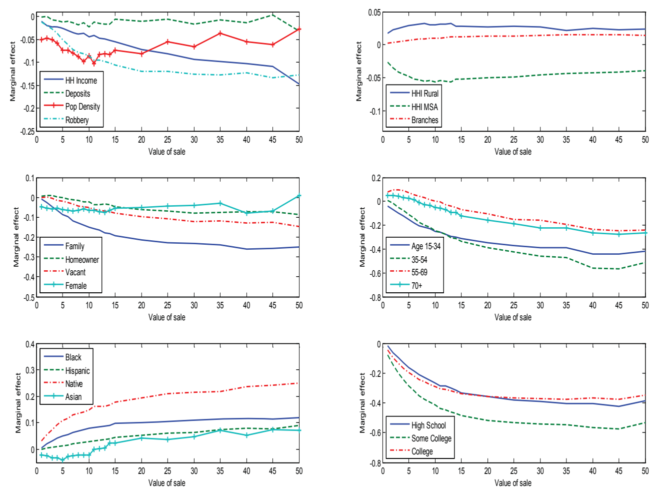


Description, and summary of findings

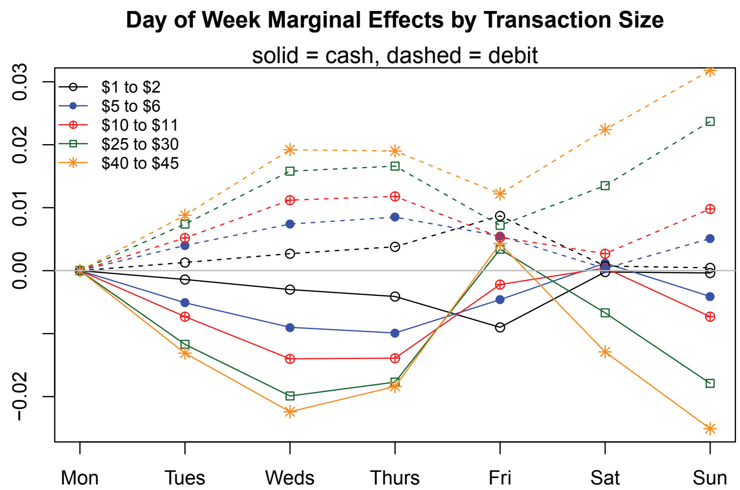
- Aggregate data to zip-code day by transaction size.
- 22 separate regressions:$1-$2, $2-$3,...,$14-15, $15-$20, $20-$25,...,>$50.
- Similar number of observations to benchmark. Number of underlying transactions between 11 and 199 million.
- Same explanatory variables but allow their coefficients to differ across size class regressions.
- The models fit data very well:
    - Marginal effects amplify with transaction size
    - Allowing coefficients to vary across transaction size is important for explaining variation in *levels* of shares, as well as *dispersion*

Amplification of marginal effects: cash

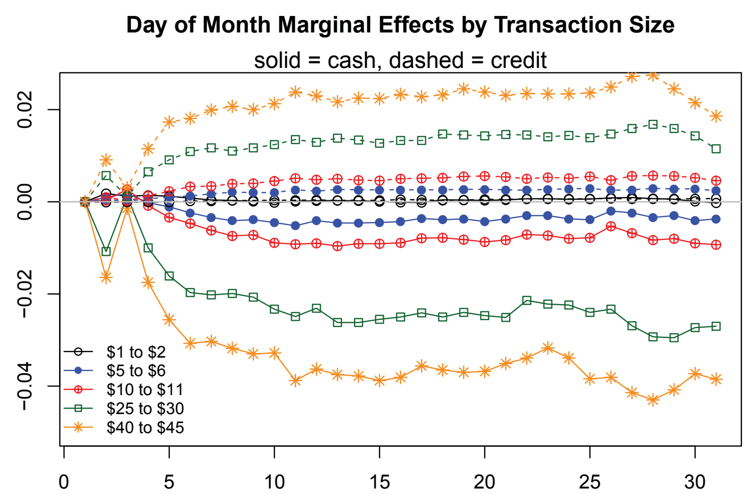
Cash



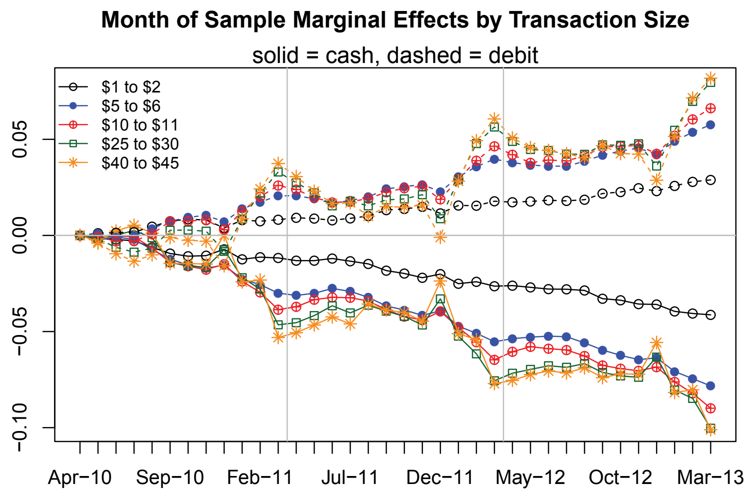
Day of week effects



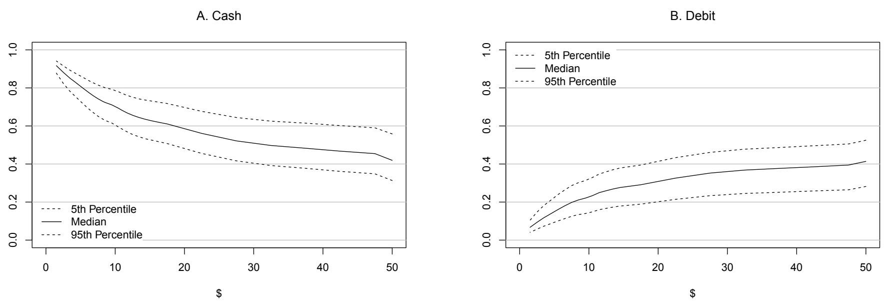
Day of month effects



Month of sample effects

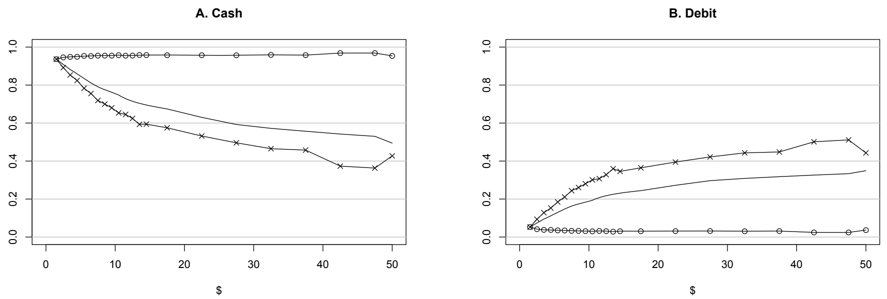


Predicted payment variation across transaction size: level and dispersion, March 2013



- model does a good job at fitting data
- how does changing coefficients of X (amplification) explain payment shares across transaction sizes?

Decomposing the level and dispersion



- "x"-lines hold fixed coeffs on zip-code-level variables
- "o"-lines hold fixed all other terms: state/time/constant
- level and dispersion explained by zip-code-level variables!

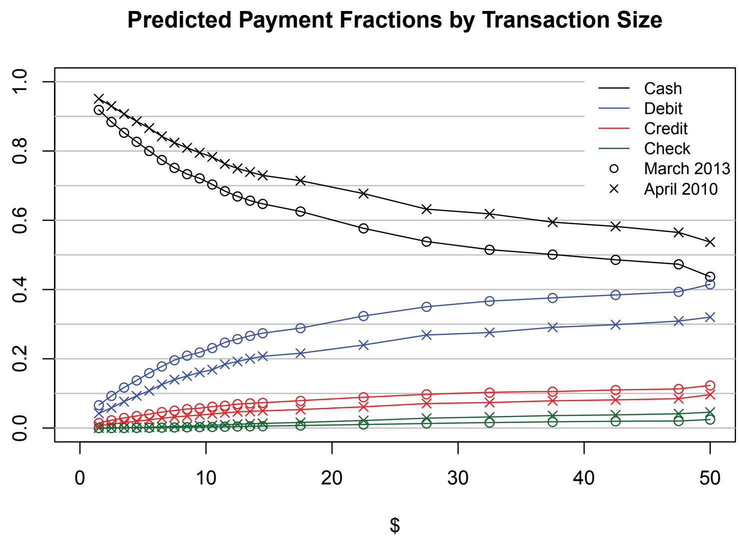
Explaining the level and dispersion

- Theories of money demand and payment choice suggest we can view consumers as each having some threshold above which they switch from using cash to non-cash payment means.
- The *level* effect: for any location-specific distribution of thresholds, at a higher transaction size there will be a higher fraction of consumers using non-cash payment means because their thresholds have been crossed.
- The *dispersion* effect: As the transaction size increases, consumers in a location with easier access to non-cash options switch increasingly further away from cash compared to locations that do not have those options.

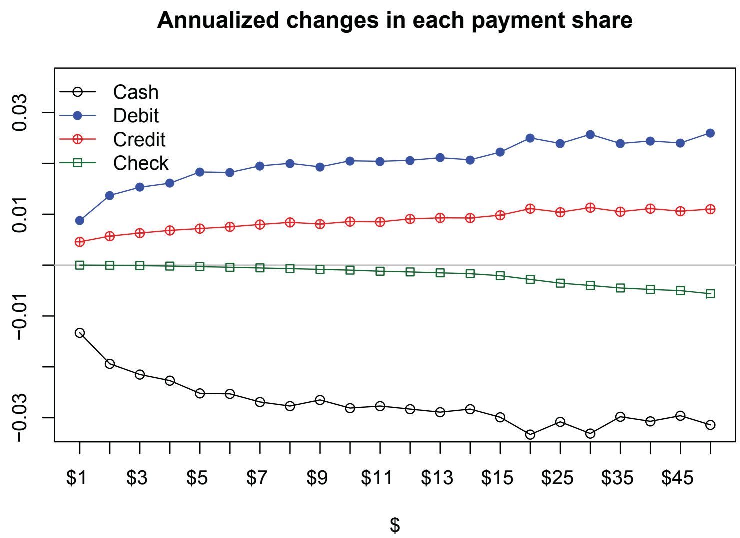
Forecasting the payment mix

- Our size-class regression models can be used to forecast the future composition of payments at the discount retailer.
- The cash component of those forecasts is related to the level of currency use in transactions, which in turn has implications for currency demand.
- We present forecasts specific to the discount retailer, which may be informative about the level of overall currency use going forward.

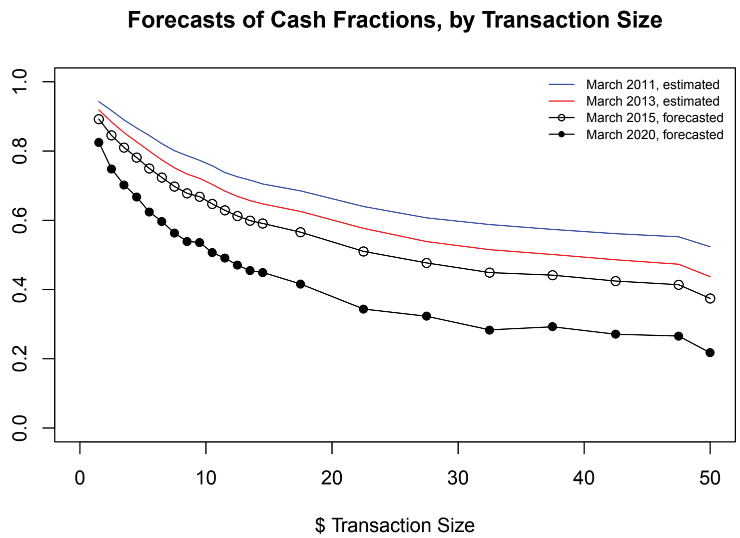
Shifts in the predicted payment mix over time



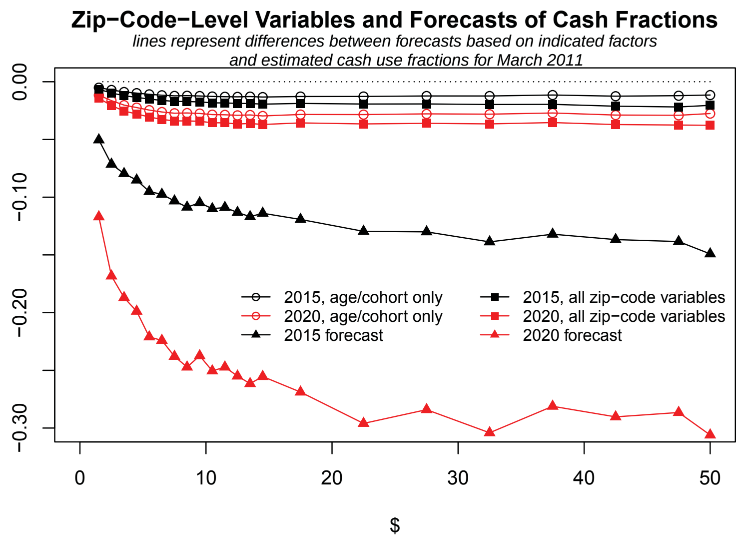
Growth of payment types by transaction size



Projecting currency use in U.S. discount retail



Projecting currency use: zip-code variables



Conclusions

- Data from a discount retailer: 3 years, thousands of locations $\Rightarrow$ 2 billion transactions.
  - Payment mix varies across time and locations
  - Payment mix varies with size of transaction
  - Cross-sectional dispersion increases with transaction size
- Estimates from FMLogit model of payment mix:
  - Consistent with basic models of money demand and payment choice (e.g. Freeman and Kydland 2000)
  - Account for both level and dispersion of payment choice across transaction sizes with coefficients that vary across transaction size (suggesting a heterogenous agent framework for money demand theories)
  - Project cash share declining 2.5pp per year, suggesting a declining number of total cash transactions

Directions for future research

- Consider theoretical models that combine payment adoption decision, payment usage decision and explicit inventory-theoretic approach to cash use.
    - All of these features seem to be important in our data, though we cannot fully disentangle them
- Collect matched merchant-consumer payments data, to estimate those models and to disentangle the aforementioned features.