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

Zhu Wang and Alexander L. Wolman

Federal Reserve Bank of Richmond

April 3, 2014

Overview

- What is the composition of in-person retail payments?
  - 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
  - 2 billion retail transactions
  - 3 yrs, thousands of zip-code locations
- Interaction between demographic variables & transaction size important for explaining payment composition.
- Cash still dominates discount retail, but share is falling at approximately 2.5 pps per year.

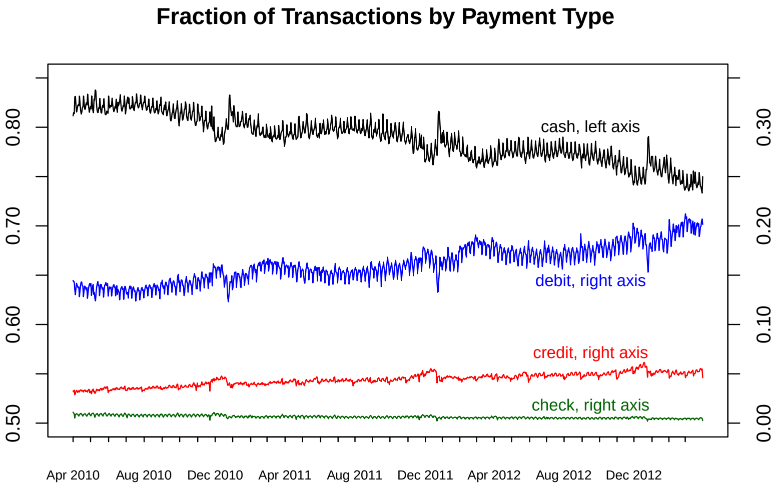
Outline

1. Data: transactions, and zip-code level expl. variables.
2. Econometric model: fmlogit for shares each payment type.
3. Results for benchmark model: data aggregated by payment type to zip-code day.
4. Results for separate models by transaction size.
5. Conclusion.

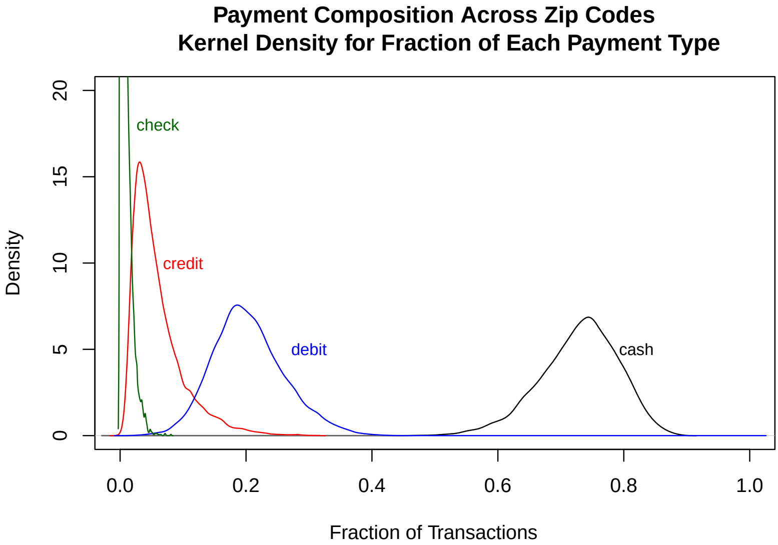
Transactions data

- Discount retailer, several ’000 stores, dozens of states.
- Data covers April 1, 2010 - March 31, 2013.
- We restrict to cash, debit, credit, check.
- More than 1.75 million transactions per day.
- Median transaction size $\approx$ \$7.

Variation across time

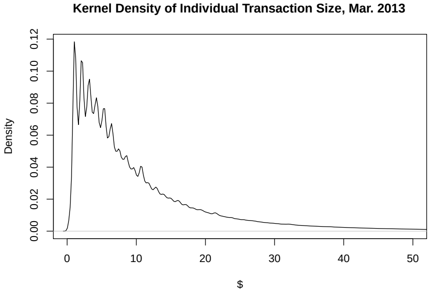


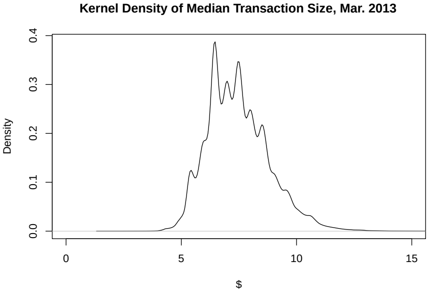
Variation across locations, March 2013



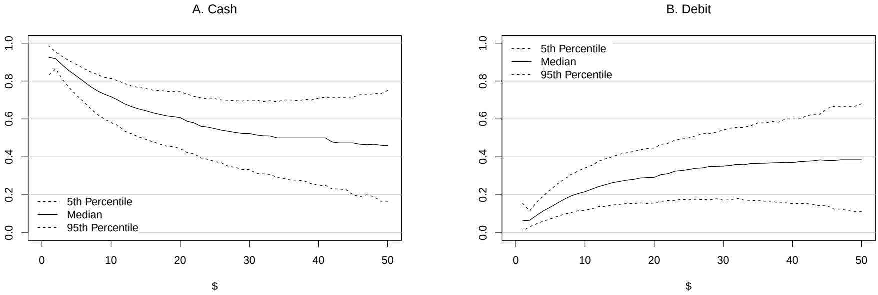
Transaction size variation, March 2013

Transactions concentrated below $15



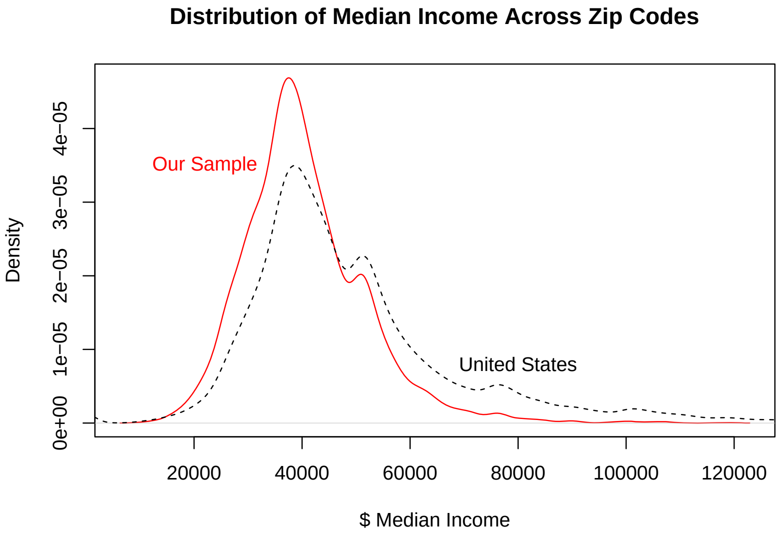


Payment shares and transaction size: level and dispersion, March 2013



Explanatory variables

- Zip-code level variables, fixed across time.
  - Variables related to money demand, cost of different payment types:
    - household income, banks/branches/deposits per capita.
    - population density, robbery rate
  - 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))

- 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:

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

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

Background for benchmark results

- Payment shares based on all transactions for a zip-code day, 4.5 million observations.
- 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.

Selected results: zip-code level variables (1)

	Cash	Debit	Credit	Check
Median transaction size	-0.017	0.012	0.005	0.001
Bank Branches per capita	0.243	-0.133	-0.113	0.003
Median household income	-0.048	0.015	0.042	-0.009
Deposits per capita	-0.036	0.035	0.016	-0.014
Banks per capita	-0.234	0.128	0.109	-0.002
Population density	-0.039	0.090	0.097	-0.148
Robbery rate	-0.046	0.063	-0.006	-0.011

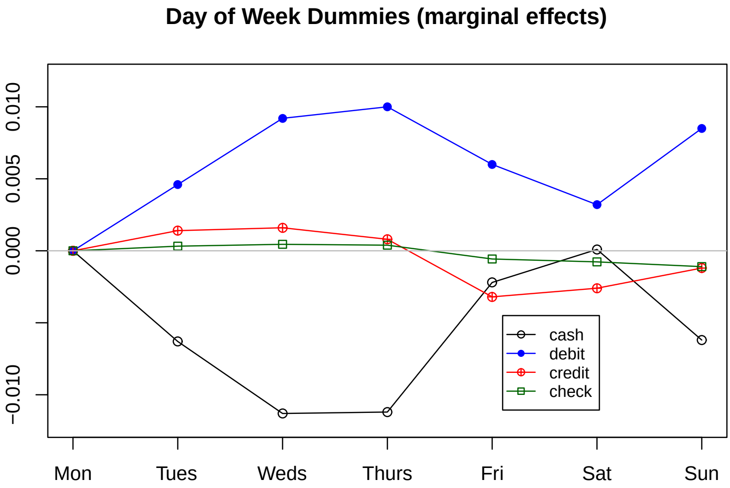
Selected results: zip-code level variables (2)

		Cash	Debit	Credit	Check
Family hhlds		-0.093	0.088	0.013	-0.008
Age share:	15-34	-0.186	0.169	0.035	-0.017
	35-54	-0.174	0.134	0.061	-0.022
	55-69	0.039	-0.003	-0.014	-0.022
	$\geq$70	-0.034	-0.030	0.058	0.006
Education:	hs	-0.202	0.137	0.059	0.006
	some college	-0.342	0.246	0.097	-0.001
	college	-0.227	0.140	0.081	0.006

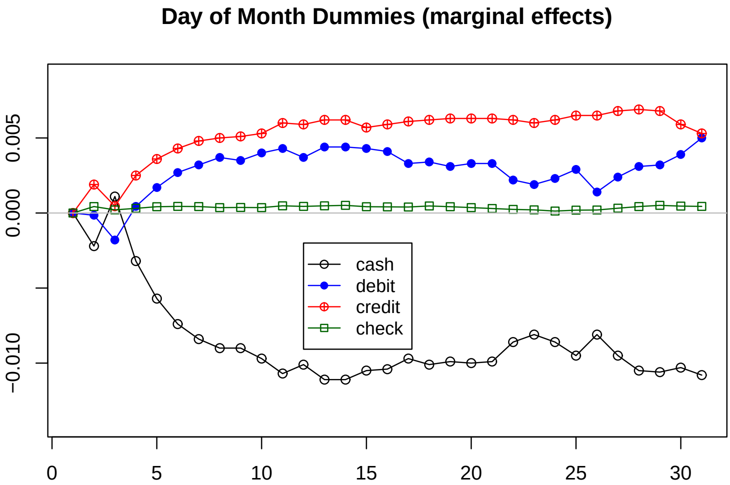
Selected results: state effects

	Cash	Debit	Credit	Check
Top States	NJ	AZ	MN	SD
	NY	ID	ND	ND
	MI	NV	SD	MN
	VT	NM	OK	OK
	DE	FL	OH	CO
Bottom States	FL	MD	IA	NH
	TX	NY	AR	NY
	NM	ND	NV	AZ
	ID	SD	MS	DE
	AZ	MN	NJ	NJ

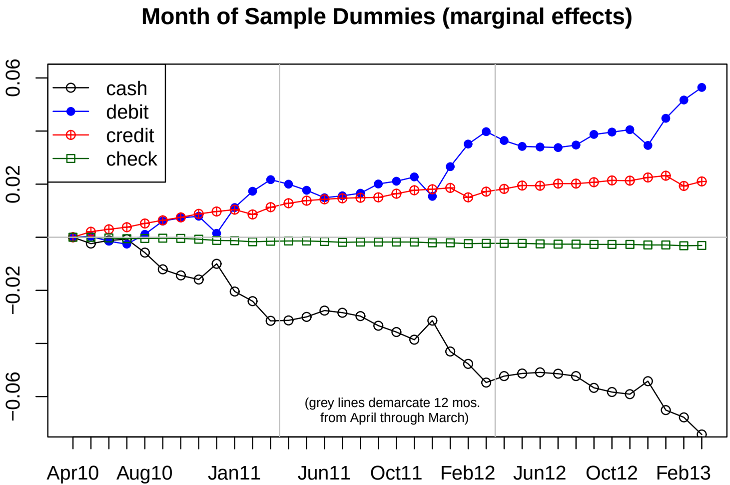
Selected results: time effects (day-of-week)



Selected results: time effects (day-of-month)



Selected results: time effects (month-of-sample)



Separate models by transaction size

- Relationship between payment shares and explanatory variables may vary with transaction size.
- Benchmark case is highly restrictive:
  - Transaction size can only shift constant, not coeffs. on X
  - Effect on levels limited to median transaction size in zip-code day
- Alternative: aggregate to zip-code day separately for \$1-\$2, \$2-\$3 etc.
  - Transaction size can affect coefficients as well as constants
  - Trans. size can matter within, not just across zip-code days

Formal Motivation

- fmlogit model implies

$$\frac{S_k}{S_M} = \exp(c + X\beta_k) => \ln\frac{S_k}{S_M} = c + X\beta_k.$$

- If allow c but not β_k to vary with transaction size (v), then

$$\ln\frac{S_k(v)}{S_M(v)} = c(v) + X\beta_k$$

and

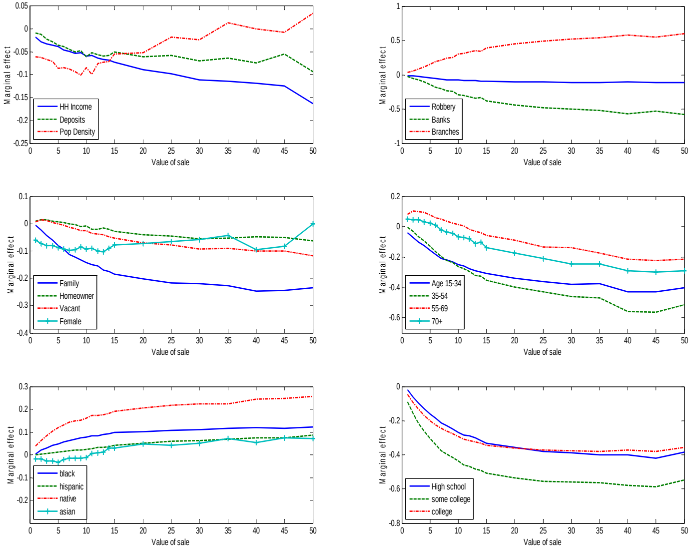
$$\frac{\partial E(\ln\frac{S_k}{S_M})}{\partial v} = c'(v) \quad \text{and} \quad \frac{\partial Var(\ln\frac{S_k}{S_M})}{\partial v} = 0.$$

- Variation in level of payment shares must come from intercept, and dispersion in payment shares must be constant across v!

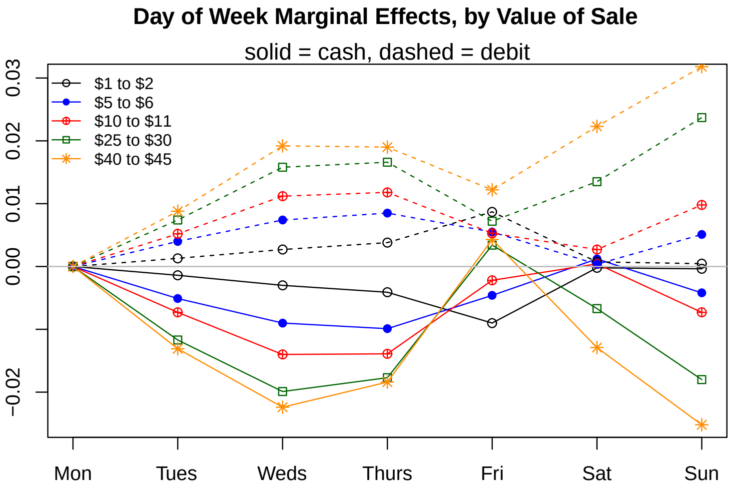
Description, and Summary of Results

- 22 separate regressions:$1-$2, $2-$3,...,$14-15, $15-$20, $20-$25,...,$>$$50.
- Same explanatory variables, except omit transaction size.
- Similar number of observations to benchmark. Number of underlying transactions between 11 and 199 million.
- Results:
  - 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*

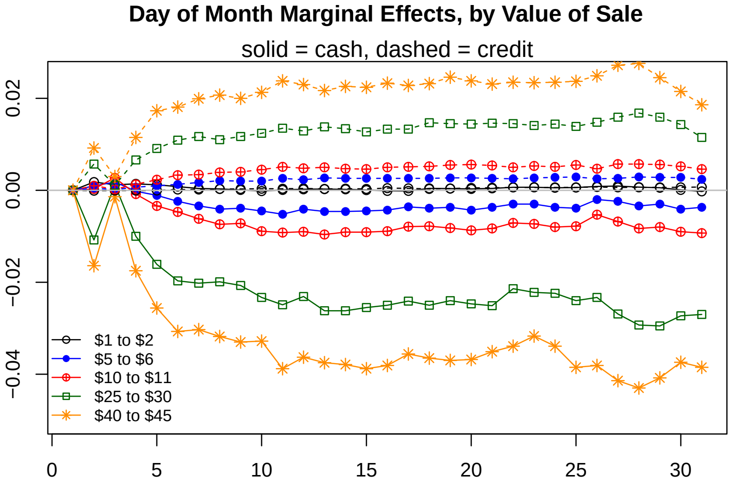
Amplification of marginal effects



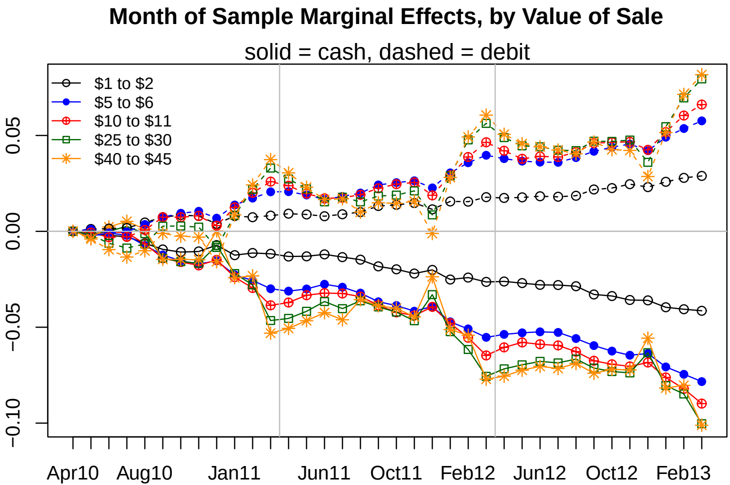
Day of week effects



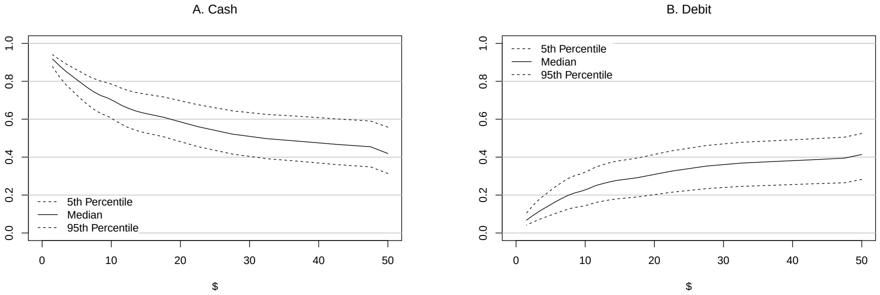
Day of month effects



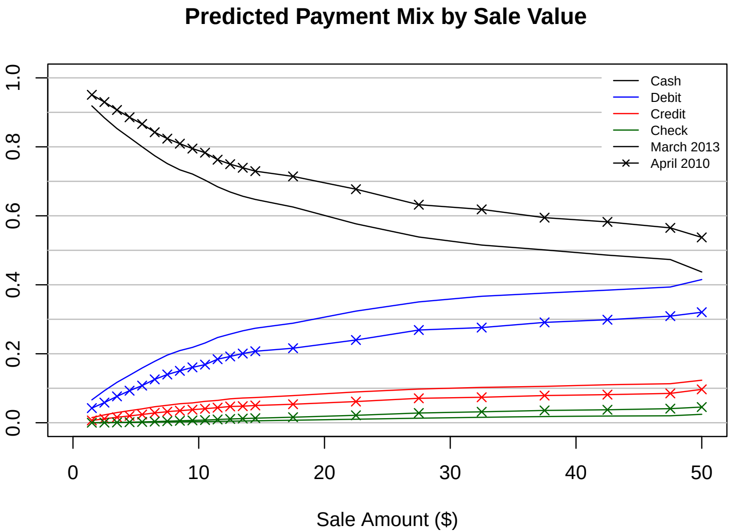
Month of sample effects



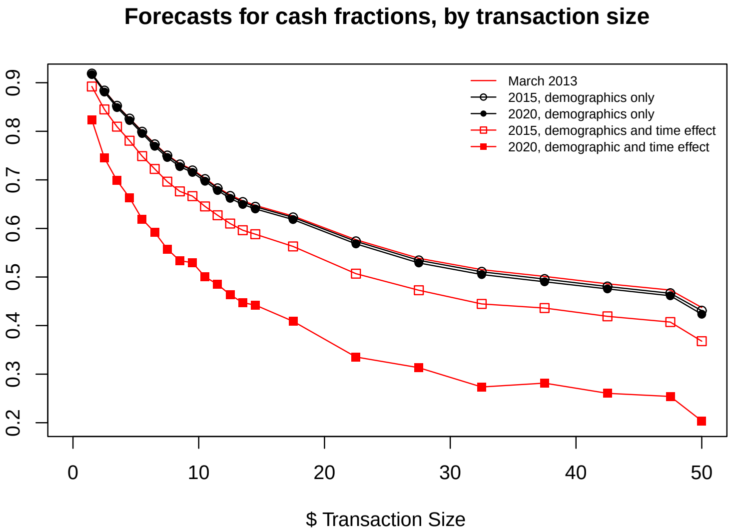
Predicted payment shares and transaction size: level and dispersion, March 2013



Shifts in the predicted payments mix



Projecting the future of cash



Conclusions

- Analyze payments at a discount retailer: 3 years, thousands of locations $\Rightarrow$ 2 billion transactions.
- Features of data:
  - 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:
  - Support generalized inventory-theoretic demand for cash, with multiple means of payment
  - Account for both level and dispersion of payment choice across transaction sizes with coefficients that vary across transaction size
  - Project cash share declining 2.5pp per year