Preference Versus Actual Choice: Evidence from a Payments Survey*

Oz Shy†
Federal Reserve Bank of Atlanta

August 31, 2026

Abstract

Using survey data on consumer payment choice and machine learning prediction algorithms, the paper analyzes the degree to which consumers' stated preference for a payment method can predict the actual use of the payment method. I find that consumers' preferred payment method remains the strongest predictor of credit and debit card payments even after individual-specific and transaction-specific variables are added to the prediction model. In contrast, predictions of cash payments are more sensitive to some transaction-specific characteristics in addition to the stated preference for paying cash.

Keywords: Preference versus actual choice; preferred payment method; actually used payment method; nonbill in-person payments; machine learning classification trees; random forest.

JEL Classifications: D91, E42, G41.

*I thank Aditi Routh for most valuable comments on an earlier draft. The views expressed here are the author's and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

†*E-mail*: Oz.Shy@atl.frb.org and ozshy@ozshy.com. Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree St. NE, Atlanta, GA 30309, U.S.A.

"You can't always get what you want." *The Rolling Stones*, 1969

1. Introduction

Do consumers make choices based on their stated preferences? This paper analyzes this question using survey data on consumer payment methods. The survey collects information on consumers' preference for a payment method (cash, credit card, debit card, and others). The survey also records the payment methods that consumers actually used at the point-of-sale (POS).

The first part of the paper constructs a simple confusion matrix and other statistics straight from the raw data. This matrix (table) displays measurements of the degrees in which consumers' actual use of payment methods match their preferred payment method for nonbill in-person payments. The restriction to in-person payments is needed because cash can only be paid in person (not online).

The remaining parts of the paper make use of two prediction algorithms: classification trees and random forest. In addition to consumers' preferred payment method, the prediction models incorporate other individual-specific and transaction-specific variables. The results show that consumers' preference remains the key predictor of actual payments made with credit and debit cards even after the additional individual-specific and transaction-specific variables are added to the prediction model. In contrast, predictions of cash payments are more influenced by whether cards are accepted than by the stated preference for paying cash.[1]

The study of consumer payment choice at the point of sale involves classifications of payment methods such as cash, credit cards, debit cards, prepaid cards, checks, direct account-to-account, and the use of mobile apps. Because these studies focus on *classifications* of payment methods, researchers have traditionally relied on discrete-choice analyses using binomial, multinomial, logit, and probit regression methods. These studies

[1]Note that preference for paying cash has been changing over time. Transition from cash to electronic payments are analyzed in Bolt et al. (2008), Van der Cruijsen et al. (2017), Jonker et al. (2018), Ardizzi et al. (2020), Kraenzlin et al. (2020), and Wisniewski et al. (2024).

measure mainly the marginal effects of transaction-specific and demographic variables on consumers' choice of payment methods.

In contrast, this paper focuses on measuring *prediction accuracy* and the relative importance of consumers' preferred payment method as a predictor of the actual use of the preferred method. For this reason, the predictions make use of classification trees and random forest techniques. The seminal paper by Breiman (2001) describes and compares the prediction performance of machine learning algorithms with that of discrete-choice regressions. The use of these techniques in economics and marketing is reviewed in Varian (2014), Kleinberg et al. (2015), Bajari et al. (2015), Mullainathan and Spiess (2017), and Athey and Imbens (2019).

This paper complements a related literature on merchant steering of consumer payment choice, see Barron et al. (1992), Bolt et al. (2010), Hayashi (2012), Briglevics and Shy (2014), Welte (2014), Stavins and Shy (2015), Welte (2016), Stavins (2018), and Greene et al. (2026). This literature examines whether merchants attempt to steer consumers away from paying with their preferred method towards payment methods that are less costly to the merchants. Relatedly, there is literature that analyzes the effects of credit card rewards (such as cashback and travel miles), see Ching and Hayashi (2010), Simon et al. (2010), Carbó-Valverde and Liñares-Zegarra (2011), Berkovich (2012), Arango et al. (2015), Agarwal et al. (2023), and Felt et al. (2023).

The above literature explains why steering is not observed very often and therefore cannot provide the full explanation for why consumers may deviate from their preferred payment method for some transactions. Therefore, unlike the above literature, the analysis in this paper focuses on providing measurements of the degrees to which a preference for a particular payment method can predict actual use of this method.

Relatedly, Wisniewski et al. (2024) analyze changes in consumer preferences to cashless payments following restrictions during the COVID-19 crisis period.

The paper is organized as follows: Section 2 describes the data. Section 3 analyzes consumer preferences for paying with cash, credit cards, debit cards, and others. Sec-

tion 4 compares straight from the raw data the actual choice of a payment method with consumers' preferred payment method for nonbill in-person payments. Section 5 constructs a decision tree that illustrates the importance of consumer preferences relative to other individual-specific and transaction-specific variables in predicting actual choice of a payment method. Section 6 analyzes payment choice predictions using random forests. Section 7 concludes. The data and R-code used in this research can be downloaded from https://github.com/ozshy/actual.

2. Data

The data used throughout this article are taken from the 2025 Survey and Diary of Consumer Payment Choice (SDCPC). The SDCPC is a representative sample of United States consumers age 18 and older taken from the University of Southern California's Understanding America Study.

The SDCPC asks respondents to record all of their transactions during three consecutive days in October. The transactions recorded in the SDCPC include purchases, bill payments, person-to-person payments, ATM withdrawals, deposits, and income receipts. On each assigned diary day, respondents report how much they paid, the means of payment that they used, and the type of merchant or person to whom the payment was made. Respondents' three-day periods were evenly distributed throughout each day in October.[2]

The October 2025 SDCPC sampled 6,079 respondents (individuals) who recorded 34,245 transactions. Some respondents or transactions were removed from the sample because of missing information. Each chart and table in this paper indicates the number of respondents and the number of payments. The data also provide information on to whom the payments were made. Each payment in the data contains a merchant (spending) category (out of 21 merchant categories) that indicates the type of spending.

The paper examines the relationship between respondents' preferred payment method

[2]The yearly data and summary reports are available for downloading from: https://www.atlantafed.org/research-and-data/surveys/survey-and-diary-of-consumer-payment-choice. The 2025 data and the R-code used in this paper are available at http://github.com/ozshy/actual.

and the payment methods that they actually used. Therefore, the full sample was reduced to include only nonbill in-person payments. This restriction is needed because cash can only be paid in person.[3] More precisely, starting from section 4, the analysis is conducted using a subsample containing 17,177 nonbill in-person payments made by 4,747 survey respondents corresponding to 10 spending/merchant categories. Appendix A provides descriptive statistics for each spending/merchant category and explains how the selection of the 10 spending/merchant categories was made.

3. Consumer preferences for payment methods

Survey respondents were asked to state their most preferred payment method (payment instrument) for making nonbill in-person payments.

3.1 Preferences by demographics

Figure 1 displays the fraction of respondents who prefer paying with cash, credit card, debit card, or other method according to 16 household income groups.

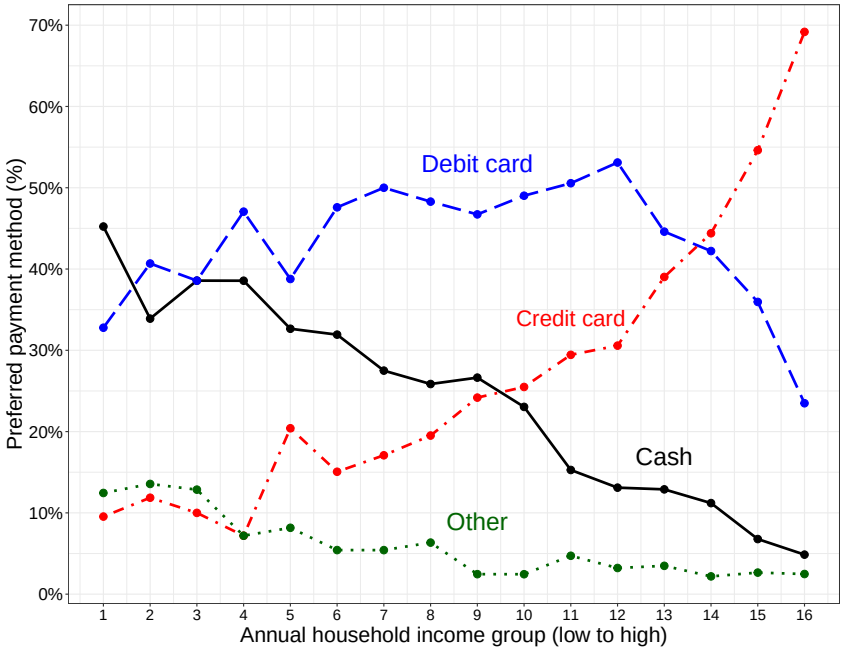
Figure 1 shows that consumer preference for paying cash declines with household income. In contrast, the percentage of respondents who prefer paying with credit cards rises with household income. The preference for paying with debit cards rises with household income until about $25,000, which reflects an increasing utilization of bank accounts. It maintains a plateau until $60,000 and then sharply declines with household income as respondents increase their preference for credit cards.

Payment methods labeled as "other" include paper checks, mobile apps, and direct bank transfers. They will not be analyzed in this paper because of their low use for nonbill payments that are made in person. It should also be noted that most payments that are made with mobile apps rely on debit and credit cards that are captured in this analysis.

Figure 2 displays preferences for payment methods by age group. Figure 2 shows that the preference for cash payments remains almost steady with age (around 15%). The

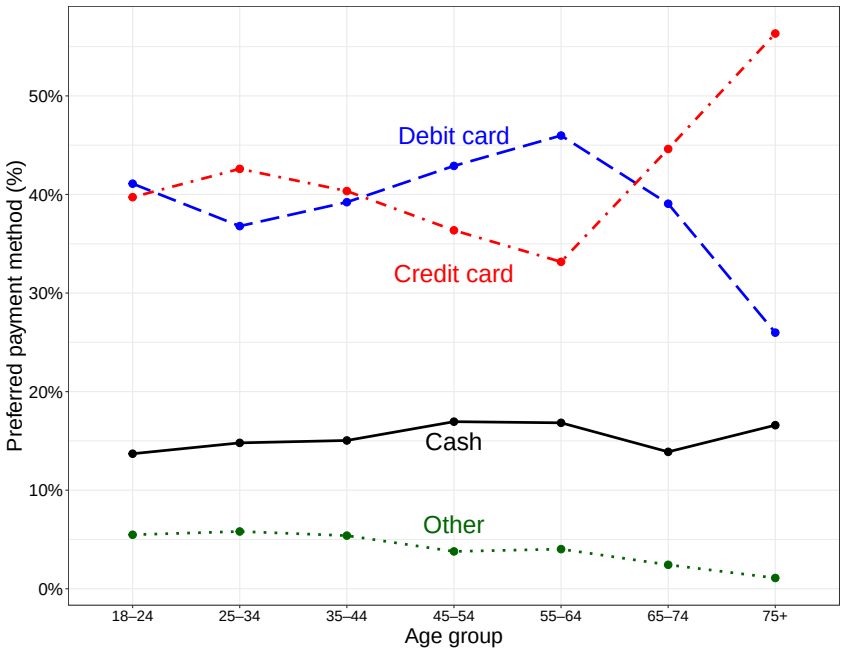
[3]Bill payments are also excluded because most bill payments are made with direct withdrawals from a bank account or online bill payments. These methods are not used for payments made in person.

Figure 1: Preferred payment method for nonbill in-person payments by household income group.



Source: Author's calculation from the 2025 Survey and Diary of Consumer Payment Choice.
Notes: Household annual income groups: 1) Less than $5,000; 2) $5,000–$7,499; 3) $7,500–$9,999; 4) $10,000–$12,499; 5) $12,500–$14,999; 6) $15,000–$19,999; 7) $20,000–$24,999; 8) $25,000–$29,999; 9) $30,000–$34,999; 10) $35,000–$39,999; 11) $40,000–$49,999; 12) $50,000–$59,999; 13) $60,000–$74,999; 14) $75,000–$99,999; 15) $100,000–$149,999; 16) $150,000 or higher. The figure displays preferences of 5,889 survey respondents.

Figure 2: Preferred payment method for nonbill in-person payments by age group



Source: Author's calculation from the 2025 Survey and Diary of Consumer Payment Choice.
Notes: The figure displays preferences of 5,879 survey respondents.

preferences for credit cards and debit cards seem to move in opposite directions which reflect substitution between the two types of cards. Preference for credit cards slightly rises whereas preference for debit card slightly falls between age 18 and 24. Beyond age 25, preference for credit cards slightly declines whereas preference for debit cards slightly increases until age 64. The share of preferences for each card type are steady (around 40%) until age 64. Beyond age 65, preference for credit card payments sharply rises whereas preference for debit card payments sharply declines with age.

3.2 Preferences for card payments, card ownerships, and card rewards

Survey respondents report whether they own or do not own credit and debit cards. That raises the question whether card ownership influences the preference for making card payments, the other way around, or not at all. Table 1 displays the number of respondents who own credit and debit cards and whether the cards are their preferred payment

method for making nonbill in-person payments.

Table 1: Card preference versus card ownership and card rewards

	Credit card		Debit card	
Card ownership	Prefer	Not prefer	Prefer	Not prefer
Total	2411	3574	2387	3598
Own card	2396	2473	2349	3038
Do not own card	15	1101	38	560
Own card with reward	2340	1964		
Own card without reward	56	509		

Source: Author's calculation from the 2025 Survey and Diary of Consumer Payment Choice.
Notes: The table captures 5,985 survey respondents. Prefer and not prefer apply to whether the card is the respondents' preferred payment method for nonbill in-person payments.

The main findings from Table 1 are as follows:

Findings 1. *(a) Card ownership does not imply that the card is the preferred payment method. Out of 4,869 credit card owners, 2,396 (49%) prefer using credit cards for nonbill in-person payments, while 2,473 (51%) prefer other methods.*

(b) Similarly, out of 5,387 debit card owners, 2,349 (44%) prefer using debit cards for nonbill in-person payments, while 3,038 (56%) prefer other methods.

(c) Preference for a card is correlated with higher card ownership. Out of 2,411 respondents who prefer paying with credit cards, 2,396 (99%) own credit cards while only 15 (1%) do not own credit cards.

(d) Similarly, out of 2,387 respondents who prefer paying with debit cards, 2,349 (98%) own debit cards while only 38 (2%) do not own debit cards.

(e) Out of the 2,396 credit card owners who prefer credit cards, 2,340 (98%) have reward credit cards. However, out of the 2,473 credit card owners who prefer to pay with other means, only 1,964 (79%) respondents receive credit card rewards.

The main implication of findings 1 is that there are no implied strong causalities between card ownership and card preferences. That is, card ownership does not imply preference

for card payments. The opposite statement is less clear because very few respondents who prefer paying with a card do not own cards. With respect to credit card rewards among credit card owners, respondents who prefer to pay with cards have relatively more reward cards than respondents who do not prefer paying with credit cards.

In the absence of strong causalities, the prediction models to be analyzed in sections 5 and section 6 will make use of all the ownership and reward variables as predictors in addition to preferences that serve as the main predictors. However, these predictors will be found to be less important than the stated preferences for predicting card payments, see figure 3 in section 5 and table 3 in section 6.

4. Can preference predict actual choice?

This section investigates the degree to which consumer preference for a payment method can predict actual use of the payment method just by looking at the raw data. Section 5 and section 6 then compare the results from this section to the predictions obtained from two prediction algorithms that incorporate additional individual-specific and transaction-specific predicting variables.

The four definitions in this section slightly modify the standard terminology that is used for measuring prediction accuracy. This terminology reflects how the choice of a payment method (actual) corresponds to individuals' preferred payment method (predictor). In table 2, columns correspond to actually used payment methods. Rows correspond to respondents' preferred payment method that serves as a predictor of an actual use.

DEFINITION 1. *The **True Positives** (TP) associated with payment method X are the number of payments made with X by respondents who prefer paying with X.*

TP payments are displayed on the first main diagonal in table 2. For example, in the sample 5,772 credit card payments were made by respondents who prefer paying with credit cards. A more informative measure is the positive prediction value (PPV) which is expressed in terms of percentages and is formally defined in (2). The PPV are shown

Table 2: Confusion matrix: Actual versus predicted use of payment methods

Preferred payment	Actually-used payment method				
method (predictor)	Cash	Credit	Debit	Other	Total Preferred
Cash	1081	291	465	141	1978
Credit	1168	5772	437	567	7944
Debit	1193	614	4304	447	6558
Other	173	186	219	119	697
Cash	55%	15%	24%	7%	101%
Credit	15%	73%	6%	7%	101%
Debit	18%	9%	66%	7%	100%
Other	25%	27%	31%	17%	100%
True Positives (TP)	1081	5772	4304	119	
False Negatives (FN)	2534	1091	1121	1155	
False Positives (FP)	897	2172	2254	578	
True Negatives (TN)	12665	8142	9498	15325	
Sensitivity (SEN)	0.30	0.84	0.79	0.09	
Specificity (SPE)	0.93	0.79	0.81	0.96	
Pos Pred Value (PPV)	0.55	0.73	0.66	0.17	
Neg Pred Value (NPV)	0.83	0.88	0.89	0.93	
Balanced Accuracy (BA)	0.62	0.82	0.80	0.53	

Source: Author's calculation from the 2025 Survey and Diary of Consumer Payment Choice.
Notes: The table captures 4,747 respondents who made 17,177 nonbill in-person payments. Some parts of this table were computed with the *caret* R-package.

on the second main diagonal and also in the row labeled as "Pos Pred Value" in table 2. The table shows that 73% of payments by respondents who prefer credit cards were made with credit cards. In contrast, only 55% of the payments made by respondents who prefer paying cash were made with cash. Similarly, 66% of the payments made by respondents who prefer paying with debit cards were made with debit cards.

DEFINITION 2. *The **False Negatives** (FN) associated with payment method X are the number of payments made with X by respondents who prefer paying with Y (for all payment methods $Y \neq X$).*

For example, table 2 shows that $1,168 + 1,193 + 173 = 2,534$ cash payments were made by respondents who prefer not paying with cash. Similarly, $291 + 614 + 186 = 1,091$ credit card payments were made by respondents who prefer not paying with credit cards.

DEFINITION 3. *The **False Positives** (FP) associated with payment method X are the number of payments made by respondents who prefer X but paid with Y (for all payment methods $Y \neq X$).*

For example, $291 + 465 + 141 = 897$ noncash payments were by respondents who prefer cash. Similarly, $1,168 + 437 + 567 = 2,172$ payments were made by respondents who prefer credit cards but paid with other methods.

DEFINITION 4. *The **True Negatives** (TN) associated with a specific payment method X are the number of payments made with Y ($Y \neq X$) by respondents who do NOT prefer X.*

True negatives can be computed from table 2 by summing up all payments that do not appear in a row and not in a column of a specific payment method. Note that the sum of all TP, FN, FP, and TN payments equals the total number of payments in the sample.

The bottom part of table 2 displays five commonly-used measures of prediction accuracy that are defined as follows:

$$\text{SEN} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad \text{SPE} = \frac{\text{TN}}{\text{TN} + \text{FP}}, \quad \text{BA} = \frac{\text{SEN} + \text{SPE}}{2}. \tag{1}$$

Intuitively, high sensitivity (SEN) means that there were very few false negatives. High specificity (SPE) means that there were very few false positives. Balanced accuracy (BA) is the average of the two measures. Also,

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad \text{NPV} = \frac{\text{TN}}{\text{TN} + \text{FN}}, \tag{2}$$

Intuitively, positive prediction value (PPV) measures the degree that a preferred payment method was actually used. In contrast, negative prediction value (NPV) measures the degree that a nonpreferred payment method was not used. Table 2 shows that paying with credit cards results in the highest positive prediction value 0.73 and the highest balanced accuracy 0.82. Debit cards scored 0.66 and 0.80, respectively. Preference for cash scored only 0.55 and 0.62, respectively. These numbers are summarized as follows:

Findings 2. *(a) Stated preference for paying with credit cards predicts actual use of credit cards better than the preferences for debit cards and cash predict actual use of debit card and cash payments, respectively.*

(b) Stated preference for paying with debit cards predicts the use of debit cards better than the preference for paying cash predicts actual cash payments.

5. Individual-specific and transaction-specific effects: Classification trees

Using the raw data only, table 2 shows how consumers' stated preference for a payment method can serve as an imperfect predictor for the payment method that consumers actually use in making nonbill in-person payments. The predictions of cash payments are relatively poor and yield only 55% positive prediction value compared with 73% and 66% positive prediction values of credit and debit card payments, respectively. This is because, consumers who prefer paying cash tend to pay cash for selected spending categories, small transaction values, and when cards are not accepted. Cash payments are analyzed in Shy (2023) and O'Brien and Phelps (2026).

The remainder of this paper investigates the degree to which consumers' preferred payment method continues to serve as the main predictor of the actually used payment

method even after the incorporation of several individual-specific and transaction-specific features. This section constructs a machine learning classification tree to illustrate how consumers' stated preference is ranked relative to other variables for predicting actual use of payment methods. Consider the following classification tree model:[4]

$$\overbrace{\text{Payment method}_{i,t}}^{\text{actually used}} \sim \overbrace{\text{Preference}_i + \text{Own credit}_i + \text{Credit reward}_i + \text{Own debit}_i}^{\text{individual specific}} + \underbrace{\text{Cash accepted}_{i,t} + \text{Credit card surcharge}_{i,t} + \text{Discount for payment method}_{i,t}}_{\text{transaction specific}} + \underbrace{\text{Amount (\$)}_{i,t} + \text{Merchant}_{i,t} + \text{Card accepted}_{i,t}}_{\text{transaction specific}}, \quad (3)$$

where subscript t denotes a particular transaction (payment) and i denotes a particular individual (respondent) in the sample.

The individual-specific predictors in the model (3) are: "Preference" (categorical: Cash, Credit, Debit, and Other); "Credit own" (binary) indicates whether the respondents owns a credit card; "Credit reward" indicates whether the credit card used most often pays a reward (such as points, miles, or cash back); "Debit own" indicates whether the respondent owns a debit card.

The transaction-specific variables in the model (3) are: "Amount" (numerical) is the payment dollar value; "Merchant" (categorical) is the payee according to 10 spending/merchant categories that are described in Appendix A; "Card accepted" and "cash accepted" (binary variables) indicate whether the payee was willing to accept these payment methods; "Credit card surcharge" (binary) indicates whether the payee added a surcharge for this credit card payment; "Discount" (binary) indicates whether the respondent received a discount for the payment method that was used in this transaction. Appendix B describes the complexity involved in collecting data on the above binary predictors. It also explains

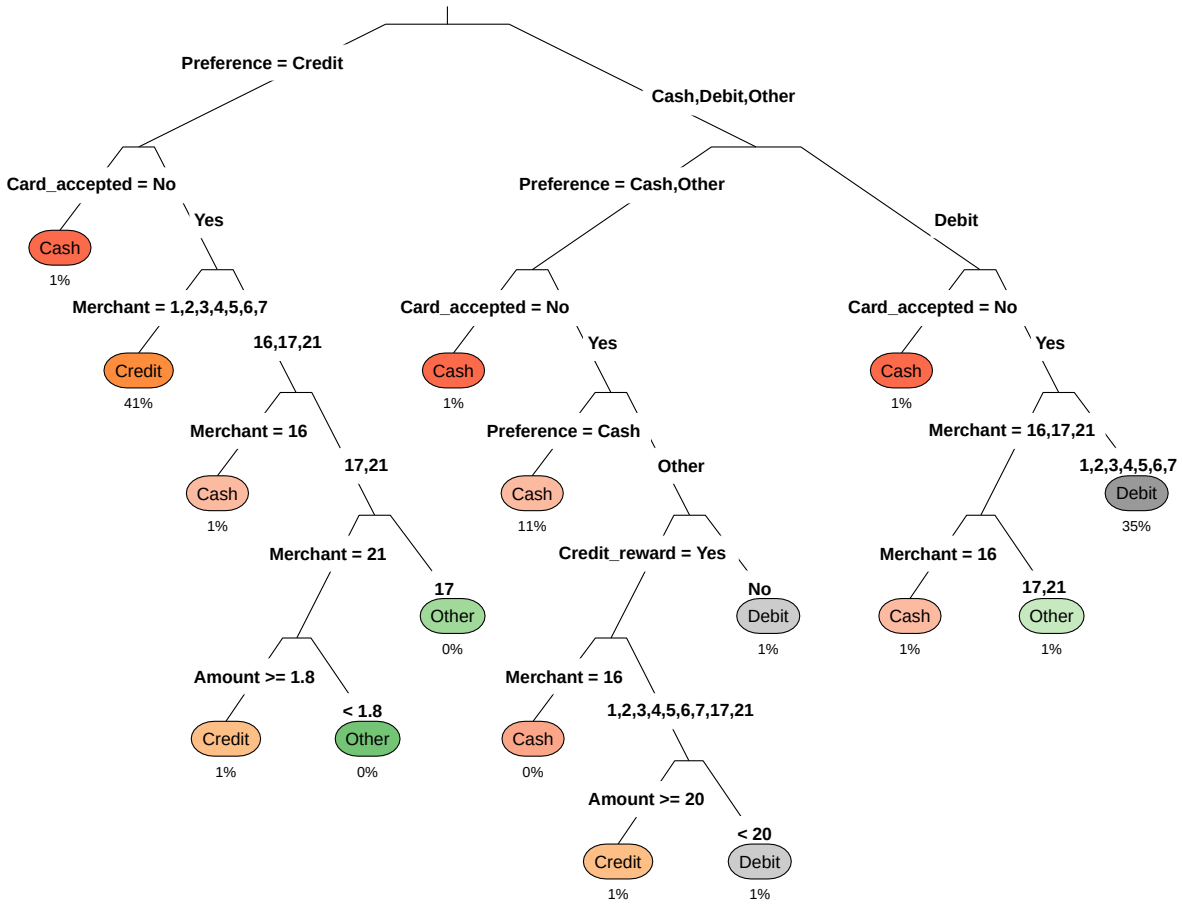
[4]Note that there is no need to control for respondents' demographic variables (such as income, age, gender, etc.) because those are already embedded into respondents' stated preference for a payment method. For example, figure 1 and figure 2 show how preferences vary with respondents' household income and age which are the two key demographic variables.

the method used to impute the missing data (NAs).

In the context of machine learning, a classification tree displays an optimized algorithm in the form of an upside-down tree. The tree illustrates how the machine (software) splits and classifies the data according to the objective of minimizing (a function of) the number of payment method classification errors according to the ten predictors listed in the model (3).

Figure 3 illustrates a decision tree. The variables/predictors appearing on higher

Figure 3: Data-driven decision tree on payment method choice



Source: Author's calculation from the 2025 Survey and Diary of Consumer Payment Choice.
Notes: The decision tree is based on the model (3). The sample contains 4,747 respondents who made 17,177 nonbill in-person payments. The percentage (%) at the bottom of each terminal node is the percentage of payments that reaches that node. Merchant categories are listed in Appendix A. The tree was constructed using the *rpart*, *rpart.plot*, and *partykit* R-packages.

branches are more important than the predictors listed on lower branches. The top branch on the left in figure 3 shows that preference for paying with credit cards serves as the main predictor for the actual use of credit cards. This branch illustrates the prediction in which consumers deviate from credit cards (the preferred method) when cards are not accepted or when they pay for spending/merchant categories 16 (person-to-person), 17 (charitable and religious donations), and 21 (public transport).

The top branch on the right in Figure 3 illustrates how preferences for cash and debit cards serve as the main predictors of cash and debit card payments, respectively. Starting with the third branch on the right and all the way down, the tree shows that consumers who prefer debit cards are predicted to deviate and pay cash if cards are not accepted or for making person-to-person payments (category 16).

The bottom split on the left and the bottom split in the middle predict minimum dollar amounts for credit card payments ($1.8 for public transport payments and $20 for a wide variety of merchant categories, respectively).

The findings from the classification tree algorithm are summarized as follows:

Findings 3. *(a) The decision tree algorithm classifies consumers' preferred payment method as the top predictor.*

(b) Card acceptance and spending/merchant category are ranked second and third, respectively.

(c) Credit card ownership, credit card rewards, and debit card ownership have smaller effects (if any) in predicting the actual choice of a payment method.

Finding 3(c) implies that credit card ownership, credit card reward, and debit card ownership predictors can be dropped from the classification model (3) with little or no effect on prediction accuracy. However, this finding highlights the limitation of the classification tree algorithm because these variables can potentially improve prediction accuracy, as shown in the next section.

6. Individual-specific and transaction-specific effects: Random forests

The decision tree displayed in figure 3 provides a visual method for displaying how a decision tree algorithm predicts actual choice of a payment method based on preferences and several individual-specific and transaction-specific variables. However, for the purpose of making predictions, a related algorithm called "random forest" can further enhance prediction accuracy. The reason is that random forest predictions are made by averaging the predictions made from hundreds of trees like the tree displayed in figure 3.

The random forest algorithm constructs a large number of classification trees where each tree splits the data according to a randomly-selected subgroup of predictors. The number of predictors in each subgroup is controlled by the researcher. The number of excluded predictors is generally chosen to be near the square root of the total number of predictors. Hence, because the model (3) has ten predictors and $\lfloor \sqrt{10} \rfloor = 3$, random forest removes randomly three predictors from each tree. This procedure decorrelates the sampled trees so that any strong predictor (such as the preferred payment method) is not chosen as the top split in each tree.

The analysis of random forest proceeds as follows: Subsection 6.1 demonstrates how this procedure ranks the relative importance of the ten predictors. Subsection 6.2 uses random forest to make predictions on the actual use of payment methods and computes several measures of prediction accuracy.

6.1 Random forest: Variable importance table

Although random forest algorithms do not draw trees, they provide a useful procedure for variable selection (in addition to improved prediction accuracy). Table 3 shows how the random forest algorithm ranks the relative importance of the ten predictors in the model (3).

In table 3, the mean decrease accuracy (MDA) for each payment method (cash, credit, debit, and other) measures by how much prediction accuracy is reduced with the elimi-

Table 3: Random forest: Predictors' degree of importance

	MDA for payment method used				Mean decrease
Predicting variable	Cash	Credit	Debit	Other	accuracy (MDA)
Preferred payment method	64.11	135.18	146.68	-1.64	150.76
Spending/merchant category	50.55	67.25	63.47	62.18	96.95
Spending amount	38.42	28.53	26.51	36.13	48.20
Own credit card	9.41	13.94	-6.90	13.96	16.68
Credit card reward	6.74	26.53	8.02	10.41	29.44
Own debit card	0.70	-8.72	49.52	6.53	48.76
Card accepted	103.77	109.35	95.15	27.31	136.47
Cash accepted	54.61	6.43	8.50	6.79	48.01
Credit card surcharge	23.97	41.95	46.91	9.83	52.63
Discount for payment method	27.81	29.68	36.71	15.94	48.71

Source: Author's calculation from the 2025 Survey and Diary of Consumer Payment Choice.
Notes: Results from a random forest model (3) based on 4,747 respondents who made 17,177 non-bill in-person payments. The table was computed using the *randomforest* R-package.

nation (permutation) of each of the ten predictors separately.[5] In simple words, the MDA measures how much worse the model performs when this variable is scrambled. The few small negative MDAs indicate that adding the predictor reduces prediction accuracy of the payment method instead of improving it.

The first row in table 3 shows that consumers' stated preferred payment method is the most important predictor of credit and debit payments. The elimination of this predictor would lead to the largest drop in prediction accuracy. For predicting card payments, card acceptance ranks second. Spending/merchant category ranks third. Credit card surcharge is also an influential predictor because consumers are likely to switch from credit to debit cards or to paying cash if credit card payments are surcharged.

For predicting cash payments, card acceptance ranks first because consumers are forced

[5] $MDA = \frac{1}{T}\sum_{t=1}^{T}(err_{t,perm} - err_t)$, where t indexes a classification tree and "perm" denotes a permutation (elimination) of a predictor. $err = 1$ denotes a prediction error of the payment method used and $err = 0$ denotes an accurate prediction. Each tree in a random forest algorithm is trained on a bootstrap sample of the training data. Each error is then computed on an out-of-bag (OOB) subsample of the training sample that was not part of the subsample used for the construction of the classification tree. The average is then computed based on all the T trees used in the procedure.

to pay cash when cards are not accepted. Preference for paying cash comes second, followed by cash acceptance, merchant/spending category, and payment dollar amount.

Finally, the column on the right in table 3 displays the overall MDAs across the utilization of the four payment methods. All these findings are summarized as follows:

Findings 4. *(a) Consumers' stated preference for credit and debit card payments is the most important predictor for nonbill in-person card payments. Card acceptance ranks second.*

(b) Spending/merchant category and credit card surcharge are ranked third and fourth for predicting credit and debit card payments.

(c) Discount for using a payment method is also an influential predictor of cash, credit, and debit card payments. Credit card rewards is also an influential predictor of credit card payments.

(d) For predicting cash payments, card acceptance ranks first and preference for paying cash ranks second. Cash acceptance and spending/merchant category are ranked next followed by spending dollar amount.

6.2 Random forest: Prediction accuracy

The random forest classification model is now used to predict consumers' choice of a payment method. Whereas table 2 displays comparisons between the preferred payment methods and the actually used payment methods based only on the raw data, table 4 displays the same comparisons but with the use of random forest predictions based on the model's ten predicting variables, see equation (3).

Table 4 was constructed in three steps: First, the full sample of all nonbill in-person payments across the 10 spending/merchant categories (described in Appendix A) was split into a training subsample and a testing subsample. The two subsamples consisted of different survey respondents. The full sample consisted of 17,177 payments made by 4,747 respondents. The training subsample consisted of 3,322 unique respondents (70%) and the testing data 1,425 unique respondents (30%). The resulting training subsample consisted of 12,023 payments and the testing subsample 5,154 payments. Second, based on the model (3), the random forest prediction algorithm was constructed using the train-

Table 4: Random forest: Actual versus predicted use of payment methods

Preferred payment	Actually-used payment method				
method (predictor)	Cash	Credit	Debit	Other	Total
Cash	504	58	117	70	749
Credit	236	1793	163	140	2332
Debit	300	190	1358	133	1981
Other	18	9	9	56	92
Cash	67%	8%	16%	9%	100%
Credit	10%	77%	7%	6%	100%
Debit	15%	10%	69%	7%	101%
Other	20%	10%	10%	61%	101%
True Positives (TP)	504	1793	1358	56	
False Negatives (FN)	554	257	289	343	
False Positives (FP)	245	539	623	36	
True Negatives (TN)	3851	2565	2884	4719	
Sensitivity (SEN)	0.48	0.87	0.82	0.14	
Specificity (SPE)	0.94	0.83	0.82	0.99	
Pos Pred Value (PPV)	0.67	0.77	0.69	0.61	
Neg Pred Value (NPV)	0.87	0.91	0.91	0.93	
Balanced Accuracy (BA)	0.71	0.85	0.82	0.57	

Source: Author's calculation from the 2025 Survey and Diary of Consumer Payment Choice.
Notes: Predictions are based on a testing subsample consisting of 5,154 payments (30% of the full sample of 17,177 payments). The table was computed using the *randomforest* and *caret* R-packages.

ing subsample only. Third, the resulting prediction algorithm was applied to the testing subsample. This three-stage procedure provides the ultimate test for the algorithm's prediction accuracy because all the prediction accuracy statistics displayed in table 4 have been generated on a completely new dataset (the testing subsample).

Comparing table 4 with table 2 reveals that the positive prediction value (PPV) for payments made in cash has improved from 0.55 to 0.67. The PPV for credit card payments has improved from 0.73 to 0.77, and from 0.66 to 0.69 for debit cards. The balanced accuracy increased from 0.62 to 0.71 for cash, from 0.82 to 0.85 for credit cards, and from 0.80 to 0.82 for debit cards. These findings are summarized as follows:

Findings 5. *(a) Random forest analysis based on consumer preference for a payment method and ten individual-specific and transaction-specific variables improve prediction accuracy. However,*

(b) For credit and debit card payments, the improvements are not sufficiently large to invalidate the significance of consumer preference as the main predictor of credit and debit card payments.

In other words, a comparison of the confusion matrices in table 2 and table 4 reveals that consumers' stated preference for a payment method does a fairly good job in predicting actual use of card payments even without making use the of ten individual-specific and transaction-specific variables in the model (3) and without making use of powerful prediction algorithms such as random forests.

7. Summary and conclusion

Using a survey on consumer payments choice, the paper analyzes the relationship between consumers' preferred payment method and the actual use of this payment method for nonbill in-person payments. Demographic predictors, such as income, age, and gender, are not used because these are already embedded in respondents' preferences (see figure 1 and figure 2). However, several other individual-specific and transaction-specific predictors are used for the purpose of measuring the degree to which prediction accuracy could be improved.

There are several reasons for why stated preferences cannot always predict actual use of payment methods in general, and use of cash in particular. One reason is that consumers' stated preference may change when asked in different contexts.[6] Ideally, consumers should be asked to state their preference for a payment method for each spending/merchant category separately. Consumers could also be asked to state their preferred payment method separately for each given dollar payment amount. However, researchers have found that asking too many survey questions about payments leads to fatigue that jeopardizes accuracy, see Jonker and Kosse (2013). Therefore, researchers may be forced to limit the scope of survey questions and focus only on respondents' general preferences.

Appendix A Spending/merchant categories

Survey respondents who participate in the SDCPC also record to whom each payment is made. Therefore, each payment in the dataset contains a merchant (spending) category indicating the type of spending. Survey respondents select one merchant category out of 21 merchant categories. However, starting from section 4, the analysis is restricted to nonbill in-person payments (the vast majority of the data). For this reason, the analysis focuses only on 10 spending/merchant categories because respondents recorded more than 250 nonbill in-person payments in these categories but not in others. Table A.1 lists and describes the 10 spending/merchant categories and also provides some statistics on the payments made in each category.

Appendix B Handling missing transaction-specific observations

The prediction model in equation (3) lists six transaction-specific predicting variables. "Amount" (numeric) is the payment dollar amount. "Merchant" (categorical) corresponds to the 10 merchant/spending categories discussed in Appendix A. However, the remain-

[6] On the other hand, some research has found "inertia" effects where consumers tend to repeatedly use the same payment method for several purchases, see Shy (2021).

Table A.1: Spending/merchant categories

	Number	Payment value (US dollars)			
Category	Payments	Median	Average	Min	Max
1. Grocery store	5107	25.00	46.03	0.21	773.32
2. Gas station	2037	25.63	29.64	0.79	513.78
3. Restaurant/bar	1690	35.00	48.52	1.00	859.34
4. Fast food/coffee shop	3405	11.75	16.05	0.60	205.94
5. General merchandise store	2703	29.33	86.17	0.48	8199.13
6. General service	604	36.00	112.98	0.25	9861.48
7. Art/entertainment	373	22.56	56.46	1.00	2748.00
16. Person-to-person	585	26.54	112.25	0.50	10000.00
17. Charitable/religious donation	262	20.00	93.14	1.00	2300.00
21. Public transport/tolls	411	2.90	5.47	0.20	150.00
Total	17177	49.29	21.74	0.20	10000.00

Source: The 2025 Survey and Diary of Consumer Payment Choice.
Notes: Index numbers correspond to how merchants are coded in the data. More details on each merchant category are available from both the survey questionnaire and codebook, https://www.atlantafed.org/research-and-data/surveys/survey-and-diary-of-consumer-payment-choice.

ing four transaction-specific (binary) variables consist of respondents' replies to questions that were asked only for specific payment methods, but not all payment methods. Therefore, the missing observations of these four predictors were imputed over all the nonbill in-person payments using the *missForest* R-package according to:

$$\overbrace{\text{Imputed variable}_t}^{\text{missing observation}} \sim \overbrace{\text{Payment method}_t + \text{Amount (\$)}_t + \text{Merchant}_t}^{\text{payment/transaction characteristics}}, \tag{B.1}$$

for every transaction t (payment) in the relevant sample. The four imputed variables (card acceptance, cash acceptance, credit card surcharge, and discount) are discussed below.[7]

Card accepted: The survey question "whether a credit or debit card would have been accepted for this transaction" was applied only to respondents who paid cash. However, we know that cards were accepted in all payments that were made with credit and debit cards. Therefore, all the 6,863 and 5,424 NAs (missing observations) for credit and debit card payments, respectively, were reclassified as "Yes" instead of NA. In addition, "I don't know" replies were reclassified as NA. That left 14,875 (96.9%) payments that were classified as "Yes" and 466 (3.1%) as "No" before the imputations. The imputations of the 1,836 remaining missing observations over all payments yielded 16,625 (96.8%) payments with "Yes" and 552 (3.2%) "No."

Cash accepted: The question "whether cash would have been accepted for this transaction" was applied only to card payments. However, we know that cash was accepted for payments that were made with cash. Therefore, all 3,613 cash payments with NA were reclassified as "Yes" instead of NA. In addition, all "I don't know" replies were classified as NA. "I'm not sure, but I think so" replies were reclassified as "Yes." "I'm not sure, but I do not think so" replies were classified as "No." That left 15,896 (93%) cash payments that were classified as "Yes" and 1,181 (7%) as "No" before imputations. The imputations

[7]For research on merchants' decisions whether to accept cards and whether to surcharge or provide discounts see Jonker (2011), Bounie et al. (2017), and Welte and Wu (2023).

of the 100 missing observations over all payments yielded 15,995 (93.1%) payments with "Yes" and 1,182 (6.9%) "No."

Credit card surcharge: Respondents who paid with credit cards were asked "did you pay an extra charge, surcharge, or convenience fee to the merchant specifically for using this credit card?" 206 (3%) answered "Yes" and 6,653 (97%) answered "No." Respondents were not asked this question if they did not pay with a credit card. That left in 10,318 NAs. The imputations of these NAs over all payments yielded 206 (1.2%) "Yes" and 16,971 (98.8%) "No."

Discount: Respondents were asked "whether a discount was received for using the chosen payment instrument." 495 (3.1%) payments were reported "Yes" whereas 15,396 (96.9%) reported "No." The imputations of the 1,286 NAs yielded 495 (2.9%) "Yes" and 16,682 (97.1%) "No."

Finally, the variable "Credit reward" in the model (3) was recorded on the individuals dataset (not on the transactions dataset). Note that the two datasets were merged in the R-code according to respondents' unique ID.

Credit reward: Respondents were asked "Does your credit card give rewards?" The merged dataset recorded 1,851 payments that were made by respondents who did not answer this question (1,851 NAs). Clearly, respondents who do not own credit cards do not have reward cards. Therefore, all NAs for non-credit card payments that were made by respondents who do not own credit cards were replaced with a "No." The remaining 29 NAs payments were replaced with a "Yes" number because most credit cards earn rewards.

References

Agarwal, Sumit, Andrea Presbitero, André Silva, and Carlo Wix, "Who pays for your rewards? Redistribution in the credit card market," 2023. Board of Governors of the Federal Reserve System, Finance and Economics Discussion Series 2023–007.

Arango, Carlos, Kim Huynh, and Leonard Sabetti, "Consumer payment choice: Merchant card acceptance versus pricing incentives," *Journal of Banking and Finance*, 2015, *55*, 130–141.

Ardizzi, Guerino, Andrea Nobili, and Giorgia Rocco, "A Game Changer in Payment Habits: Evidence from Daily Data During a Pandemic," 2020. Bank of Italy, Economic Research and International Relations Area.

Athey, Susan and Guido Imbens, "Machine learning methods that economists should know about," *Annual Review of Economics*, 2019, *11* (1), 685–725.

Bajari, Patrick, Denis Nekipelov, Stephen Ryan, and Miaoyu Yang, "Machine learning methods for demand estimation," *American Economic Review*, 2015, *105* (5), 481–485.

Barron, John, Michael Staten, and John Umbeck, "Discounts for cash in retail gasoline marketing," *Contemporary Economic Policy*, 1992, *10* (4), 89–102.

Berkovich, Efraim, "Card rewards and cross-subsidization in the gasoline and grocery markets.," *Review of Network Economics*, 2012, *11* (4). Article 4.

Bolt, Wilko, David Humphrey, and Roland Uittenbogaard, "Transaction pricing and the adoption of electronic payments: A cross-country comparison," *International Journal of Central Banking*, 2008, *4* (1), 89–123.

_ , Nicole Jonker, and Corry Van Renselaar, "Incentives at the counter: An empirical analysis of surcharging card payments and payment behaviour in the Netherlands," *Journal of Banking and Finance*, 2010, *34* (8), 1738–1744.

Bounie, David, Abel François, and Leo Van Hove, "Consumer payment preferences, network externalities, and merchant card acceptance: An empirical investigation," *Review of Industrial Organization*, 2017, *51* (3), 257–290.

Breiman, Leo, "Statistical modeling: The two cultures," *Statistical Science*, 2001, *16* (3), 199–231.

Briglevics, Tamás and Oz Shy, "Why don't most merchants use price discounts to steer consumer payment choice?," *Review of Industrial Organization*, 2014, *44* (4), 367–392.

Carbó-Valverde, Santiago and José M Liñares-Zegarra, "How effective are rewards programs in promoting payment card usage? Empirical evidence," *Journal of Banking and Finance*, 2011, *35* (12), 3275–3291.

Ching, Andrew and Fumiko Hayashi, "Payment card rewards programs and consumer payment choice," *Journal of Banking and Finance*, 2010, *34* (8), 1773–1787.

der Cruijsen, Carin Van, Lola Hernández, and Nicole Jonker, "In love with the debit card but still married to cash," *Applied Economics*, 2017, *49* (30), 2989–3004.

Felt, Marie-Hélène, Fumiko Hayashi, Joanna Stavins, and Angelika Welte, "Regressive effects of payment card pricing and merchant cost pass-through in the United States and Canada," *Journal of Banking and Finance*, 2023, *154*, 106968.

Greene, Claire, Oz Shy, and Joanna Stavins, "Merchant steering of consumer payment choice," 2026. Federal Reserve Bank of Atlanta, Working Paper 2026–2.

Hayashi, Fumiko, "Discounts and surcharges: Implications for consumer payment choice," 2012. Federal Reserve Bank of Kansas City, Payments System Research Briefing, June.

Jonker, Nicole, "Card acceptance and surcharging: the role of costs and competition," *Review of Network Economics*, 2011, *10* (2), 1–35.

_ **and Anneke Kosse**, "Estimating cash usage: The impact of survey design on research outcomes," *De Economist*, 2013, *161* (1), 19–44.

_ **, Lola Hernández, Renate de Vree, and Patricia Zwaan**, "From cash to cards: How debit card payments overtook cash in the Netherlands," 2018. Netherlands Central Bank, Occasional Studies No. 1.

Kleinberg, Jon, Jens Ludwig, Sendhil Mullainathan, and Ziad Obermeyer, "Prediction policy problems," *American Economic Review*, 2015, *105* (5), 491–495.

Kraenzlin, Sébastien, Christoph Meyer, and Thomas Nellen, "COVID-19 and Regional Shifts in Swiss Retail Payments," *Swiss Journal of Economics and Statistics*, 2020, *156* (1), 1–20.

Mullainathan, Sendhil and Jann Spiess, "Machine learning: An applied econometric approach," *Journal of Economic Perspectives*, 2017, *31* (2), 87–106.

O'Brien, Shaun and Hailey Phelps, "2026 findings from the Diary of Consumer Payment Choice," 2026. Federal Reserve Financial Services.

Shy, Oz, "Consumer use of multiple payment methods," *Review of Industrial Organization*, 2021, *58* (3), 339–355.

_ , "Cash is alive: How economists explain holding and use of cash," *Journal of Economic Literature*, 2023, *61* (4), 1465–1520.

Simon, John, Kylie Smith, and Tim West, "Price incentives and consumer payment behaviour," *Journal of Banking and Finance*, 2010, *34* (8), 1759–1772.

Stavins, Joanna, "Consumer preferences for payment methods: Role of discounts and surcharges," *Journal of Banking and Finance*, 2018, *94*, 35–53.

_ **and Oz Shy**, "Merchant steering of consumer payment choice: Evidence from a 2012 diary survey," *Journal of Behavioral and Experimental Economics*, 2015, *55*, 1–9.

Varian, Hal, "Big data: New tricks for econometrics," *Journal of Economic Perspectives*, 2014, *28* (2), 3–27.

Welte, Angelika, "Merchant payment steering: Evidence from the 2009 Bank of Canada Methods of Payment survey." PhD dissertation, Carleton University 2014.

_ , "Wait a minute: the efficacy of discounting versus nonpecuniary payment steering," *Journal of Financial Market Infrastructures*, 2016, *4* (4), 17–25.

_ **and Joy Wu**, "The 2021–22 merchant acceptance survey pilot study," 2023. Bank of Canada, Staff Discussion Paper 2023–1.

Wisniewski, Tomasz, Michal Polasik, Radoslaw Kotkowski, and Andre Moro, "Switching from cash to cashless payments during the COVID-19 pandemic and beyond," *International Journal of Central Banking*, 2024, *20* (3), 303–371.