How Do Individuals Repay Their Debt? The Balance-Matching Heuristic*

John Gathergood† Neale Mahoney‡ Neil Stewart§ Jörg Weber¶

September 14, 2017

Abstract

We study how individuals repay their debt using linked data on multiple credit cards from five major issuers. We find that individuals do not allocate repayments to the higher interest rate card, which would minimize the cost of borrowing. Instead, individuals seem to allocate repayments using a balancing-matching heuristic by which the share of repayments on each card is matched to the share of balances on each card. We show that balance matching captures more than half of the predictable variation in repayments, performs substantially better than other models, and is highly persistent within individuals over time. Consistent with these findings, we show that machine learning algorithms attribute the greatest variable importance to balances and the least variable importance to interest rates in predicting repayment behavior.

Keywords: credit cards, consumer borrowing, rational behavior, balance matching, heuristics

JEL Codes: D12, D14, G02, G20

* We thank our discussant Ben Keys for thoughtful and constructive feedback, and we are grateful to David Laibson, Devin Pope, and Abby Sussman for helpful comments. Hanbin Yang provided excellent research assistance.

† University of Nottingham, School of Economics; Network for Integrated Behavioural Science. Email: john.gathergood@nottingham.ac.uk.

‡ University of Chicago, Booth School of Business and NBER. Email: neale.mahoney@gmail.com.

§ University of Warwick, Department of Psychology. Email: neil.stewart@warwick.ac.uk.

¶ University of Nottingham, School of Economics; Network for Integrated Behavioural Science; Centre for Decision Decision Research and Experimental Economics. Email: joerg.weber@nottingham.ac.uk.

1 Introduction

Borrowing decisions underpin a broad set of economic behavior. Individuals borrow to smooth their consumption over the life-cycle, invest in human capital, and purchase durable goods, among other reasons. Thus, understanding how individuals borrow has implications for many fields of economic research, and for consumer financial policy.

This paper aims to shed light on this question by studying how individuals choose to repay debt – and thus implicitly how to borrow – across their portfolio of credit cards. We have a dataset with rich information on credit card contract terms, billing information and repayments for 1.4 million individuals in the United Kingdom over a two-year period. Unlike other leading credit card datasets, our data allows us to link multiple credit card accounts held by the same individual.[1] We study how individuals choose to allocate repayments across their credit cards, holding the total repayment amount fixed.

The credit card repayment decision is an ideal laboratory for studying borrowing because optimal behavior – that is, behavior that minimizes interest charges – can be clearly defined. Consider individuals with exactly two cards. Holding the total amount repaid on both cards in a particular month fixed, it is optimal for individuals to make the minimum payment on both cards, repay as much as possible on the higher interest rate card, and only allocate further payments to the lower interest rate card if they are able to pay off the higher interest rate card in full. What sets the credit card repayment decision apart from many other financial decisions is that optimal behavior does not depend on preferences (such as risk preferences or time preferences).[2] This allows us to evaluate models of optimal and heuristic behavior without having to jointly identify (heterogeneous) preferences.

We start by showing that Ponce et al.'s (2017) finding of non-optimal credit card borrowing

[1] For instance, neither the OCC's Consumer Credit Panel nor the CFPB's Credit Card Database are designed to permit linking of accounts held by the same individual. The credit bureau datasets that combine information from multiple accounts held by the same individual do not have information on interest rates or repayments. There are a number of opt-in panels such as the Mint.com data and Lightspeed Research's "Ultimate Consumer Panel" that have information on multiple cards, but only for a self-selected sample of individuals.

[2] For example, optimal mortgage choices are dependent on risk preferences (in the decision to use an adjustable or fixed rate mortgage) and time preferences over the real option to refinance in the future (see, Campbell and Cocco, 2003). There are very few institutional settings in which optimal mortgage choices can be clearly defined, such as in the Danish mortgage market (see, Andersen et al., 2015). The optimal credit card *spending* allocation is dependent on rewards programs, such as cash-back or airline points. Even when the terms of the rewards program are known, the optimal spending allocation depends on individuals' (idiosyncratic) value of features.

in Mexico is highly robust to the U.K. credit card market where we have data. Our baseline analysis focuses on individuals who hold exactly two cards in our data. For these individuals, the average difference in Annual Percentage Rate (APR) between the high and low interest rate cards is 6.5 percentage points, approximately one-third of the average 19.7% APR in our sample. If these individuals were completely unresponsive to interest rates, it is natural to assume that they would allocate 50.0% of their payments to each card on average. To minimize interest charges, we calculate that individuals should allocate 70.8% of the payments to the high APR card.[3] We show that individuals allocate only 51.2% of their payments to the high APR card, behavior that is virtually indistinguishable from the completely non-responsive baseline. Establishing this result is not the main focus of our analysis, but a necessary first step before going on to investigate alternative models.

If individuals do not optimally allocate their credit card repayments, what explains their repayment behavior? One potential explanation is that individuals face a fixed cost of optimization – such as the time, psychological, or cognitive costs associated with determining the optimal repayment allocation (Sims, 2003). For some individuals, the reduction in interest costs may be too low to rationalize this fixed cost. We show, however, that the share of misallocated repayments does not decline even for individuals with the largest differences in interest rates across cards (more than 15 percentage points) or for individuals who repay the largest amounts (more than £800 in a month). The observed behavior, thus, seems inconsistent with a fixed-cost model of optimization. We also find that the degree of misallocation does not decrease with time since account opening, indicating that learning cannot explain the observed behavior.

The main contribution of this paper is to evaluate heuristics that might better explain the observed allocation of credit card payments. We first consider a balance-matching heuristic under which individuals match the share of repayments on each card to the share of balances on each card. Making payments in proportion to balances might result from the fact that balances are saliently displayed on credit card statements. The balance-matching heuristic is also closely related to "matching" behavior that has been observed in other domains (discussed below), and

[3] The number is less than 100% because we require individuals to make the minimum payment on the low interest rate card and because individuals who can pay off the full balance on the high APR card should allocate remaining payments to the low interest rate card.

thus may result from a deeper underlying tendency for proportionality in decision-making. We do not propose balance matching as a precise description of individual repayment behavior, but instead as an approximate repayment heuristic which individuals might adopt to guide their repayment decisions. We also consider four alternative heuristics, such as the "debt snowball method" (under which payments are concentrated on the card with the lowest balance), which is recommend by some financial advisors.

We assess the explanatory power of these different repayment models using standard measures of goodness-of-fit (root mean square error, mean absolute error) and by calculating the correlation between predicted and observed repayments. To provide a lower benchmark, we calculate goodness-of-fit under the assumption that the percentage of repayments on the high APR card is randomly drawn from a uniform distribution with support on the 0% to 100% interval. To provide an upper benchmark, we use machine learning techniques to find the repayment model that maximizes out-of-sample fit using a rich set of explanatory variables.

We find that balance matching captures more than half of the "predictable variation" in repayment behavior. That is, based on the the range determined by the lower benchmark of random repayments and the upper benchmark of the machine learning models, we find that balance matching is closer to the upper benchmark on all of our measures. We also show that the optimal repayment rule and the other heuristic models do not come close to balance matching in their ability to match the data, capturing less than a quarter of the predictable variation for most of our measures.

In addition to providing us with an upper benchmark, the machine learning models also provide use with a "model free" method for assessing the relative importance of interest rates versus balances in predicting repayment behavior. Consistent with the poor fit of the optimal repayment rule, we find that interest rates have the lowest variable importance (i.e., partial R-squared) in our machine learning models. Indeed, in the decision tree model that maximizes out-of-sample fit, interest rates do not even enter the model. Consistent with balance matching results, we find that balances have the highest variable importance, with importance factors nearly twice as large as any of the other explanatory variables.

We also evaluate each of our models in "horse race" type analysis where we determine the best fit model on an individual × month basis. In binary tests, balance matching is the best fit model for twice as many observations as either the uniform, optimal, and other heuristics models. Balance matching performs comparably to the machine learning models. We also show that balance matching exhibits a high degree of persistence within individuals over time, suggesting that balance matching is more than a good statistical model but is actually capturing a stable feature of individual decision-making.

An alternative explanation for the balance-matching result could arise from individuals anchoring their repayments to minimum payment amounts (Keys and Wang, 2017). If minimum payments are proportional to balances and individuals allocate repayments across cards based on relative minimum payments (or set payments at multiples of minimum payment amounts), then anchoring on minimum payments could produce the observed balance-matching behavior. We evaluate this alternative explanation by exploiting non-linearities in minimum payment rules. Most minimum payment amounts are calculated as the maximand of a fixed amount (the "floor") and a percentage of the balance (the "slope"). For individuals with lower balances, the floor is binding and minimum payments do not vary with balances. We show that for individuals with binding "floors," the allocation of repayments is strongly correlated with balance-matching amounts but not with the ratio of minimum payments, indicating that minimum payments are not driving our findings. We note that while minimum payments are not driving our findings, our analysis does *not* imply that minimum payments are irrelevant for repayment behavior. Our point is simply that minimum payments do not seem to be generating a spurious balance matching result.

Our findings are related to a number of strands of literature. Our result on non-optimal repayments is closely related to the aforementioned Ponce et al. (2017) study and a working paper on the same topic by Stango and Zinman (2015). Ponce et al. (2017) study borrowing using linked data from Mexico and also find that borrowing is highly non-optimal. Stango and Zinman (2015) use data from Lightspeed Research's "Ultimate Consumer Panel", an opt-in sample of U.S. borrowers. In contrast to our results, they find that individuals are much more likely to make optimal allocations when the stakes are large. However, it is unclear whether these findings are unique to the opt-in sample of individuals that they study. The first, modest, contribution of our paper is to show that the non-optimal behavior documented in Ponce et al. (2017) in Mexico extends to the U.K. market that we study.

Our main result on balance matching relates to a literature in psychology and economics on heuristics in individual decision-making. The fact that individuals focus on balances, which are prominently displayed on credit card statements, connects to a literature on how saliently placed information can provide an anchor for choices (e.g., Tversky and Kahneman, 1974; Ariely et al., 2003; Bergman et al., 2010). Our finding on balance matching also shares a resemblance with a long line of research on probability matching. For instance, Rubinstein (2002) shows, in an experimental study, that subjects diversify when choosing between gambles with a 60% and 40% chance of winning, even though the option with a 60% chance of winning dominates any other strategy (see Vulkan, 2000 for a review of this literature). Balance matching is also reminiscent of the classic Benartzi and Thaler (2001) result on how investors in defined-contribution saving plans allocate funds such that the proportion invested in stocks depends strongly on the proportion of stock funds in the choice set.[4]

The caveats to our analysis largely stem from the fact that we focus on the allocative decision of how individuals split repayments across their portfolio of credit cards. While this decision greatly simplifies the analysis, our estimates of the degree of non-optimal behavior should be interpreted as lower bounds relative to a counterfactual in which individuals could additionally reallocate debt payments across non-credit card loans (such as mortgages or automobile loans) or make adjustments on the extensive margin (e.g., by adjusting the tradeoff between debt repayment and consumption). Our focus on the allocative decision also naturally leads us to consider "allocative heuristics," such as balance matching, rather than heuristics that determine behavior on the extensive margin. For example, balance matching could arise from individuals repaying a fixed percentage (e.g., 10%) of their balances, a rule-of-thumb that would lead to inefficient behavior on both the allocative and extensive margins.

The rest of the paper proceeds as follows. Section 2 describes our data and presents summary statistics for our baseline sample. Section 3 presents our results on the optimality of repayment behavior. Section 4 examines rounding and a $1/n$ rule for repayments. Section 5 lays

[4] See DellaVigna (2009) for a review of the evidence on choice heuristics using field data.

out alternative heuristics for debt repayment, including the balance-matching heuristic. Section 6 tests between these repayment models. Section 7 presents sensitivity analysis. Section 8 concludes.

2 Data

2.1 Argus Credit Card Data

Our data source is the Argus Information and Advisory Services' "Credit Card Payments Study" (CCPS). Argus provided us with detailed information on contract terms and billing records from five major credit card issuers in the UK. These issuers have a combined market share of over 40% and represent a broad range of credit card products and market segments. We have obtained a 10% representative sample of all individuals in the CCPS who held a credit card between January 2013 and December 2014 with at least one of the five issuers. Unlike other leading credit card datasets, the CCPS provides us with anonymized individual-level identifiers that allow us to link multiple accounts held by the same individual.[5]

2.2 Sample Restrictions

Our interest lies in understanding how individuals make repayment decisions across their portfolio of credit cards. Our unit of analysis is the individual × month. All of the credit cards in our data require payments on a monthly basis. We consider cards to be in the same "month" if their billing cycles conclude in the same calendar month. We examine sensitivity to this assumption in Section 7.

For our baseline analysis, we focus our analysis on the roughly 250,000 individuals who hold exactly two cards for at least part of the sample period. We do this to simplify the exposition. Relative to individuals with three or more cards, it is easier to explain how we calculate optimal repayments and easier to display deviations from optimality in figures and tables in the two-card sample. We do, however, show that our main findings are robust to considering individuals

[5] As discussed in Footnote 1, neither the OCC's Consumer Credit Panel nor the CFPB's Credit Card Database are designed to permit matching of multiple individually-held accounts, and credit bureau datasets typically do not have information on interest rates or repayments. Opt-in panels such as Lightspeed Research's "Ultimate Consumer Panel" have information on multiple cards, but only for a self-selected sample of individuals.

with three or more cards.

We restrict our attention to observations in which individuals face *economically meaningful* decisions about how to allocate their repayment across cards. This is crucial for our analysis: for instance, individuals who repay balances on both cards in full each month ("transactors") never generate interest charges and therefore do not face an allocative decision problem with economic consequences. Specifically, we restrict the sample to observations where the individual (i) holds a revolving balance on both of their cards, (ii) makes at least the minimum repayment on both cards, (iii) pays more than the minimum repayment on at least one card, and (iv) does not pay both cards down in full. For these observations, the individual faces a choice over how much to repay on each card, with interest charges necessarily incurred on *at least* one card. For this sample we can cleanly define optimal repayment behavior and also behavior under heuristic rules, all of which allocate repayments across cards while honouring these sample restrictions.[6] Applying these sample restrictions provides a baseline sample of approximately 395,000 individual $\times$ month observations which represent 68% of all revolving balances across all two $-$ card individual $\times$ month observations.

Table 1 provides summary statistics on the baseline sample. The average difference in APR (for purchases) between the high and low interest rate card is 6.3 percentage points, or approximately one-third of the 19.7% average purchase APR in the sample. Yet despite this substantial difference in prices, by most measures, individuals borrow more on the *high* APR card. While monthly repayments are marginally larger on the high APR card (£260 versus £230), revolving balances are larger on the high APR card (£2,197 versus £2,049). This is all the more striking given that average credit limits are almost three times larger than revolving balances (£6,386 versus £6,010), indicating that the average borrower would be able to shift all of their borrowing to the lower APR card without exceeding their credit limit.

[6] By construction, we draw a sample of observations in which individuals neither underpay their card (i.e. pay less than the minimum on either card) or overpay their card (i.e. pay more than the balance on either card). The rules for credit card repayment we consider in the paper all allocate repayments within these bounds by explicitly incorporating these four restrictions within the rule. Hence we rule out strategically non-paying, or over-paying, a card.

3 Actual and Optimal Repayment

We start by comparing the actual and interest-cost-minimizing allocation of repayments across cards. We refer to the interest-cost-minimizing allocation as the "optimal" allocation because it is hard to think of a (reasonable) scenario where minimizing interest costs would not be optimal. Holding the total repayment amount on both cards fixed, it is optimal for individuals to make the minimum required payment on both cards, repay as much as possible on the card with the higher interest rate, and only allocate further payments to the lower interest rate card if they are able to pay off the high interest rate card in full.[7]

We focus on repayments, rather than other measures of credit card use like spending or revolving balances, because, for repayments, we can clearly define optimal behavior. In contrast, optimal spending behavior may depend upon rewards programs, which we do not observe in our data.[8] We also do not focus on the optimality of revolving balance allocations because revolving balances are a "stock" that cannot typically be quickly adjusted.[9] Thus, determining whether revolving balances are "optimal" would require us to take a stand on how individuals could reallocate revolving balances through counterfactual spending and repayment decisions over a sequence of months (if not years).

Figure 1 plots actual and optimal repayments for individuals with different numbers of cards. As discussed in Section 2, we restrict the samples to individual $\times$ months in which individuals face an economically meaningful allocative decision. Panel A plots the distribution of actual and optimal payments in the two-card sample. The distribution of actual repayments appears close to symmetric, with a mass point at 50%, and smaller mass points 33% and 67%. In contrast, the distribution of optimal repayments is heavily weighted towards the high APR card.

Panels B to D show radar plots of the average percentage of actual and optimal payments

[7] We explicitly rule out the possibility that choosing not to make the minimum payment on the low interest rate card could be optimal. Failing to repay the minimum repayment on the low-APR card would incur a penalty fee and a marker on the individual's credit file.

[8] While issuers typically incur only a small cost for the rewards they provide – approximately 1%, see Agarwal et al. (2015) – individuals might value rewards (such as airline points) at a high enough value to affect optimal spending decisions.

[9] In particular, with the exception of balance transfer products in the prime credit card market, individuals can only reallocate their stock of revolving balances by adjusting the flow of spending and repayment on a month-by-month basis.

on each card for the samples of individual × months in which individuals hold 3, 4, and 5 cards. In each of the plots, the cards are ordered clockwise from highest to lowest APR (starting at the first node clockwise from 12 o'clock). As in the two-card sample, the actual percentage of payments is very similar across cards, but it would be optimal to allocate a substantially higher percentage of payments to the highest APR card and a substantially lower percentage to the card with the lowest APR.

Summary data for actual and optimal repayments for the two-card sample is shown in Table 2. On average, individuals should allocate 70.8% of repayments to the higher-APR card, whereas they actually allocate 51.2% to that card. Hence, individuals misallocate 19.6% of their total monthly payment on average. In Figure A1 we plot misallocated repayments in *excess of the minimum* payment. That is, we subtract out the amount required to make the minimum payment on each card and then calculate the share of the remaining amount that is allocated across cards. On average, individuals should allocate 97.1% of payments in excess of the minimum to the high-APR card, whereas in practice they actually allocate 51.5% to that card.[10] Summary data for payments in excess of minimum are shown in Table A2.

3.1 Fixed Costs of Optimization

One potential explanation for the non-optimality of repayments is that some individuals face a fixed cost of optimization – such as the time, psychological, or cognitive costs associated with determining the optimal repayment strategy (Sims, 2003). For some individuals, the reduction in interest payments from cost-minimizing may be too low to rationalize incurring this fixed cost.

To investigate this potential explanation, we examine the correlation between the percentage of misallocated repayments and the economic stakes of the repayment decision. We define misallocated payments as difference between optimal and actual payments on the high APR card. We examine two measures of the economic stakes: (i) the difference in APR across cards and (ii) the total repayments made that month. Since the gains from optimizing are increasing in the economic stakes, under the fixed cost explanation, the percentage of misallocated repay-

[10] The number is not exactly 100% because sometimes individuals can pay off the full balance by allocating a smaller amount, in which case they should allocate the remaining amount to the low interest rate card.

ments should be declining in both measures. Moreover, for individuals with large economics stakes, we would expect the degree of misallocation to be close to zero.

Panel A of Figure 2 shows a binned-scatter plot of the percentage of misallocated payments against the difference in APR between the high and low interest rate cards. The binned-scatter plot is constructed by partitioning the x-axis variable into 20 equal-sized groups and plotting the mean of the y-axis and x-axis variables for each group.[11] The flat relationship indicates that individuals are not less likely to misallocate repayments even when there is a large APR difference (more than 15 percentage points).[12]

Panel B of Figure 2 shows a binned-scatter plot of the percentage of misallocated payments against total repayments across both cards. Again, there is no evidence of a decreasing relationship. Indeed, the relationship is increasing due to the fact that individuals who make the largest payments can cover the minimum on the low interest rate card with a smaller percentage of their overall allocation and thus *should* allocate an even larger fraction of payments to the high APR card.[13]

Another potential explanation for the observed non-optimal behavior is that individuals learn over time (e.g., since opening a card), and that our analysis of the cross-sectional distribution of repayments masks this learning behavior. A model with time-varying adjustment costs (in the spirit of Calvo, 1983) would also generate a gradual reduction in the degree of misallocation over time.

Panel C of Figure 2 examines this explanation by showing a binned-scatter plot of the percentage of misallocated payments against the age (in months) of the high APR card. For this analysis, we restrict the sample to individuals who open a high APR during our sample period and for whom we can observe economically meaningful allocation decisions for 10 consecutive months. In the plot, the horizontal axis starts in the second month after opening, since this is the first month in which individuals could have a balance on the high APR card to repay. The plot shows no evidence of a reduction in the percentage of misallocated repayments over time.

[11] See Chetty et al. (2014) for more details on the binned-scatter plot methodology.

[12] Panel A of Figure A1 illustrates the relationship between misallocated payments in excess of the minimum payment and the difference in APR across cards.

[13] Panel B of Figure A1 illustrates the relationship between misallocated payments in excess of the minimum payment and the total repayment across both cards. There is a slight downward slope, but certainly not the type of relationship that would be predicted by a fixed-cost-of-optimization model.

This finding suggests that neither learning nor time-varying adjustment costs can explain the observed non-optimizing behavior.[14]

4 Rounding and the 1/n Rule

The spike in repayments at 50% (see Panel A of Figure 1) suggests that some individuals use a simple $1/n$ heuristic in which they make equal-sized repayments across cards, analogous to the $1/n$ heuristic documented in defined-contribution savings decisions (Benartzi and Thaler, 2001). However, an alternative explanation is that individuals round payments to £50, £100, £200, and so on. If an individual rounds up a payment on card A from £80 to £100 and rounds down a payment on card B from £120 to £100, then the individual would appear as if she intended to make equal-sized payments, even though, absent rounding, the share of payments on each card would be substantially different from 50%.

Figure 3 investigates this competing explanation for the spike at 50%. Panel A plots the distribution of payments in £s, and shows substantial evidence of rounding. We calculate that 19.2% of payments take on values that are multiples of £100, and 33% of payments take on values that are multiples of £50 (which obviously includes payments that are multiples of £100). Panel B shows the percentage of payments on the high APR card for the subset of accounts that make payments that are multiples of £50; Panel C shows the percentage of payments on the high APR card for the subset of accounts that pay other ("non-round") amounts.

The plots show that the peaks at 50% repayment on the high APR card (as well as 33% and 66%) are heavily concentrated among individuals who make round number repayments (defined as multiples of £50). Among the remainder of the sample, who do not make round number payments, there is only a small spike at 50%, and no discernible spike at 33% or 66%.

We therefore view the spike at 50% as a red herring that, while temping to over-interpret, does not provide compelling evidence for the $1/n$ heuristic. Instead, we think that the more likely explanation for the spike at 50% is rounding behavior. Of course, since we do not have random variation in whether individuals round, we cannot rule out the possibility that some

[14] Panel C of Figure A1 illustrates the relationship between misallocated payments in excess of the minimum payment and the age (in months) of the higher APR card.

individuals would have allocated 50% on the higher APR card if they had counterfactually not rounded their payments. Thus, although we view this behavior as unlikely, we include the $1/n$ rule in some of the analysis that follows.

5 Balance Matching and Other Heuristics

If individuals do not optimally allocate their credit card repayments, what explains their behavior? In the remainder of this paper, we evaluate heuristics that might better explain the allocation of credit card repayments. In this section, we introduce the set of heuristics that we consider. In Section 6, we evaluate the explanatory power of these models.

5.1 Balance Matching

We first consider a balance-matching heuristic by which individuals match the share of repayments on each card to the share of balances on each card. Let $k = \{A, B\}$ index cards, q_k indicate balances and p_k indicate payments. In a two-card setting, balance-matching payments are given by

$$\frac{p_A}{p_B} = \frac{q_A}{q_B}. \tag{1}$$

subject to the constraint that the individual pays at least the minimum on both cards and no more than the full balance on either card.[15]

Why would repayments follow a balance-matching heuristic? First, as shown in Figure A3, balances are perhaps the most prominently displayed element on credit card statements. The psychological theory of *anchoring* (Tversky and Kahneman, 1974) suggests that individuals might make payments in relation to this saliently displayed amount (instead of less saliently displayed interest rates).[16,17]

[15] As discussed earlier, we explicitly exclude in our sample selection observations which might be best explained by heuristics under which individuals might underpay either card (pay less then the minimum), or overpay either card (pay more than the balance).

[16] A second reason why balances may be more salient is that balances are denote in the same units as repayments (£s), whereas APR take on different units (%).

[17] Balances also enter the minimum payment formula. Therefore, at least in principle, repayments might depend on balances indirectly through the minimum payment amount. We discuss this issue in Section 7 and show that this channel is unlikely to explain our results.

Second, the balance-matching heuristic is closely related to "matching" behavior that has been observed in other domains, and thus may result from a deeper underlying tendency for proportionality in decision-making. For instance, the *probability matching* literature finds that individuals place bets in proportion to the probability of payoffs, even though betting on the option with the highest probability of payoff first-order stochastically dominants any other decision.[18] Herrnstein's (1961) *matching law* is based on the observation that pigeons peck keys for food in proportion to the time it takes the keys to rearm rather than concentrating their effort on the key that rearms most quickly. Balance matching is also reminiscent of the classic Benartzi and Thaler (2001) result on how investors in defined-contribution saving plans allocate funds such that the proportion invested in stocks depends strongly on the proportion of stock funds in the choice set.[19]

Of course, we do not propose balance matching as a precise description of individual repayment behavior. Pigeons do not measure the time is takes keys to rearm with a stopwatch and we do not mean to suggest that individuals use long division to calculate the share of repayments that should be allocated to each card. Instead, we propose that individuals approximate balance matching in their repayment behavior. Indeed, since credit card balances are fairly stable over time, an individual could approximate a balance matching rule without knowing the exact balance on each card in any given month.

5.2 Other Heuristic Models of Repayment

We also consider four alternative heuristics that capture intuitive economic and non-economic approaches to the allocation of payments. Some of these heuristics are based on the capacity of a credit card, which we define as the difference between the credit limit and current balance in £s.

- **Heuristic 1**: *Repay the card with the lowest capacity.* Allocate payments to the lowest capacity card, subject to paying the minimum on both cards. Once capacity is equalized across cards, allocate additional payments to both cards equally. Intuitively, by focusing

[18] For example, Rubinstein (2002) shows in an experimental study that subjects diversify across independent 60%-40% gambles even though betting on the gamble with a 60% probability of payout is a strictly dominant strategy. See Vulkan (2000) for a review on this literature.

[19] See DellaVigna (2009) for a review of the evidence on choice heuristics using field data.

payments on the card with the lowest capacity, this heuristic reduces the risk that an accidental purchase will put an individual over their credit limit, which would incur an over-limit fee.

- **Heuristic 2**: *Repay the card with highest capacity.* Allocate payments to the highest capacity card, subject to paying the minimum on both cards. Once the highest capacity card is fully repaid, allocate remaining payments to the other card. Intuitively, by allocating payments to the card with the highest capacity, this heuristic creates maximum “space” for making a large purchase on a single card (e.g., buying a television).
- **Heuristic 3**: *Repay the card with the highest balance.* Allocate payments to the highest balance card, subject to paying the minimum on the other card. Once balances are equalized across cards, allocate additional payments to both cards equally. If individuals dislike having a credit card with a large balance, this heuristic reduces the maximum balance they are carrying, and thus might explain repayment behavior.
- **Heuristic 4**: *Repay the card with the lowest balance (“debt snowball method”).* Allocate payments to the lowest balance card, subject to paying the minimum on the other card. Once the balance on the lowest balance card is paid down to zero, allocate any additional payments to the other card. This heuristic is sometimes referred to as the debt snowball method by financial advisors. Proponents argue that paying off a card with a low balance generates a “win” that motivates further repayment behavior.

6 Testing Repayment Models

We evaluate balance matching and the other models using two statistical approaches. First, we assess the explanatory power of each of the models using standard measures of goodness-of-fit. Second, we evaluate the performance of our models in “horse race” type analysis where we determine the best fit model on an individual × month basis.

6.1 Goodness-of-Fit

For our analysis of goodness-of-fit, we focus on the two-card sample. We consider three measures of goodness-of-fit: the square root of the mean squared error (RMSE), the mean absolute error (MAE), and the correlation between predicted and actual payments (Pearson correlation).[20] Our findings are consistent across each of these measures.

We start by establishing lower and upper benchmarks for model fit. For a lower benchmark, we calculate goodness-of-fit under the assumption that the percentage of repayments on the high APR card is randomly drawn from a uniform distribution with support on the 0% to 100% interval. To provide an upper benchmark, we use machine learning techniques to construct a set of purely statistical models of repayment behavior. Specifically, we estimate decision tree, random forest, and extreme gradient boosting models of the percentage of payments allocated to the high APR card. We use APRs, balances, and credit limits on both cards as input variables and "tune" the models to maximize out-of-sample power. Technical details are provided in Appendix I.

Table 3 shows our measures of goodness-of-fit under each of the models. The lower benchmark of uniformly distributed payments has a RMSE of 36.4, MAE of 29.9, and a Pearson correlation of 0. The optimal model yields only a very small improvement in the RMSE and a modest improvement in the MAE. The optimal model does generate an economically meaningful increase in the Pearson correlation, although this is partly due to the fact that the lower benchmark is constructed to have a Pearson correlation of 0.

The balance-matching model fits repayments substantially better than both the lower benchmark of uniformly distributed payments and the optimal repayment model. Panel A of Figure 4 shows the distribution of actual and balance matching payments on the high APR card. The balance-matching heuristic naturally does not fit the spike in repayments at 50%, but otherwise fits the marginal distribution of actual payments fairly well. Panel B of Figure 4 displays the joint distribution of actual and balance matching payments using a contour plot. The ridge along the 45-degree line indicates that the distributions are correlated. Figure A4 shows that the balance-matching heuristic fits actual repayments particularly well in the samples of 3, 4, and 5 cards.[21] As shown in Table 3, the goodness-of-fit measures for the balance matching model fall slightly more than halfway between the lower and upper benchmarks,

[20] The Pearson correlation is the square root of the R-squared.

[21] Summary statistics for balance matching repayments on the higher cost card and actual repayments are shown in Table A3.

suggesting that balance matching captures more than half of the predictable variation in repayment behavior.

The other heuristics do not come close to balance matching in their ability to fit the data. Figure A5 shows the marginal density (left column) and joint density (right column) under each of the other heuristics. One common feature of these other heuristics is that they predict that individuals should often concentrate their repayments on a single card only. For instance, under Heuristic 1 (repay the card with the lowest capacity), individuals should fully allocate repayments, in excess of the minimum, to the card with the lowest capacity until the point where both cards have equal capacity remaining. Individuals, however, seem to avoid "corner solutions" in their repayment behavior. The marginal density plots (left column) show that these heuristics over-predicted the share of individuals who allocate a very small (less 10%) or very large (great than 90%) share of payments to the high APR card.[22] Based on our measures of goodness-of-fit, Table 3 shows that these heuristically typically fall less than a quarter of the way between the lower and upper benchmarks, or capture less than a quarter of the predictable variation in repayment behavior.

There are two ways to view this analysis from the perspective of the balance-matching model. The glass half full view is is that being able to capture more than half of the predictable variation in repayment behavior with a simple balance matching model is useful. Balance matching is a useful description of behavior because it is easy to understand, reinforces existing theories of behavior (e.g., probability matching, Herrnstein's matching law), and might provide intuition for individual behavior in yet-to-be-studied environments. The glass half empty perspective is that machine learning techniques provide substantially higher predictive power. Thus, if the goal is prediction – rather than understanding human behavior – machine learning techniques may be preferable.

6.1.1 Balances and APRs in Machine Learning Models

In addition to providing us with an upper benchmark, the machine learning models also provide us with a "model free" method for assessing the relative importance of different variables in

[22] Summary data for repayments under each heuristic is shown in Table A4.

predicting repayment behavior. Consistent with the balance matching results, the machine learning models confirm that balances are hugely important for predicting behavior. Table A5 shows that for the random forest and extreme gradient boosting models, balances have highest variable importance (i.e., partial R-squared), with importance factors nearly twice as large as any of the other explanatory variables. This indicates for *any* model to have high predictive power, it would need to have balances play an important role.

Consistent with the poor fit of the optimal repayment rule, we find that APRs have the lowest variable importance. In particular, Figure A6 shows that APRs (on either card) do not enter the decision tree model that maximizes out-of-sample fit. Alternatively put, while optimal behavior depends almost fully on "prices," the decision tree model *completely ignores* prices in predicting the allocative decision. Table A5 shows that for the random forest and extreme gradient boosting models, the variable importance is lowest for APRs.[23] Table A6 shows that including the APR variables only marginally improves the goodness-of-fit of these models.

6.2 Horse Races Between Alternative Models

The goodness-of-fit analysis effectively measures the distance between observed repayments and predicted repayments under each of our models. An alternative approach to evaluating the models is to conduct "horse race" type analysis where we determine the best fit model on an observation-by-observation basis. A model that fits a small number of observations very poorly, but a larger number quite well, would perform poorly under most distance metrics (and especially those with increasing loss functions) but might perform well using this alternative approach.

6.2.1 Methodology

Let i denote individuals and t denote months. Let p_{it} indicate the actual share of payments that is allocated to the high APR card and let $\hat{p}^j_{it}$ indicate the share of payments on the high APR card predicted by model $j \in J$. To test between alternative models, we estimate specifications

[23] Table A7 confirms that APRs and balances on cards are not collinear. In cases where variables are collinear the interpretation of variable importance may be spurious.

of the form:

$$p_{it} = \left(\sum_{j \in J} \lambda_{it}^{j} \hat{p}_{it}^{j} \right) + \epsilon_{it} \qquad \text{s.t.} \quad \lambda_{it}^{j} \in \{0, 1\} \quad \text{and} \quad \sum_{j} \lambda_{it}^{j} = 1, \tag{2}$$

where the λ_{it}^{j} are indicators that "turn on" for one and only one of the candidate models $j \in J$. We vary the set of alternatives J to allow for horse races among different competing sets of models. Intuitively, for each observation p_{it}, this procedure picks the model j that best fits observed repayment behavior at the individual × month level.

It is worth pointing out that our ability to identify the best-fit model at the individual × month level is due to the unique nature of the credit card repayment decision. As discussed in Section 1, what sets credit card repayments apart from many other financial decisions is that optimal behavior does not depend on preferences (such as risk preferences or time preferences). If we needed to recover preferences, then identifying the best fit model would require *jointly* estimating preferences and behavior under the different models. The best we could do would be to estimate heterogeneous preferences for different demographic subgroups and then ask which model best fits behavior on a subgroup-by-subgroup level. What allows us to perform this exercise at the individual × month level is that we do not need to recover preferences to generate predicted behavior under the different repayment models.

There are a few subtle issues involved in the estimation. The first involves "corner solutions." When a model predicts that an individual should pay less than the minimum amount or more than the entire balance on a given card, we set the payments to the corner value and assign the remaining payments to the other card (leaving total repayment across both cards fixed). The second issue involves "ties". When there are multiple models that are tied for closest to the actual repayments, the λ_{it}^{j} are not identified, and so we drop these observations from our analysis. We estimate the model by minimizing the absolute deviation between the observed and predicted values.[24]

[24] In practice, since we are estimating a separate set of coefficients for every individual × month, the estimates would be identical under a quadratic (or any other monotonically increasing) loss function.

6.2.2 Results

Table 4 shows results of this horse race analysis in the pooled sample of individual × months. Panel A compares each of our models one-by-one against the lower benchmark where the percentage of repayments on the high APR card is randomly drawn from a uniform distribution with support on the 0% to 100% interval. In a binary comparison, balance matching is the best fit model for 67% of observations, or about twice the percentage of the uniform benchmark. The optimal model and the other heuristics are closest for slightly more than half of the observations, and therefore only perform slightly better than the uniform benchmark. In binary comparisons, the machine learning models have the best fit for between 63% and 72% of observations, which is similar to balance matching.[25]

Panel B of Table 4 compares each of the models one-by-one to the balance matching model.[26] In a horse race with the optimal model, balance matching has the best fit for slightly more than two-thirds of observations. When compared with the other heuristic models, balance matching is also the best fit model for approximately two-thirds of observations. Balance matching performs comparably to the machine learning models, with balance matching exhibiting the best fit for 45% to 56% of observations.

To the extent that we think of the competing models as actually representing different models of individual decision-making, we would naturally expect the best-fit model to be persistent within individuals over time. Table 5 shows the within-person transition matrix for the best-fit model. The sample is restricted to individual × months where we observe repayment behavior for at least two months in a row. For this exercise, we allow the set J to encompass all of the candidate models, and we fix the uniformly distributed repayment to be constant within a individual over time.

The table shows the balance matching exhibits a high degree of persistence – both in absolute value and relative to the other models of repayment behavior. Among individuals whose repayments are best fit by the uniform model in a given month, 24% make repayments

[25] Since the machine learning models were tuned to minimize RMSE, it is natural for these models to perform relatively better when evaluated using RMSE (and other distance metrics) than when evaluated using this type of horse race analysis.

[26] We exclude a comparison of balance matching and the uniform model, since it was shown in Panel A.

that are closest to the uniform model in the next month. This persistence likely reflects the fact that balances and repayments are sticky over time – if the uniform model happens to be accurate in a given month, and balances and payments are sticky, then the uniform model will mechanically be accurate in the next month as well.

The balance-matching model exhibits three-fold greater persistence than the uniform model. Among individuals whose repayments are closest to balance matching in a particular month, 83% make payments that are closest to balance matching in the next month. The high degree of persistence suggests that balance matching is more than a good statistical model but is actually capturing a stable feature of individual decision-making. The optimal model and the other heuristics exhibit persistence in the same range as the uniform model, which may just reflect stickiness in balances and repayments over time. The only other model that exhibits strong persistence is the $1/n$ rule, suggesting that this model also captures a stable feature of individual behavior (or that there is a stable tendency tendency to round repayments).

Taken together, our goodness-of-fit analysis supports the view that balance matching is a powerful predictor of credit card repayments, capturing more than half of the predictable variation in repayment behavior and performing substantially better than the alternative models. In the horse race analysis, balance matching performs at a similar level to the machine learning models, and is highly persistent over time, suggesting it is more than a good statistical model but is actually capturing a stable feature of individual decision-making.

7 Sensitivity Analysis

7.1 Balance Matching and Minimum Payments

An alternative explanation for the balance-matching result could arise from individuals anchoring their repayments to minimum payment amounts (Keys and Wang, 2017). Like balances, minimum payments are prominently displayed on credit card statements (see Figure A3). If minimum payments are proportional to balances and individuals allocate repayments across cards based on relative minimum payments (or set payments at multiples of minimum payment amounts), then anchoring on minimum payments could produce the observed balance-matching behavior.

We evaluate this alternative explanation by exploiting non-linearities in minimum payment rules. Most minimum payment amounts are calculated as the maximand of a fixed amount (the "floor") and a percentage of the balance (the "slope"). For instance, a typical minimum payment formula might be:

$$\text{Minimum Payment} = \max\{£25, 2\% \times \text{Balance}\}.$$

Consider an individual with two cards that both use this formula to calculate minimum payment amounts. Define the *minimum payments-matching* heuristic as the payment allocation under which the share of repayments on each card matches the share of minimum payments amounts on each card. If minimum payments are on the "slope" part of the formula (balances greater than £1,250), and the slopes are identical (2% for both cards), then the balance-matching payments will be near – perfectly correlated with minimum payments-matching payments and it will be impossible to tease apart these mechanisms.[27]. If the slopes differ, then the correlation with balance matching payments will be weaker. If minimum payments are on the "floor" part of the formula (balances less than £1,250), then the balance-matching allocation will not be correlated with the minimum payment-matching allocation. Hence it should be possible to separately identify these mechanisms where cards have different slopes, or where the minimum payment is on the floor.

Figure 5 shows binned scatter plots of the correlation between the actual percentage of payments on the high APR card and predicted payments under the balance-matching and minimum payments-matching heuristics. The top row shows these relationships for individuals where both cards are on the slope of the minimum payment formula and the slope is the same on both cards, the middle row shows these relationships for individuals where the slopes differ and the bottom row shows these relationships for individuals where both cards on the floor part of the formula. Table A9 reports correlation coefficients between balance matching payments and minimum payment matching payments in each of these samples (Panel A) as well as between balance matching payments and actual payments in each sample (Panel B) and

[27] The correlation will not be 1 because the balance matching heuristic incorporates a minimum repayment floor on repayments on both cards

between minimum payment matching payments and actual payments in each sample (Panel C).

In the same slope sample, the balance matching and the minimum payment matching payments are near-perfectly correlated ($\rho = 0.96$). As a result, actual payments exhibit an identical correlation with balance-matching payments ($\rho = 0.62$) and minimum payment matching payments ($\rho = 62$), and we cannot identify whether the observed behavior stems from balance matching or minimum payment matching behavior.

In the different slope sample, the balance matching and the minimum payment matching payments are more weakly correlated ($\rho = 0.84$). The correlation between actual and balance matching payments in the different slopes sample ($\rho = 0.42$). However, the correlation between actual and minimum payment matching is much weaker in the different slopes sample ($\rho = 0.26$) than in the same slope sample.

In the floor sample, there is a much weaker correlation between the balance matching payment and the minimum payment matching payment ($\rho = 0.51$). The correlation between actual and balance matching payments in the floor sample ($\rho = 0.51$) is similar to that in the slope samples. However, the correlation between actual and minimum payment matching is much weaker in the floor sample ($\rho = 0.20$) than in the slope sample.

It thus follows that observed repayment behavior is driven by balance matching and not by individuals setting payments in relationship to minimum payments.[28] We note that while minimum payments do not seem to be driving our findings, our analysis does *not* imply that minimum payments are irrelevant for repayment behavior. Our point is simply that minimum payments do not seem to be generating a spurious balance matching result. Indeed, while not directly comparable, our finding of a modest correlation between actual and minimum payments matching repayments is consistent with (Keys and Wang, 2017), who estimate that 9% to 20% of account-holders anchor their repayments to minimum payment amounts.

[28] To complement the analysis above, Table A8 shows horse race analysis separately for the "floor" and 'slope" samples. In binary tests against the optimal repayment model, the share of observations best fit by balance matching is very similar in both samples. These results further support our conclusion that the observed behavior is driven by balance matching and not anchoring on minimum payments.

8 Conclusion

In this paper, we used linked data on multiple cards from five major credit card issuers in the U.K. to study borrowing behavior in the credit card market. We showed that the allocation of repayments is highly non-optimal, with individuals allocating only 51.5% of their payments to the high APR card, relative to optimal repayments of 70.5%. This finding builds on Ponce et al. (2017), who showed evidence of similar non-optimal behavior in credit card data from Mexico.

The main contribution of our paper was to show that, in contrast to the optimal repayment rule, actual repayment behavior can be explained by a *balance matching* heuristic by which individuals match the share of repayments on each card to the share of balances on each card. In particular, we showed that balance matching captures more than half of the predictable variation in repayments, performs substantially better than other models, and is highly persistent within individuals over time.

We provided additional support for the importance of balances – and irrelevance of interest rates – using machine learning models. Consistent with the poor fit of the optimal repayment rule, we find that interest rates have the lowest variable importance in our machine learning models. Indeed, in the decision tree model that maximizes out-of-sample fit, interest rates do not even enter the model. Consistent with balance matching results, we find that balances have the highest variable importance, with importance factors nearly twice as large as any of the other explanatory variables.

References

Agarwal, S., S. Chomsisengphet, N. Mahoney, and J. Stroebel (2015). Regulating consumer financial products: Evidence from credit cards. *Quarterly Journal of Economics 130*(1), 111–164.

Andersen, S., J. Campbell, K. M. Nielsen, and T. Ramadorai (2015). Inattention and Inertia in Household Finance: Evidence from the Danish Mortgage Market. *NBER Working Paper 21386*.

Ariely, D., G. Loewenstein, and D. Prelec (2003). "Coherent Arbitrariness": Stable Demand Curves Without Stable Preferences. *The Quarterly Journal of Economics 118*(1), 73–106.

Benartzi, S. and R. H. Thaler (2001). Naive diversification strategies in defined contribution saving plans. *American Economic Review 91*(1), 79–98.

Bergman, O., T. Ellingsen, M. Johannesson, and C. Svensson (2010). Anchoring and cognitive ability. *Economics Letters 107*(1), 66–68.

Breiman, L. (2001). Random forests. *Machine learning 45*(1), 5–32.

Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics 12*(3), 383–398.

Campbell, J. Y. and J. F. Cocco (2003). Household Risk Management and Optimal Mortgage Choice. *The Quarterly Journal of Economics 118*(4), 1449–1494.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review 104*(9), 2593–2632.

DellaVigna, S. (2009). Psychology and Economics: Evidence from the Field. *Journal of Economic Literature 47*(2), 315–372.

Friedman, J., T. Hastie, and R. Tibshirani (2001). *The elements of statistical learning*, Volume 1. Springer series in statistics New York.

Friedman, J. H. (2001, October). Greedy function approximation: a gradient boosting machine. *Annals of statistics 29*(5), 1189–1232.

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior 4*(3), 267–272.

Keys, B. J. and J. Wang (2017). Minimum payments and debt paydown in consumer credit cards. *NBER Working Paper*.

Ponce, A., E. Seira, and G. Zamarripa (2017). Borrowing on the Wrong Credit Card: Evidence from Mexico. *The American Economic Review 107*(4), 1335–1361.

Rubinstein, A. (2002). Irrational diversification in multiple decision problems. *European Economic Review 46*(8), 1369–1378.

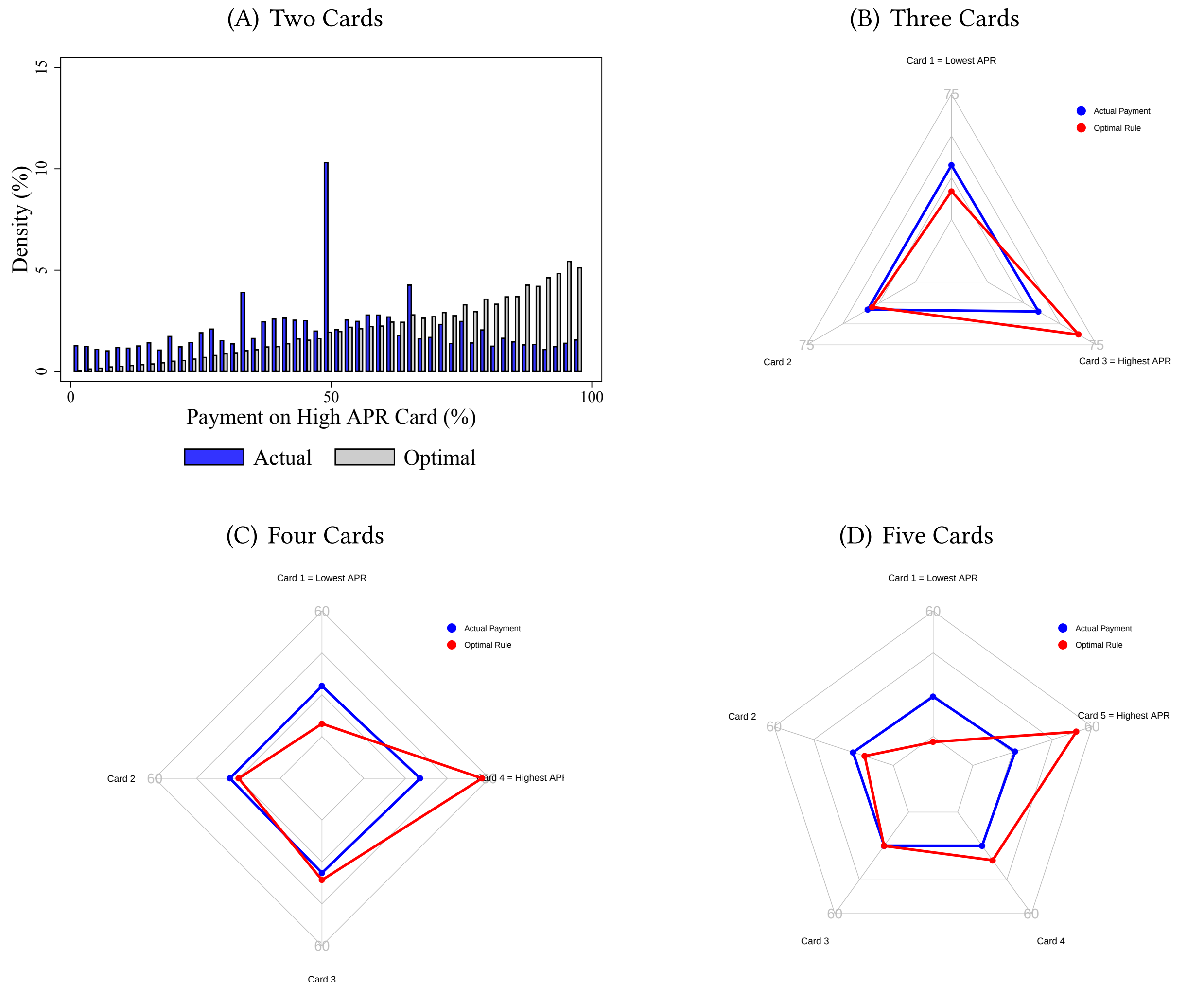
Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics 50*(3), 665–690.

Stango, V. and J. Zinman (2015). Bounding the costs of bounded rationality: Evidence from credit cards. *Working Paper*.

Tversky, A. and D. Kahneman (1974). Judgment under Uncertainty: Heuristics and Biases. *Science 185*(4157), 1124–1131.

Vulkan, N. (2000). An Economist's Perspective on Probability Matching. *Journal of Economic Surveys 14*(1), 101–118.

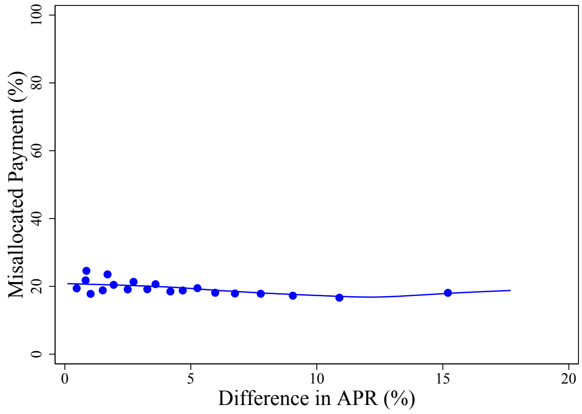
Figure 1: Actual and Optimal Payments



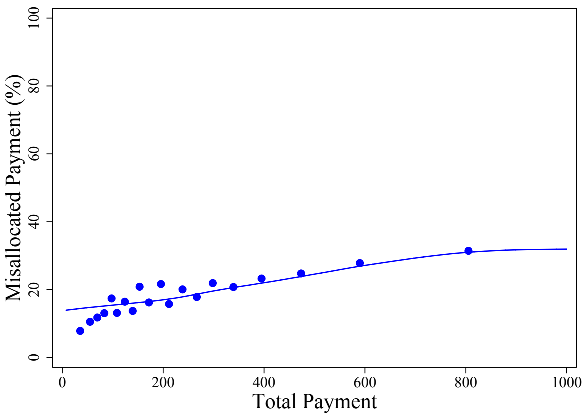
Note: Panel A shows the distribution of actual and optimal payments on the high interest rate card in the two-card sample. Panels B to D show radar plots of mean actual and optimal payments in the samples with 3 to 5 cards. In the radar plots, cards are ordered clockwise from highest to lowest APR (starting at the first node clockwise from 12 o'clock). All samples are restricted to individual × months in which individuals face an economically meaningful allocative decision. See Section 2 for details.

Figure 2: Misallocated Payments by Economics Stakes

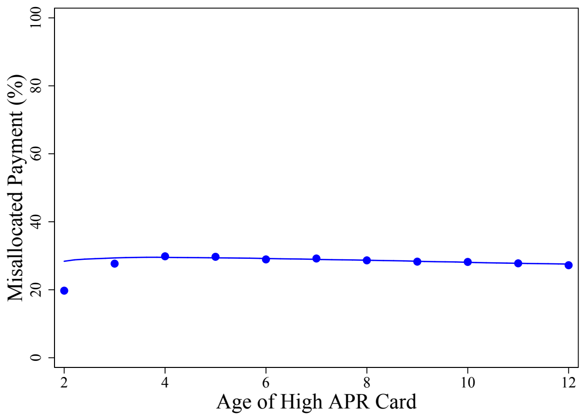
(A) Misallocated vs. Difference in APR



(B) Misallocated vs. Total Payments



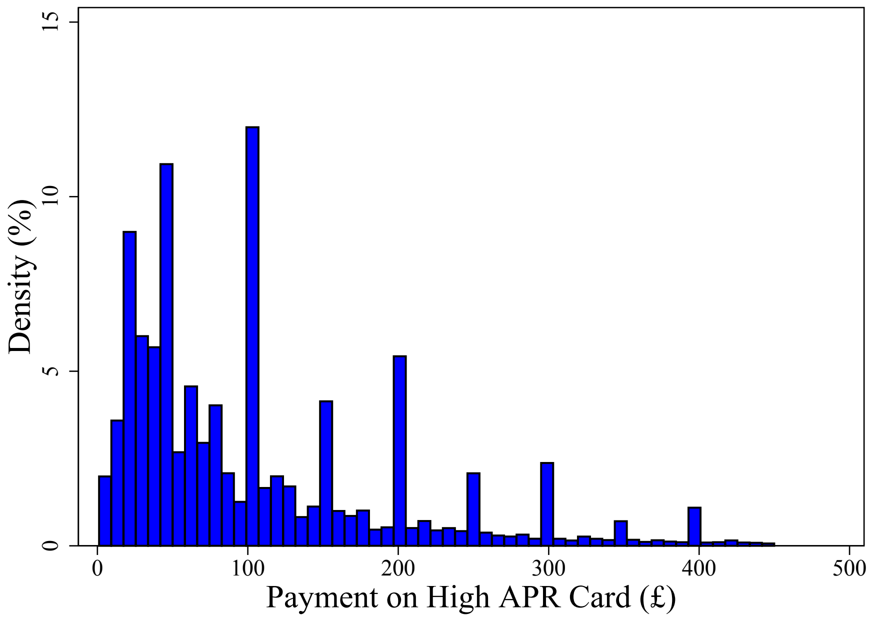
(C) Misallocated vs. Age of High-APR Card



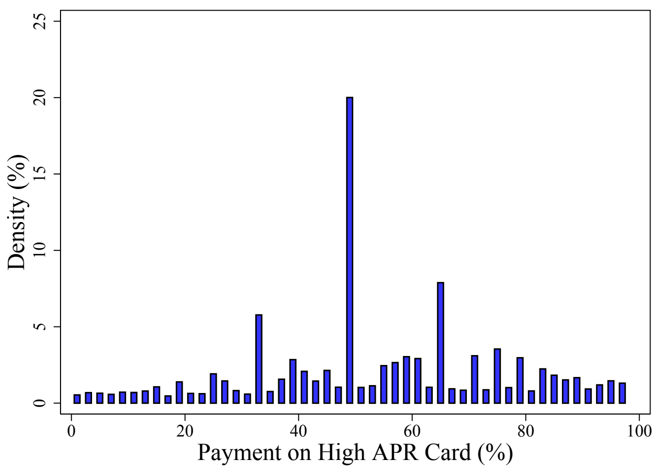
Note: Figure shows binned scatterplots (with 20 equally sized bins) of misallocated payments against the difference in annualized percentage interest rate (APR) across cards (Panel A) and the total value of payments within the month in pounds (Panel B), and the age of the high-APR card (Panel C). Local polynomial lines of best fit, based on the non-binned data, are also shown. Two-card sample restricted to individual × months in which individuals face an economically meaningful allocative decision. See Section 2 for details.

Figure 3: Rounding and the $1/n$ Rule

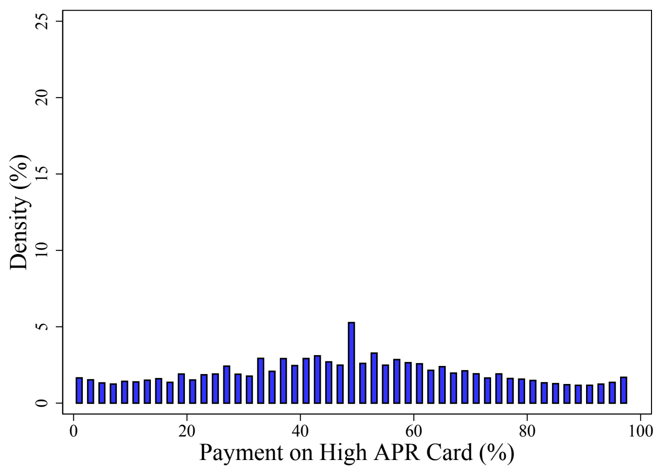
(A) Density of Payments (£s)



(B) Density of Payments (%), Round Number Values



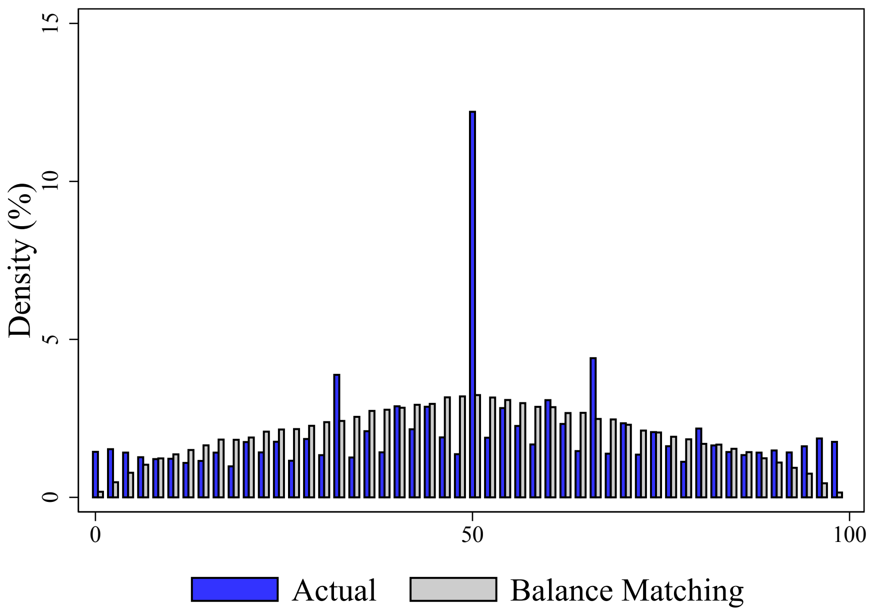
(C) Density of Payments (%), Non-Round Number Values



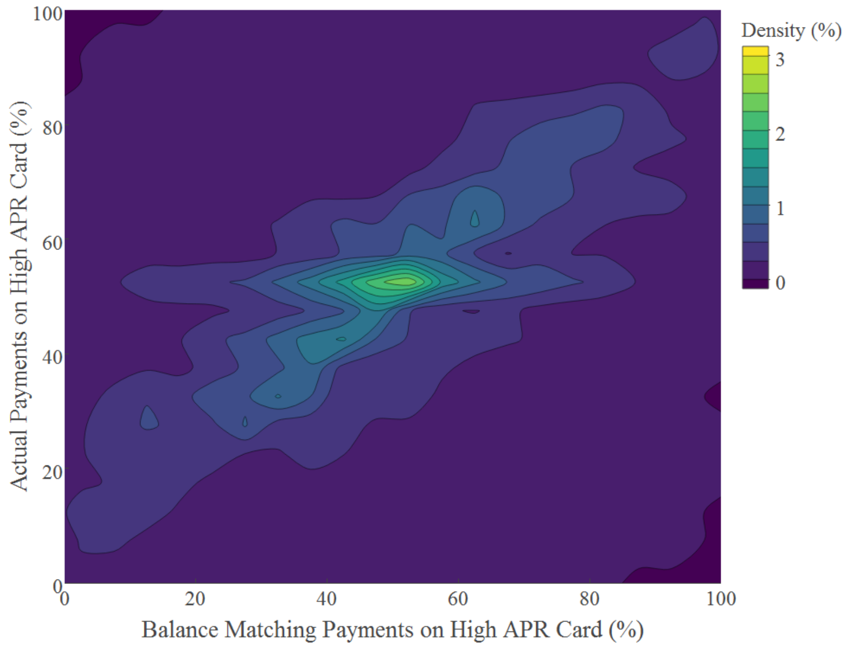
Note: Panel A shows the distribution of payments on the high-APR card in £s (excluding the top decile). Panel B plots the distribution of payments on the high-APR card in percent, among individuals who make round number payments (exact multiples of £50). Panel C plots the distribution of payments on the high-APR in percent, among individuals who do not make paments in exact multiples of £50. Two-card sample restricted to individual × months in which individuals face an economically meaningful allocative decision. See Section 2 for details.

Figure 4: Balance Matching

(A) Marginal Densities of Actual and Balance-Matching Payments



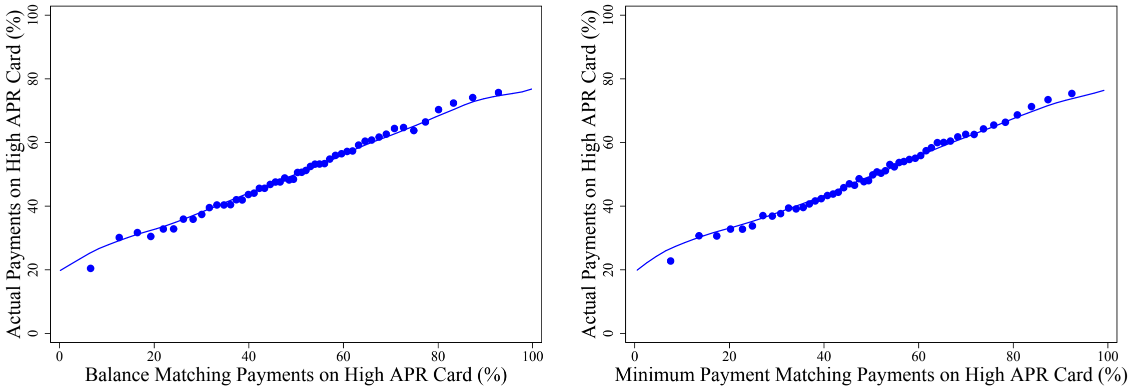
(B) Joint Densities of Actual vs. Balance-Matching Payments



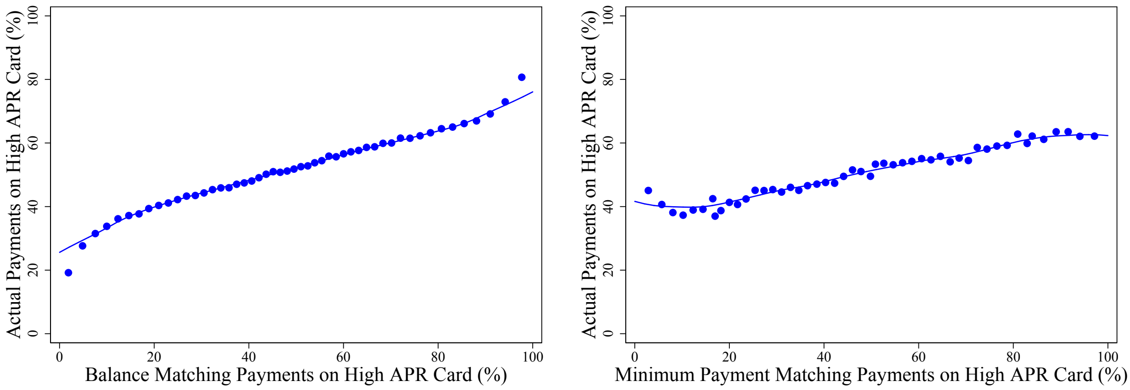
Note: Panel A shows the distribution of actual and balance-matching payments on the high APR card. Panel B plots the joint density of actual and balance-matching payments. Two-card sample restricted to individual × months in which individuals face an economically meaningful allocative decision. See Section 2 for details.

Figure 5: Balance Matching and Minimum Payment Matching in the Floor and Slope Samples

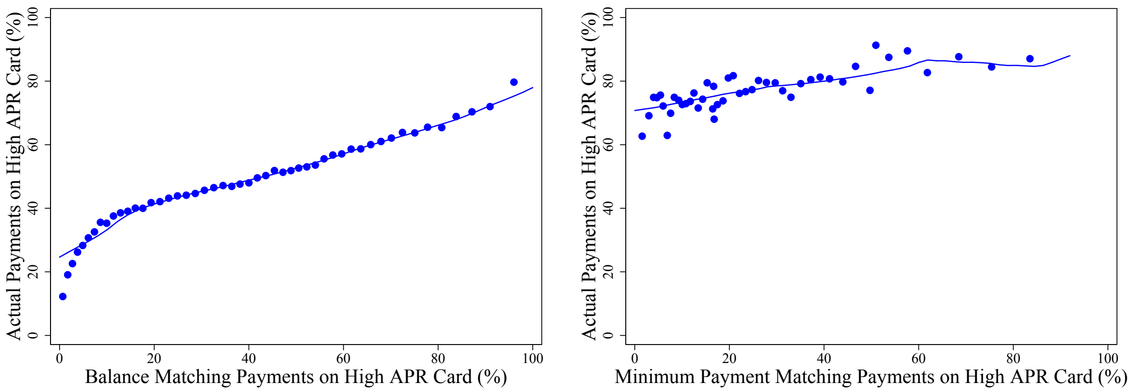
(A) Same Slope Sample



(B) Different Slope Sample



(C) Floor Sample



Note: Panels show binned scatterplots of the actual percentage of monthly payment allocated to the high APR card (y-axis) and the percentage of total monthly payment allocated to the high APR card under different rules (x-axis). “Same Slope” sample comprises account × months in which the minimum payment is determined by the percentage formula on both cards and the percentage is identical across cards.“Different Slope” sample comprises account × months in which the minimum payment is determined by the percentage formula on both cards and the percentage differs across cards “Floor” sample comprises account × months in which the minimum payment determined by the floor value on both cards held by the individual, e.g. £25.

Table 1: Summary Statistics

	(1) High APR card		(2) Low APR card		(3) Difference	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Card Characteristics						
APR: purchases (%)	22.86	4.80	16.56	6.40	6.30	5.85
APR: cash advances (%)	26.08	4.12	23.72	5.27	2.36	6.31
Monthly credit limit (£)	6,385.65	4,443.43	6,010.49	4,090.44	375.16	4,861.46
Spending (£)						
Purchases	128.04	432.21	116.88	399.07	11.15	570.92
Purchases if > £0	380.50	672.82	360.70	631.81	-3.15	798.61
Cash advances	6.44	73.42	5.84	73.99	0.60	97.66
Cash advances if > £0	216.44	368.41	215.82	396.14	-8.50	350.25
Payments (£)						
Repayments	259.85	735.98	230.14	660.25	29.72	916.19
Interest paid (£)						
Purchases	38.44	59.51	28.89	48.27	9.55	61.68
Cash advances	1.48	10.75	0.91	7.11	0.58	11.87
Card cycle (£)						
Closing balance	3,018.47	3,116.00	3,026.54	2,961.87	-8.07	3,479.06
Balance revolving	2,197.68	2,892.93	2,049.44	2,791.24	148.24	3,084.89
Minimum amount due	63.19	68.91	56.65	57.91	6.54	71.42
Card Status						
Account charge-off rate (%)	1.80	3.03	1.65	2.57	0.13	3.12
Tenure (months since account opened)	104.81	78.15	78.59	70.22	26.23	84.63
Number of account-months	394,061		394,061		394,061	

Note: Summary statistics for two-card analysis sample, defined as account × months in which the individual enters the account cycle with i) revolving debt on both cards, ii) pays at least the minimum on both cards, iii) pays more than the minimum on at least one card, and iv) ends the cycle with revolving debt on at least one card. Account charge-off rate is the predicted probability of the credit card charging-off within the next six months.

Table 2: Actual and Optimal Payments on the High APR Card

	Mean	Std. Dev.	*Percentiles* 10th	25th	50th	75th	90th
i) As % Total Monthly Payment							
Actual Payment (%)	51.21	24.21	16.85	33.33	50.00	67.94	84.77
Optimal Payment (%)	70.77	22.15	38.13	55.96	75.25	89.51	95.83
Difference (%)	19.56	23.76	0.00	0.72	9.98	32.50	54.64
ii) Payment in £							
Actual Payment (£)	259.85	735.98	25.00	45.64	100.00	200.00	450.00
Optimal Payment (£)	377.76	852.12	32.78	65.00	138.44	307.09	808.84
Difference (£)	117.91	422.71	0.00	1.00	17.99	75.00	238.51

Note: Summary statistics for actual and optimal payments on the high APR card. The top panel shows values as a percentage of total payments on both cards in that month. The bottom panel shows values in £. Two-card sample restricted to individual × months in which individuals face an economically meaningful allocative decision. See Section 2 for details.

Table 3: Goodness-of-Fit for Different Models

	RMSE	MAE	Correlation
i) Main Models			
Uniform Draw (0, 100)	36.52	29.99	0.00
Optimal	35.19	25.48	0.31
Balance Matching	23.83	17.02	0.47
ii) Alternative Heuristics			
Heuristic 1 (Pay Down Lowest Capacity)	36.43	27.27	0.08
Heuristic 2 (Pay Down Highest Capacity)	33.47	23.86	0.29
Heuristic 3 (Pay Down Highest Balance)	35.23	25.91	0.27
Heuristic 4 (Pay Down Lowest Balance)	34.09	24.59	0.10
iii) Machine Learning			
Decision Tree	19.35	14.97	0.53
Random Forest	16.22	11.56	0.71
XGBoost	17.15	12.90	0.66

Note: Goodness-of-fit for different models of the percentage of payments on the high-APR card. The first column shows the Root Mean Squared Error (RMSE), the second column shows that Mean Absolute Error (MAE), and third column shows the Pearson Correlation Coefficient, which can also be interpreted as the square root of the R-squared. Two-card sample restricted to individual × months in which individuals face an economically meaningful allocative decision. See Section 2 for details.

Table 4: Horse Races Between Alternative Models

Panel (A) Uniform vs. Other Rules

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Win %									
Uniform	33.57	45.26	50.00	44.72	46.98	46.49	38.07	31.28	34.41
Balance Matching	66.43								
Optimal		54.74							
Heuristic 1 (Pay Down Lowest Capacity)			50.00						
Heuristic 2 (Pay Down Highest Capacity)				55.28					
Heuristic 3 (Pay Down Highest Balance)					53.02				
Heuristic 4 (Pay Down Lowest Balance)						53.51			
Decision Tree							61.93		
Random Forest								68.72	
XGB									65.59

Panel (B) Balance Matching vs. Other Rules

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Win %								
Balance Matching	67.54	72.19	65.90	74.43	64.01	55.55	46.95	50.84
Optimal	32.46							
Heuristic 1 (Pay Down Lowest Capacity)		27.81						
Heuristic 2 (Pay Down Highest Capacity)			34.10					
Heuristic 3 (Pay Down Highest Balance)				25.57				
Heuristic 4 (Pay Down Lowest Balance)					35.99			
Decision Tree						44.45		
Random Forest							53.05	
XGB								49.16

Note: Table shows percentage of individual × month observations that are closest-fit in a horse race between payment rules. Target variable is the percentage of monthly repayment on the high APR card. Each column shows a pairwise horse race between rules, with cells reporting the percentage of observations closest-fit to the rule (shown by row). For individual × months in which the payment rule predicts a level of payment outside the feasible range, values are adjusted to minimum or maximum corners. Horse race uses two-card sample restricted to individual × months in which individuals face an economically meaningful allocative decision. See Section 2 for details

Table 5: Transition Matrix for Best-Fit Model

		Current Period							
		Uniform	Balance Matching	Optimal	H 1	H 2	H 3	H 4	1/n Rule
Previous Period	Uniform	21.92%	58.53%	0.26%	2.37%	1.00%	2.04%	1.10%	12.78%
	Balance Matching	8.39%	82.05%	0.34%	1.60%	1.55%	1.34%	0.94%	3.78%
	Optimal	8.19%	68.76%	11.62%	4.00%	0.76%	1.14%	0.95%	4.57%
	Heuristic 1 (Pay Down Lowest Capacity)	11.91%	50.01%	0.45%	23.39%	0.56%	3.13%	3.77%	6.78%
	Heuristic 2 (Pay Down Highest Capacity)	6.48%	62.31%	0.11%	0.68%	25.20%	3.10%	0.00%	2.12%
	Heuristic 3 (Pay Down Highest Balance)	11.91%	50.29%	0.59%	3.93%	4.21%	18.30%	0.61%	10.16%
	Heuristic 4 (Pay Down Lowest Balance)	10.46%	56.41%	0.46%	5.44%	0.04%	0.65%	22.63%	3.92%
	1/n Rule	15.87%	30.61%	0.23%	1.74%	0.51%	1.80%	0.62%	48.62%

Note: Table shows transition matrix for the best-fit payment model. Two-card sample restricted to individual × months in which individuals that face an economically meaningful allocative decision for at least two consecutive months. See Section 2 for details.

Online Appendix

I Machine Learning Models

This section provides details of machine learning models we use to fit repayment behaviors. We estimate decision tree, random forest and extreme gradient boosting. For all of these models, our target variable is the percentage of payments allocated to the high APR card in the two-card sample. We use APRs, balances, and credit limits on both cards as explanatory variables, and tune the models with cross-validation to maximize out-of-sample power.

Decision Tree Tree-based methods partition the sample space into a series of hyper-cubes, and then fit a simple model in each partition. The decision tree is grown through iteratively partitioning nodes into two sub-nodes according to a splitting rule. In our case, the splitting criterion is to find one explanatory variable as well as a cut-off value that minimize the sum of squared errors in the two sub-nodes combined. In theory, the tree can have one observation in each final node, but this tree will have poor performance out-of-sample. In practice, the decision tree is grown until the reduction in squared error falls under some threshold. Then, it calculates the average percentage of payments allocated to high APR cards in each final node.

We use the r package "rpart" to fit the decision-tree model.[29] To avoid overfitting the data, we "prune" the decision tree by tuning the complexity parameter through cross-validation. The complexity parameter requires each split to achieve a gain in R-squared greater than the parameter value. We pick the complexity parameter threshold that minimizes mean squared error in 10-fold cross-validation. That is, we split the sample randomly into 10 disjoint subsets. For each of these 10 subsets, we use the remaining 90% of the data to train the tree, and calculate the error on each 10% subset.[30]

Random Forest The machine learning literature has proposed several variations on the tree model. One approach which has been found to work very well in practice is random forest

[29] See `https://cran.r-project.org/web/packages/rpart/vignettes/longintro.pdf` for a complete description of the function.

[30] See Friedman et al. (2001) Chapter 9, for further information on tree-based methods.

(Breiman, 2001). Random forest builds a large number of trees and averages their predictions. It introduces randomness into the set of explanatory variables considered when splitting each node. The algorithm first draws a number of bootstrapped samples, and grows a decision tree within each sample. At each node, it randomly selects a subset of m explanatory variables in the split search, and chooses the best split among those m variables. Lastly, it makes predictions by averaging the results from each tree.

We use the r package "randomForest" to grow a forest of 100 trees.[31] For each tree, we calculate the out-of-sample error using the rest of the data not included in the bootstrapped sample. The average prediction error over these 100 trees is minimized to fine tune "m", the number of explanatory variables in the subset we consider in each split search. Increasing the number of trees does not significantly improve prediction accuracy.

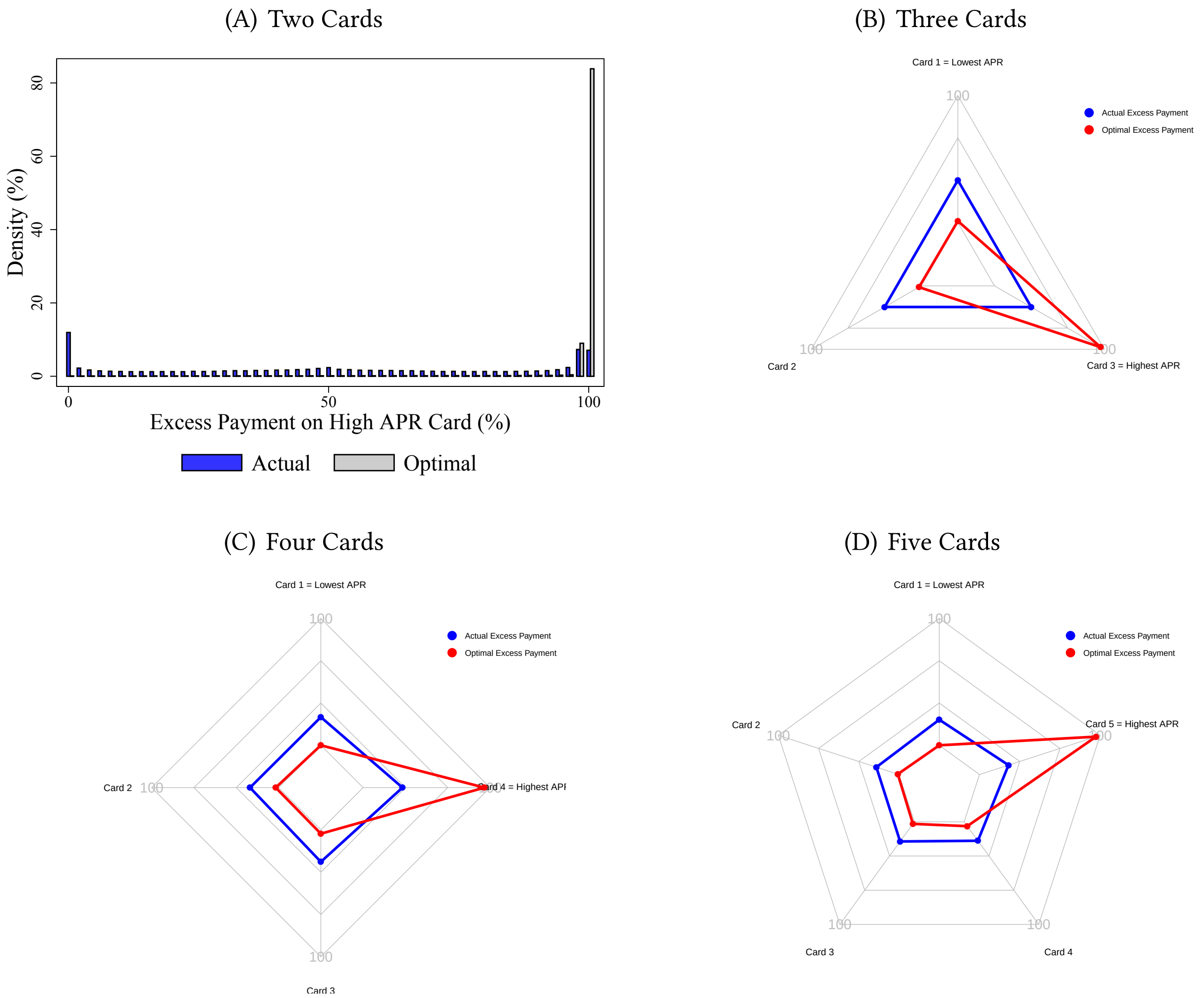
Extreme Gradient Boosting Extreme gradient boosting and random forest are both based on a collection of tree predictors. They differ in their training algorithm. The motivation for boosting is a procedure that combines the outputs of many "weak" classifiers to produce a powerful "committee" (Friedman et al., 2001). Instead of growing a number of trees independently, boosting applies an additive training strategy, by adding one new tree at a time. At each step, the new decision tree puts greater weights on observations that are misclassified in the previous iteration. Finally, it averages predictions from trees at each step. This algorithm effectively gives greater influence to the more accurate tree models in the additive sequence. We use the r package "xgboost" and fine tune the number of iterations over a 10-fold cross-validation.[32] The rest of parameters such as the learning rate are kept at their default values. Perturbation of these values does not have material impact on out-of-sample errors.[33]

[31] See `https://cran.r-project.org/web/packages/randomForest/randomForest.pdf` for a complete description of the function.

[32] See `http://cran.fhcrc.org/web/packages/xgboost/vignettes/xgboost.pdf` for a complete description of the function.

[33] For a more detailed introduction of extreme gradient boosting, see `http://xgboost.readthedocs.io/en/latest/model.html`. Friedman (2001) is the first paper that introduced the term "gradient boosting". Friedman et al. (2001), Chapter 10 also introduces a boosting algorithm.

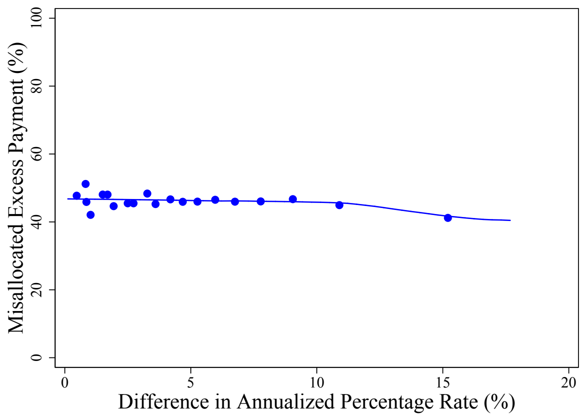
Figure A1: Actual and Optimal *Excess* Payments



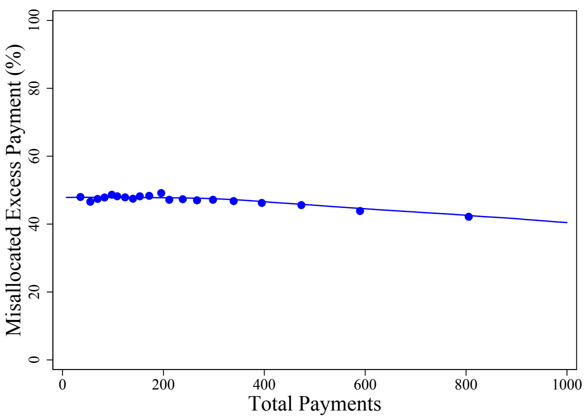
Note: Panel A shows the distribution of actual and optimal excess payments on the high interest rate card in the two-card sample. Panels B to D show radar plots of mean actual and optimal excess payments in the samples of individuals with 3 to 5 cards. Excess payments are calculated as the percentage of payments on a given card after subtracting out repayments needed to pay the minimum amounts due. In the radar plots, cards are ordered clockwise from highest to lowest APR (starting at the first node clockwise from 12 o'clock). Samples restricted to individual × months in which individuals face an economically meaningful allocative decision. See Section 2 for details.

Figure A2: Misallocated *Excess* Payments by Economics Stakes

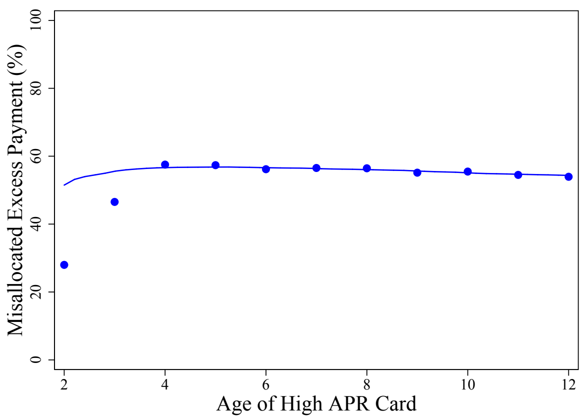
(A) Misallocated vs. Difference in APR



(B) Misallocated vs. Total Payments

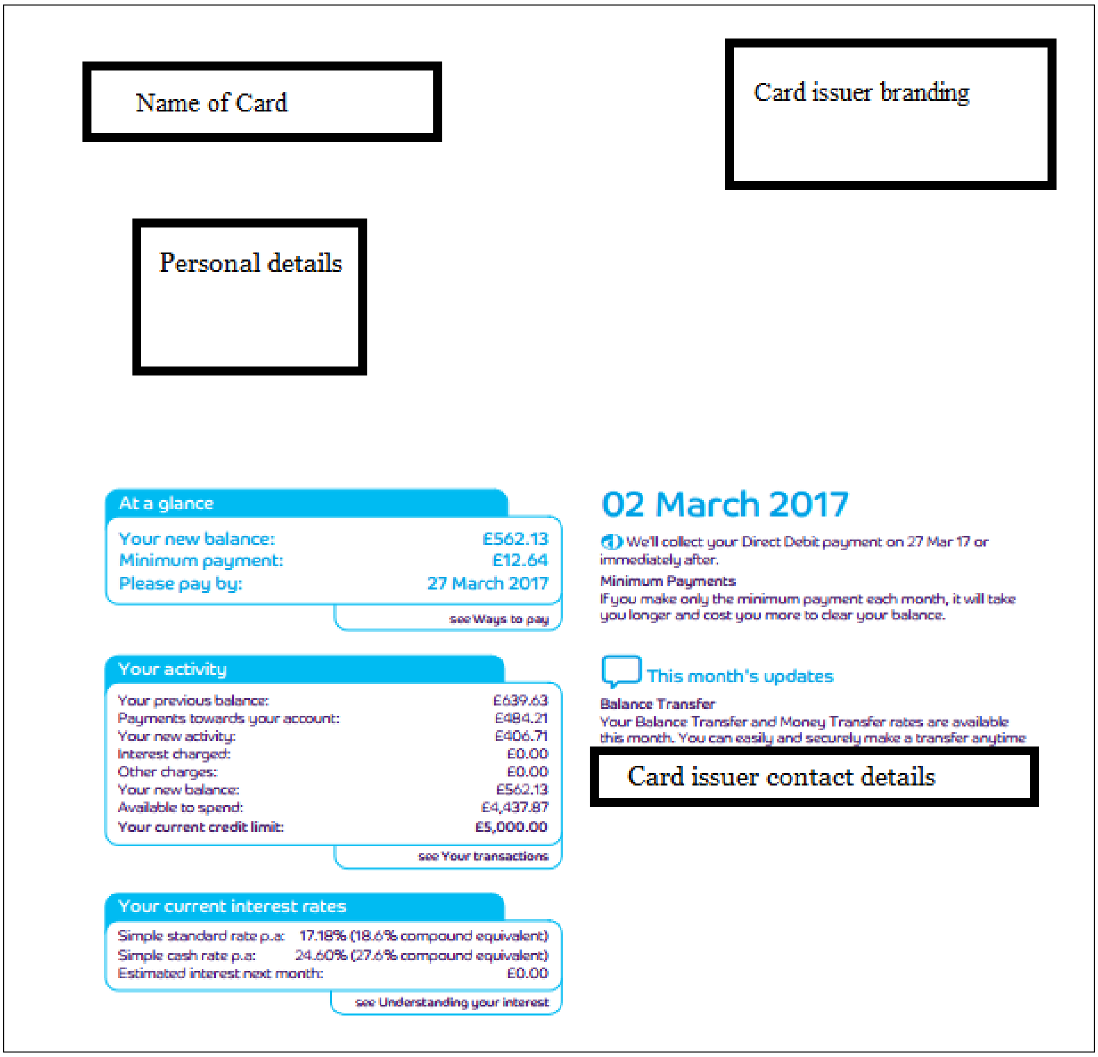


(C) Misallocated vs. Age of High Cost Card



Note: Figure shows binned scatterplots (with 20 equally sized bins) of misallocated payments against the difference in annualized percentage interest rates across cards (Panel A) and the total value of payments within the month in pounds (Panel B), and the age of the high-cost card (Panel C). Local polynomial lines of best fit, based on the non-binned data, are also shown. Excess payments are calculated as the percentage of payments on a given card after subtracting out repayments needed to pay the minimum amounts due. Two-card sample restricted to individual × months in which individuals face an economically meaningful allocative decision. See Section 2 for details.

Figure A3: Example Credit Card Statement



Name of Card

Card issuer branding

Personal details

At a glance

Your new balance:	£562.13
Minimum payment:	£12.64
Please pay by:	27 March 2017

see Ways to pay

Your activity

Your previous balance:	£639.63
Payments towards your account:	£484.21
Your new activity:	£406.71
Interest charged:	£0.00
Other charges:	£0.00
Your new balance:	£562.13
Available to spend:	£4,437.87
Your current credit limit:	**£5,000.00**

see Your transactions

Your current interest rates

Simple standard rate p.a:	17.18% (18.6% compound equivalent)
Simple cash rate p.a:	24.60% (27.6% compound equivalent)
Estimated interest next month:	£0.00

see Understanding your interest

02 March 2017

We'll collect your Direct Debit payment on 27 Mar 17 or immediately after.

Minimum Payments

If you make only the minimum payment each month, it will take you longer and cost you more to clear your balance.

This month's updates

Balance Transfer

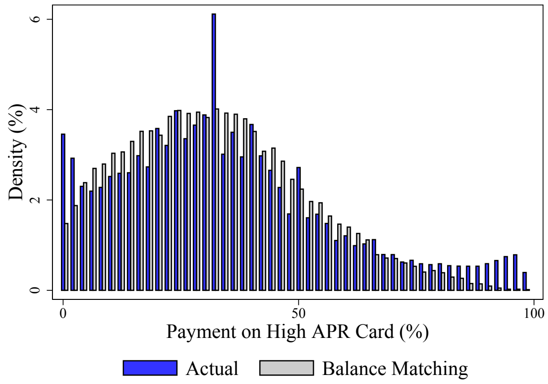
Your Balance Transfer and Money Transfer rates are available this month. You can easily and securely make a transfer anytime

Card issuer contact details

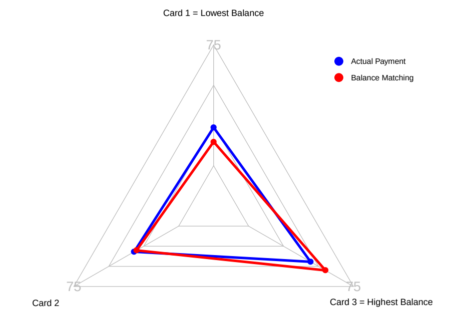
Note: The figure shows an extract of one of the author's credit card statements, with card issuer branding, contact details and card holder personal identifying information obscured.

Figure A4: Distribution of Actual and Balance-Matching Payments on Multiple Cards

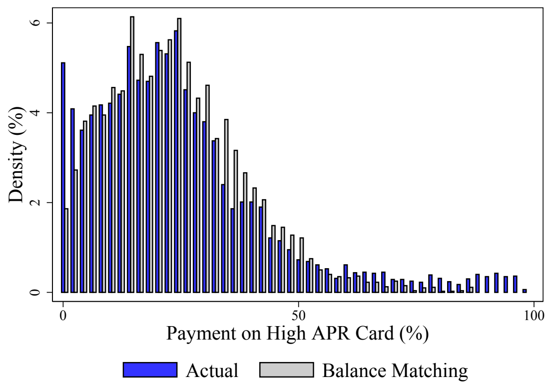
(A) Three Cards



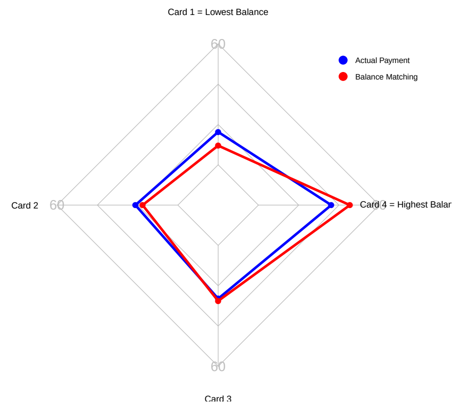
(B) Three Cards



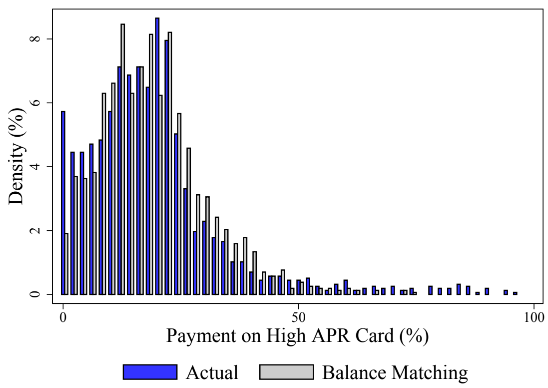
(C) Four Cards



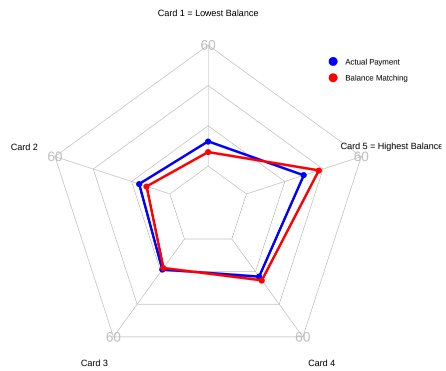
(D) Four Cards



(E) Five Cards



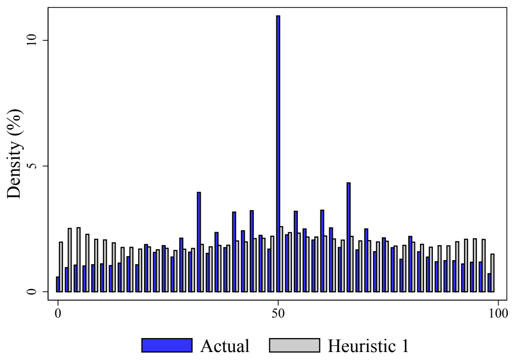
(F) Five Cards



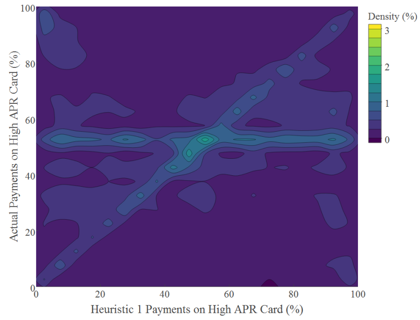
Note: Left column shows the marginal distributions of actual and balance-matching payments on the high APR card. Right column shows radar plots of the mean percentage of actual and balance-matching payments allocated to each card. In the radar plots, cards are ordered clockwise from highest to lowest balance (starting at the first node clockwise from 12 o'clock). Sample is restricted to individual × months in which individuals face an economically meaningful allocative decision. See Section 2 for details.

Figure A5: Payments Under Different Heuristics

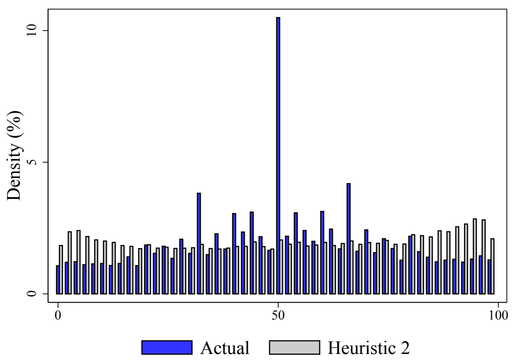
(A) Heuristic 1: Pay Down Lowest Capacity



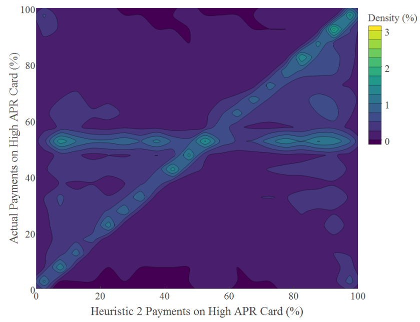
(B) Heuristic 1 Joint Density



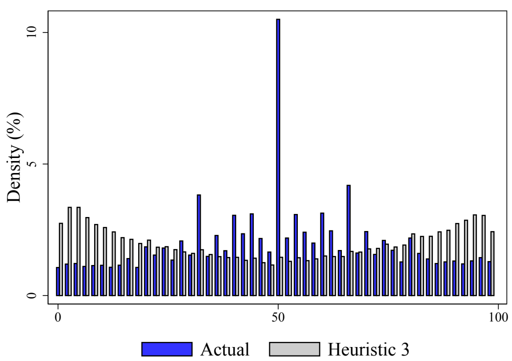
(C) Heuristic 2: Pay Down Highest Capacity



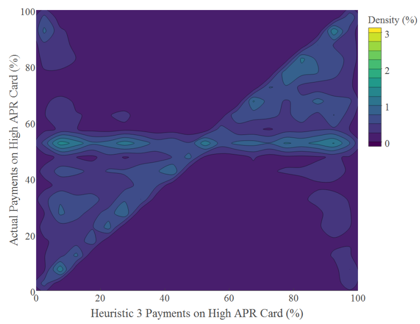
(D) Heuristic 2 Joint Density



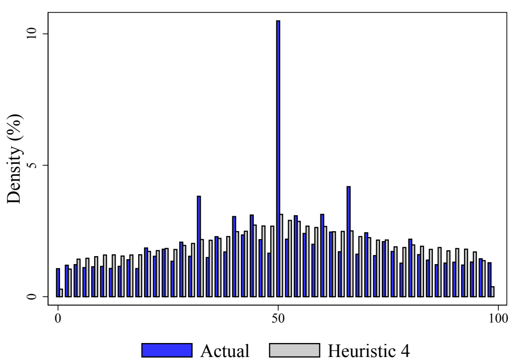
(E) Heuristic 3: Pay Down Highest Balance



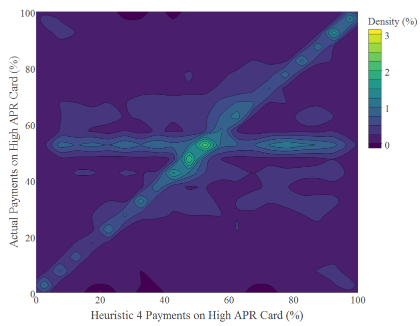
(F) Heuristic 3 Joint Density



(G) Heuristic 4: Pay Down Lowest Balance

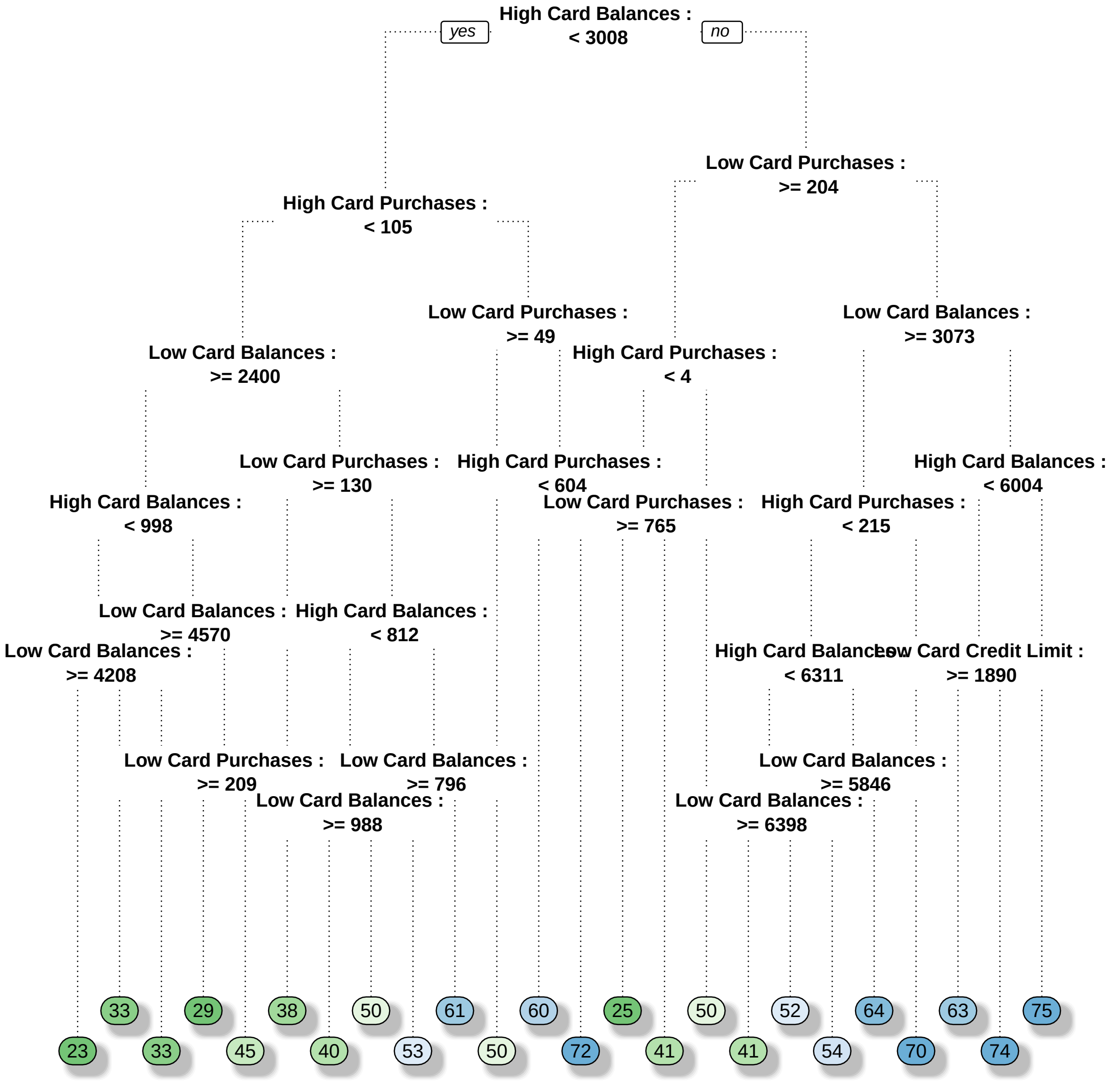


(H) Heuristic 4 Joint Density



Note: Left column shows the marginal distributions actual payments and payments under the heuristics. Right column shows the joint density. See Section 5 for definitions of the heuristics. Two-card sample restricted to individual × months in which individuals face an economically meaningful allocative decision. See Section 2 for details.

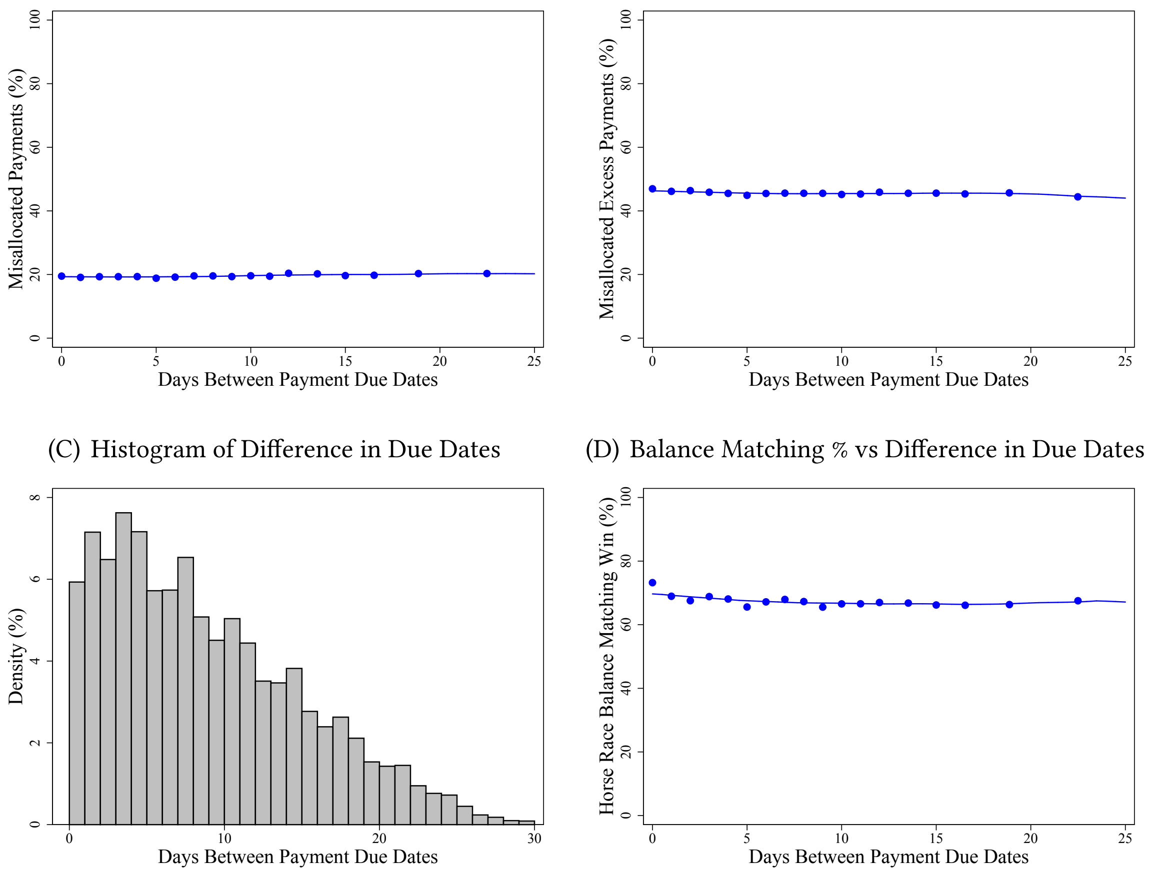
Figure A6: High APR Card Payment Decision Tree



Note: Figure shows the decision (regression) tree for high APR card repayment. Top row is tree root. Nodes show variable and split value at each branch. Bottom rows show predicted values at end of each branch.

Figure A7: Misallocated Payments and Balance Matching by Difference in Due Dates

(A) Misallocated Payments vs. Diff. Due Dates (B) Excess Misallocated Payments vs. Diff. Due Dates



Note: Panels A shows a binned-scatter plot of misallocated payments against the absolute difference in due dates of each of the credit cards. Panel B shows a binned-scatter of misallocated excess payments against the absolute difference in due dates. Panel C shows the distribution of the absolute difference in due dates. Panel D shows a binned-scatter plot of the percentage of observations that are best fit by balance-matching payments (instead of optimal payments) against the absolute difference in due dates. Local polynomial lines of best fit, based on the non-binned data, are also shown in the binned-scatter plots. Excess payments are calculated as the percentage of payments on a given card after subtracting out repayments needed to pay the minimum amounts due. Two-card sample restricted to individual × months in which individuals face an economically meaningful allocative decision. See Section 2 for details.

Table A1: 2-Card Sample Selection

	Account X months		*Customers*		*Aggregate debt*	
	Count	*%*	*Count*	*%*	*Count*	*%*
Unrestricted Sample	7,876,760	100%	229,260	100%	330,495,722	100%
Drop if Equal Interest Rates	315,070	4%	2,293	1%	6,609,914	2%
Drop if No Debt on Either Card	3,544,542	45%	71,071	31%	0	0%
Drop if Debt on Only One Card	1,260,282	16%	16,048	7%	49,574,358	15%
of which:						
Debt on High Card Only	708,908	9%	9,170	4%	29,744,615	9%
Debt on Low Card Only	551,373	7%	6,878	3%	19,829,743	6%
Drop if Pays Min Only	1,654,120	21%	25,219	11%	39,659,487	12%
Drop if Pays Both in Full	315,070	4%	2,293	1%	9,914,872	3%
Restricted Sample	788,122	10%	112,796	49%	226,059,074	68%

Note: Table shows number and percent of observations for unrestricted sample (all two-card customer months in the Argus data), restricted sample and data dropped under each sample selection rule.

Table A2: Actual and Optimal *Excess* Payments on the High APR Card

			Percentiles				
	Mean	Std. Dev.	10th	25th	50th	75th	90th
i) As a % Total Monthly Payment							
Actual Excess Payment (%)	51.46	34.74	0.88	19.91	51.23	84.80	99.82
Optimal Excess Payment (%)	97.08	12.93	100.00	100.00	100.00	100.00	100.00
Difference (%)	45.61	35.04	0.00	11.48	45.44	75.71	98.41
ii) Payments in £							
Actual Excess Payment (£)	196.66	731.51	0.23	2.32	22.78	88.79	351.00
Optimal Excess Payment (£)	314.57	845.95	1.92	14.48	66.65	223.30	739.02
Difference (£)	117.91	422.71	0.00	1.00	17.99	75.00	238.51

Note: Summary statistics for actual and optimal excess payments on the high APR card. Excess payments are calculated as the percentage of payments on a given card after subtracting out repayments needed to pay the minimum amounts due. The top panel shows values as a percentage of total excess payments on both cards in that month. The bottom panel shows values in £. Two-card sample restricted to individual × months in which individuals face an economically meaningful allocative decision. See Section 2 for details.

Table A3: High APR Card Actual, Optimal and Balance Matching Payments, Multiple Cards

	Mean	Std. Dev.	*Percentiles* 10th	25th	50th	75th	90th
i) 3 Cards							
Actual Payments (%)	35.05	21.49	8.70	19.71	33.25	46.67	64.86
Optimal Payments (%)	62.76	24.34	28.14	44.60	65.22	83.39	93.58
Balance Matching Payments (%)	32.79	19.77	7.75	17.33	31.13	45.47	59.86
i) 4 Cards							
Actual Payments (%)	26.97	18.67	6.08	14.06	23.64	35.71	50.65
Optimal Payments (%)	56.55	24.15	23.11	37.93	57.46	75.84	89.27
Balance Matching Payments (%)	25.32	16.44	5.33	13.32	23.30	34.51	47.05
i) 5 Cards							
Actual Payments (%)	21.23	16.87	4.66	10.04	17.78	27.12	40.63
Optimal Payments (%)	51.99	23.76	19.25	33.23	53.13	70.30	84.48
Balance Matching Payments (%)	18.41	12.51	3.12	8.92	17.22	25.18	34.39

Note: Summary statistics for actual payments, optimal, and balance-matching payments on the high-APR card in the samples with 3 to 5 cards. Samples are restricted to individual × months in which individuals face an economically meaningful allocative decision. See Section 2 for details.

Table A4: High APR Card Payments Under Heuristics Models, Two Cards

	Mean	Std. Dev.	*Percentiles* 10th	25th	50th	75th	90th
i) Heuristic 1 (Pay Down Lowest Capacity Card)							
Heuristic 1 Rule Payment (£)	202.79	500.38	14.65	30.14	75.00	174.61	390.06
Actual – Heuristic 1 Rule Payment (£)	49.08	512.93	−93.79	−20.52	0.00	32.01	175.15
Heuristic 1 Rule Payment (%)	48.91	29.29	7.38	23.16	49.98	73.27	89.70
Actual – Heuristic 1 Rule Payment (%)	2.81	33.36	−37.50	−12.20	0.00	17.48	46.93
ii) Heuristic 2 (Pay Down Highest Capacity Card)							
Heuristic 2 Rule Payment (£)	260.13	704.26	17.00	33.38	82.59	199.21	500.00
Actual – Heuristic 2 Rule Payment (£)	−0.26	439.71	−101.00	−20.60	0.00	15.00	88.91
Heuristic 2 Rule Payment (%)	51.97	30.01	9.13	25.36	53.18	79.28	92.31
Actual – Heuristic 2 Rule Payment (%)	−0.76	29.32	−37.50	−10.62	0.00	8.07	35.04
iii) Heuristic 3 (Pay Down Highest Balance Card)							
Heuristic 3 Rule Payment (£)	246.99	659.84	11.78	25.18	73.11	195.65	500.00
Actual – Heuristic 3 Rule Payment (£)	12.58	425.02	−102.33	−20.00	0.00	25.00	131.89
Heuristic 3 Rule Payment (%)	49.67	32.28	6.33	18.52	49.87	80.65	93.00
Actual – Heuristic 3 Rule Payment (%)	1.55	31.16	−36.22	−10.48	0.00	14.31	40.81
iv) Heuristic 4 (Pay Down Lowest Balance Card)							
Heuristic 4 Rule Payment (£)	250.28	631.74	25.00	42.83	90.31	195.00	471.92
Actual – Heuristic 4 Rule Payment (£)	9.59	463.11	−104.00	−19.56	0.00	18.27	107.13
Heuristic 4 Rule Payment (%)	51.64	25.33	15.36	32.24	52.15	71.50	86.80
Actual – Heuristic 4 Rule Payment (%)	−0.44	29.73	−38.90	−10.71	0.00	9.15	37.50

Note: Summary statistics for actual payments on the high interest card as a percentage of total monthly repayments and payments under heuristic rules.

Table A5: Machine Learning Models Variable Importance

Random Forest		Gradient Boost	
Variable	Importance	Variable	Importance
High Card Balance	0.25	High Card Balance	0.26
Low Card Balance	0.21	Low Card Balance	0.25
Low Card CL	0.16	Low Card Purchases	0.17
High Card CL	0.10	High Card Purchases	0.17
High Card Purchases	0.11	Low Card CL	0.05
Low Card Purchases	0.10	High Card CL	0.04
High Card APR	0.04	High Card APR	0.03
Low Card APR	0.03	Low Card APR	0.02

Note: Table summarizes the importance of input variables in explaining high card repayments in random forest and extreme gradient boosting models. Rows show the proportion of the total reduction in sum of squared errors in the outcome variable resulting from the split of each variable across all nodes and all trees.

Table A6: Goodness of Fit Measures for Machine Learning Models

	RMSE	MAE	Correlation
i) Decision Tree			
Without APR	19.35	14.97	0.53
With APR	19.35	14.97	0.53
ii) Random Forest			
Without APR	17.09	12.49	0.67
With APR	16.22	11.56	0.71
iii) Gradient Boosting			
Without APR	17.52	13.21	0.64
With APR	17.15	12.90	0.66

Note: Table shows the godness of fit for the machine learning algorithms with and without the high and low card APR variables included in the dataset fed into the machine. The target variable is the fraction of payment to the card with higher APR. We show Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and Pearson Correlation Coefficient in each of the three columns.

Table A7: Correlation Matrix Machine Learning Model Input Variables

	APR(H)	APR(L)	Bal(H)	Bal(L)	Pur(H)	Pur(L)	Lim(H)	Lim(L)
APR(H)	1							
APR(L)	0.495***	1						
Bal(H)	0.108***	0.135***	1					
Bal(L)	0.101***	0.0876***	0.370***	1				
Pur(H)	-0.0630***	-0.0563***	0.0437***	0.0718***	1			
Pur(L)	-0.0496***	-0.0264***	0.0926***	0.0481***	0.00744	1		
Lim(H)	-0.0466***	0.0341***	0.646***	0.240***	0.151***	0.101***	1	
Lim(L)	-0.0455***	0.0503***	0.268***	0.690***	0.0744***	0.124***	0.373***	1

Note: Table shows correlation matrix (Pearson correlation coefficients) for input variables to the machine learning models.

Table A8: Horse Race by Minimum Payment

	"Floor" Sample	"Slope" Sample
Optimal	30.95	35.15
Balance Matching	69.05	64.85

Note: Table shows percentage of individual × month observations that are closest-fit in a horse race between payment rules. Target variable is the percentage of monthly repayment on the high APR card. Each column shows a pairwise horse race between rules, with cells reporting the percentage of observations closest-fit to the rule (shown by row). For individual × months in which the payment rule predicts a level of payment outside the feasible range, values are adjusted to minimum or maximum corners. "Floor" minimum sample comprises account × months in which the minimum payment determined by the floor value on both cards held by the individual, e.g. £25. "Slope" minimum sample comprises account × months in which the minimum payment is determined by the percentage formula on both cards.

Table A9: Correlations Between Payment Rules

	Panel (A) Balance Matching vs. Min. Pay Matching		
	(1)	(2)	(3)
	Same Slopes	Different Slopes	Floor
Correlation	0.96	0.86	0.56
	(0.00)	(0.00)	(0.02)
	Panel (B) Balance Matching vs. Actual		
	(1)	(2)	(3)
	Same Slopes	Different Slopes	Floor
Correlation	0.63	0.53	0.54
	(0.00)	(0.00)	(0.00)
	Panel (C) Min. Pay Matching vs. Actual		
	(1)	(2)	(3)
	Same Slopes	Different Slopes	Floor
Correlation	0.62	0.29	0.22
	(0.00)	(0.01)	(0.02)

Note: Table shows correlation coefficients (standard errors in parenthesis) between balance matching payments minimum payment matching payments and actual payments for the high APR expressed as a percentage of the total monthly repayment. "Same Slope" sample comprises account × months in which the minimum payment is determined by the percentage formula on both cards and the percentage is identical across cards."Different Slope" sample comprises account × months in which the minimum payment is determined by the percentage formula on both cards and the percentage differs across cards "Floor" sample comprises account × months in which the minimum payment determined by the floor value on both cards held by the individual, e.g. £25.