

Selecting the Optimal Credit Card Portfolio

ECO 6936

Capstone in Business Analytics II

Remco A. Scheepmaker

July 24, 2024

1 Introduction

Credit cards are an important part of the American economy and culture. According to the Consumer Financial Protection Bureau (CFPB, 2023), consumers spent \$3.2 trillion on purchases using credit cards in 2022. Total credit card debt recently passed \$1 trillion, 82 percent of which is revolving (i.e. bearing interest, the remaining 18 percent is paid by the due date without being charged interest). The combination of income from interest, interchange fees, and annual fees makes credit cards very profitable for the major banks, but they can also be profitable for a certain fraction of their users. A key feature of many credit cards is their rewards structure, which returns a percentage of the user's spend back to the user in the form of cash back, reward points, or miles.¹ Credit card rewards programs are designed by banks to attract new customers and improve their loyalty, as well as to increase income through interchange fees and interest by stimulating (over)spending. The CFPB estimates that the total dollar value of the rewards earned in 2022 exceeded \$40 billion, and the average rewards-earning account redeemed \$167.² These "earnings" are much less than the \$130 billion charged to consumers in interest and fees, making credit cards such a profitable enterprise for banks.

¹I consider points and miles to be the same for this project. Both can generally be redeemed for cash, statement credit, or used to book travel, either by transferring to loyalty programs of travel partners, or by booking flights and hotels directly through the bank's own travel portal.

²According to Experian, the average American has 3.9 credit card accounts (<https://www.experian.com/blogs/ask-experian/average-number-of-credit-cards-a-person-has/>, accessed June 1, 2024).

As we will see in more detail in the next section, credit card rewards work partially through what is called the “reverse Robin Hood” mechanism, since to some extent it is the poor who, by paying interest and fees, subsidize the rewards of the rich ([The Wall Street Journal, 2010](#)). More accurately, it is primarily the financially naïve who are sponsoring the financially sophisticated, where the level of sophistication seems best measured by credit scores from the Fair Isaac Corporation (FICO), as opposed to income ([Agarwal, Presbitero, Silva, and Wix, 2023](#)). For those interested in travel using credit card points, or just saving some money through cashback rewards, it is therefore crucial to learn good financial habits and eventually become part of the financially sophisticated demographic with high FICO scores.

Once credit card rewards are a net benefit (meaning the value of the rewards are higher than the costs of fees and interest), it might be worthwhile to explore how we can optimize one’s choice of credit cards to maximize this benefit. Which credit cards should different consumers select, and is there an optimal number before the marginal benefit of adding more cards becomes too small? These are the topics I study in this project. My personal credit card portfolio has expanded from two to eleven credit cards over the last five years, increasing my net benefit from less than two to about five percent of my entire annual spend. The analysis that I applied to my own spending budget, in order to optimize my own credit card portfolio, inspired me to create an optimization algorithm that can be applied at scale to a larger selection of credit cards, and to different types of consumers. In this paper I present this algorithm, code it in R, and use it to model “optimal” portfolios of rewards credit cards for various combinations of user preferences and spending characteristics. I then study the properties of these modeled portfolios by combining data from the Bureau of Labor Statistics (BLS) on income and spending with data on credit card reward multipliers and values. By performing a sensitivity analysis, I find that spending on four to six credit cards is the sweet spot for most people, before the marginal benefit of adding more cards to the portfolio drops below \$50. From a Monte Carlo simulation I find that the average return on spend (ROS) is 3.8 percent, with a standard deviation of 0.8 percent. In terms of ROS, the benefit of using credit cards can be significantly higher (up to ~6.5 percent) for people who do not spend a lot but make full use of static benefits, such as travel credits and airport lounge access, or for people who spend above average on travel, since the travel categories have the highest (uncapped) point multipliers. Finally, I present a Shiny app that I developed in R, to serve as an

online tool that people can use to get a recommended credit card portfolio, based on their personal preferences and spending budget. This app can benefit many sophisticated credit card users who would like to stretch their budgets, travel more, or who simply enjoy optimizing their personal finances.

The structure of this paper is as follows. In Sect. 2, I first discuss some literature and background on credit scores and the payment structure of rewards programs. In Sect. 3, the theoretical model that will form the basis of my optimization algorithm is presented. In Sect. 4, the theory is broken down into an empirical specification that depends on variables and parameters that will need to come from data. In Sect. 5, the sources for these data are described. In Sect. 6, I present the results, which includes the sensitivity analysis and Monte Carlo modeling, as well as a description of the R Shiny app that has also been deployed online. Finally, the conclusions are discussed in Sect. 7, and the appendices are reserved for some tables that show additional details of the data, as well as the code that makes up the recommendation algorithm.

2 Literature

The latest “Consumer Credit Card Market” report from the [CFPB \(2023\)](#) contains a wealth of information about how Americans use their credit cards, separated by FICO credit scores according to the levels shown in Table 1. The CFPB reports that more than 90 percent of the purchase volume is on rewards cards, and that the average rewards-earning account contained \$156 in 2022. But, as we can see in Fig. 1, accounts with high credit scores have considerably higher rewards balances compared to accounts with low credit scores. The CFPB also finds that people with prime plus or superprime credit scores keep zero, or very low, revolving balances, while their purchase volumes and payment rates are actually the highest. This is consistent with practicing good financial habits in order to enjoy the benefits of credit card rewards and reach a high FICO score.³ The most important habit regarding rewards is to pay off statement balances *in full* by the due date each month (so no interest on revolving balances is paid). This implies only using credit cards for purchases that you can afford to pay off completely the next month.

While it is true that applying for new credit cards results in the banks “pulling” the applicant’s

³Although the exact algorithm is proprietary, there are many online resources that explain what a FICO score consists of, and how one can improve it, e.g. <https://www.myfico.com/credit-education/credit-scores>.

	CFPB	Agarwal <i>et al.</i> (2023)
deep sub-prime	< 580	
sub-prime	580–619	< 660
near-prime	620–659	660–719
prime	660–719	720–779
prime plus	720–799	
super-prime	> 799	> 779

Table 1: FICO score thresholds used by the CFPB and Agarwal *et al.* (2023).

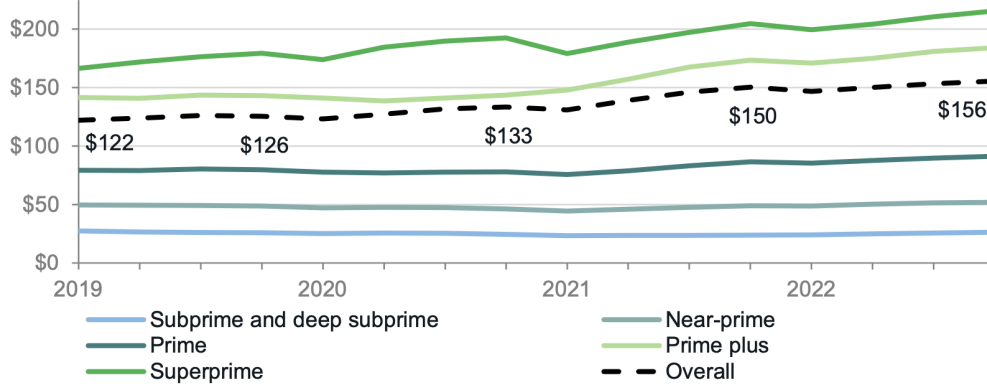


Figure 1: Quarterly dollar value of average rewards balances for different credit scores. Source: CFPB (2023).

credit report, lowering the credit score, it is important to note that this effect is only temporary. In the long run, opening many credit cards to optimize rewards can actually have a positive impact on credit scores, as the “utilization ratio” per card is lowered (the amount of spending on the card as a percentage of the credit limit). As an example, Fig. 2 shows a time series of my personal credit scores with the dates of nine credit card applications marked. Even with this above-average number of credit card applications, the overall trend in credit scores is positive.

Regarding the payment structure of rewards programs, Fumiko Hayashi (2009) shows a good overview of the payment fee flows between the merchant, the merchant acquirer (the bank that processes card payments for the merchant), the card issuer (the cardholder’s bank) and the cardholder. In an example where the cardholder makes a \$100 purchase, the merchant receives \$97.50. The \$2.50 in fees is split between the merchant acquirer (\$0.50) and an interchange fee of \$2.00 that is paid from the merchant acquirer to the card issuer. The card issuer pays a fraction of the interchange fee (say \$1.00) as a reward to the cardholder, and also pays the card network (such as

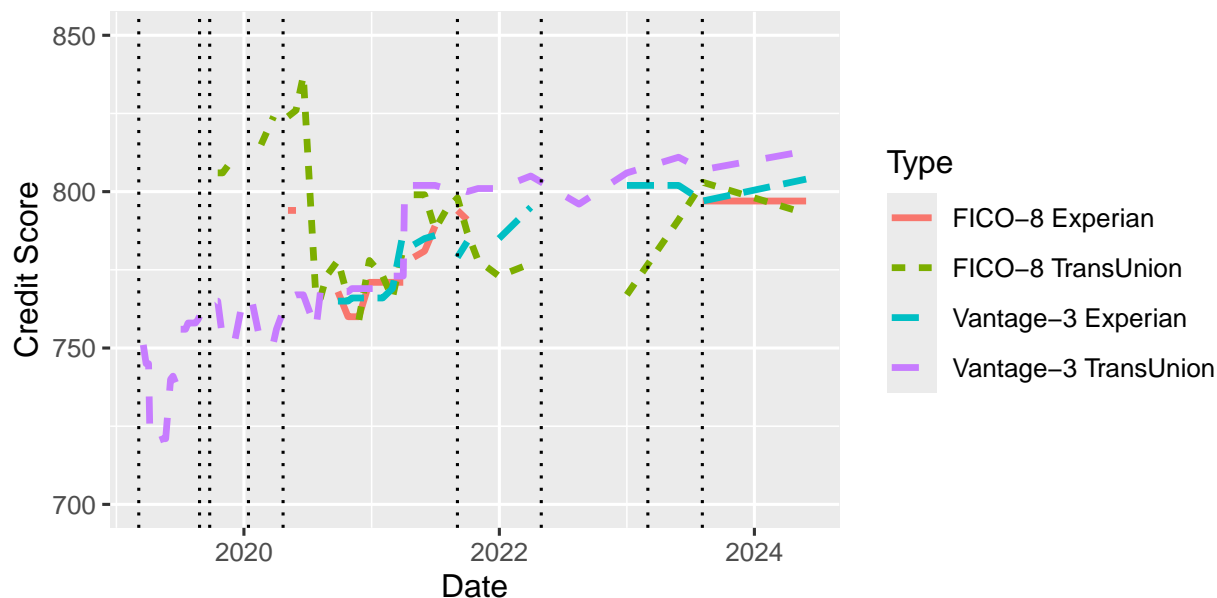


Figure 2: Time series of my personal credit scores. The vertical dotted lines mark the dates when new credit card accounts were opened.

Visa or MasterCard), who set the interchange fees.⁴ The merchant is likely to pass on these fees as higher retail prices to the consumer (regardless of the payment method). Hayashi (2009) concludes that it is not completely clear who ultimately pays for the rewards programs, but it seems likely that credit card rewards programs are not efficient markets. For rewards programs to be Pareto efficient, the use of credit cards would have to lead to cost savings for the merchants (due to less handling of cash, which comes with security risks and depositing costs). Only the savings beyond the operating costs of handling credit card transactions would then be returned to customers in the form of rewards, in which case everyone is better off without anyone being worse off. Such a scenario seems unlikely when many merchants complain about interchange fees being too high.

Schuh, Shy, and Stavins (2010) find that it is cash-using households who are paying for the rewards of card-using households, which, due to the positive correlation between income and use of credit cards, translates to lower incomes paying for the rewards of higher incomes (the “reverse Robin Hood” mechanism). However, Agarwal, Presbitero, Silva, and Wix (2023) show that this redistribution takes place from low to high FICO scores *regardless* of income. As Fig. 3 shows,

⁴American Express transactions work differently, as American Express operates its own network and simultaneously acts as both card issuer and merchant acquirer.

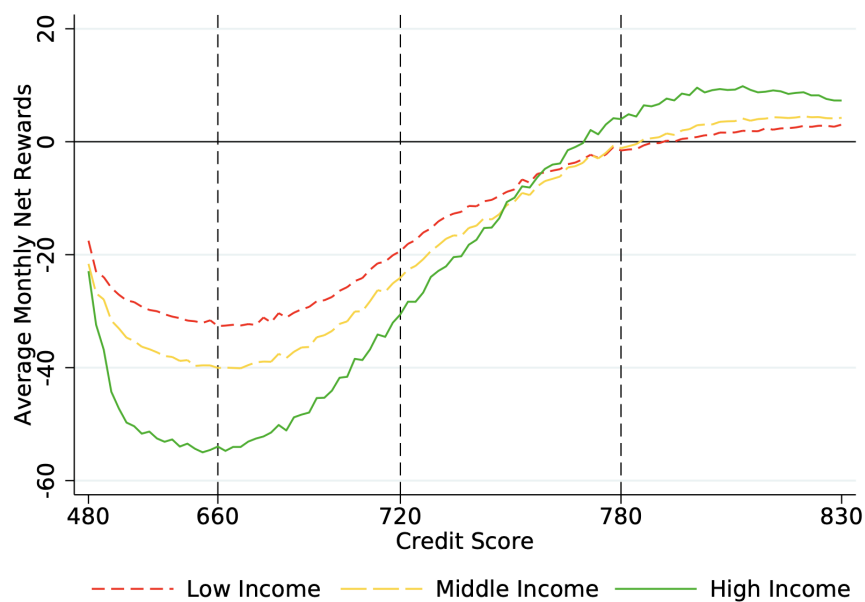


Figure 3: Average monthly net rewards by income and credit score. Source: [Agarwal et al. \(2023\)](#).

it is primarily the super prime cardholders who have positive monthly net rewards, while the sub-prime and near-prime cardholders pay the most for using rewards cards. Interestingly, high-income, super-prime cardholders benefit the most (\$20.1 in net rewards per month), primarily at the expense of *high-income* cardholders with low credit scores.

[Agarwal, Presbitero, Silva, and Wix \(2023\)](#) use FICO scores as a proxy for financial sophistication, and conclude that “sophisticated individuals profit from reward credit cards at the expense of naïve customers.” They also estimate the aggregate annual redistribution to be \$15 billion and find that this redistribution takes place from less to more educated, poorer to richer, and high to low minority areas, widening existing disparities.

3 Theoretical Model

Above we have seen that, in order to maximize the net benefits of credit card rewards, we should show “sophisticated” financial behavior. Assuming we have such a sophisticated credit card user who never pays interest, and who spends her entire discretionary budget (staying within her means) using rewards credit cards, I will model the maximum net benefit for a different number of credit cards and for different types of users (e.g. users who will redeem credit card points for

travel versus users who are only interested in cashback).

All reward credit cards come with certain point multipliers in specific categories, that will be multiplied with the spend in that category to calculate the number of points. For cashback rewards, the value of a point is per definition one cent, but for certain cards the value of a point can vary between a lower value (base value) and a higher value when redeemed for travel. Cards might also have an annual fee, or come with certain static benefits such as airport lounge access, travel or food credits, and discounts toward certain subscriptions such as Disney⁺. In this project I will ignore sign-up bonuses (SUBs), which were 9.1 percent of total reward earnings in 2022 (CFPB, 2023), since these high (but one-time) SUBs will always result in a very high marginal benefit for relatively little spend. Including them would lead to the conclusion that it is always beneficial to open another account with a SUB, into perpetuity.⁵

The problem of optimizing a credit card portfolio comes down to maximizing the net benefit by selecting, say at most, K credit cards from a dataset of N possible cards. By itself, an algorithm to solve such a problem would, at best, need exponential-time resources, and is therefore NP-hard and intractable (Paarsch and Golyaev, 2016, Sect. 6.2.3). Luckily for us, people usually apply for credit cards one at a time, and over the course of months or years. We can therefore easily discipline our problem by assuming that people start with a single credit card, and subsequently add cards to their portfolio. Smaller portfolios are also preferred for the sake of simplicity (and limited wallet space). This additional structure turns our problem of selecting the optimal credit card portfolio into a tractable *greedy* algorithm: the problem is broken down into smaller subproblems (“what is the best card to add now?”), and the best solution (highest additional value) is selected first. It also uses elements from *dynamic programming*, since we will need to store the intermediate solutions, and we will be using tables to calculate the marginal benefit that can still be extracted from the remaining cards.

The first subproblem is to select the best single card, and use it for all our spending categories. For adding an additional card, we have to consider all remaining cards and compare the value for each spending category to the previous card that we had already selected.⁶ The card

⁵People who keep opening new credit card accounts just for the SUB are known as *churners*.

⁶In code, I implemented this by creating a matrix with the values per card and spending category. The values of the previously selected card are subtracted from this matrix, and negative values are set to zero. The remaining positive values show where additional value can be extracted from the corpus of cards.

with the highest additional net benefit (summing all categories and benefits, and subtracting fees) is added to the portfolio, and a table with the best card to use for each category is updated. To select K cards, we have to repeat the search through all N possible cards K times, making this a feasible algorithm with complexity $\mathcal{O}(N)$.

4 Empirical Specification

The theoretical model described in Sect. 3 can be broken down into the following empirical specification. For every credit card k we have for the value earned per spending category c :

$$y_{kc} = x_{kc} m_{kc} [\eta v_{t,k} + (1 - \eta) v_{b,k}], \quad (1)$$

where x_{kc} is the spend on card k in category c , m_{kc} is the card multiplier for that category, $v_{t,k}$ and $v_{b,k}$ are the card's highest and lowest point redemption values, and η quantifies which fraction of the points are used for the higher-valued travel redemptions (a user-specific variable). The total value from spending is simply found by summing over the C possible categories:

$$s_k = \sum_{c=1}^C y_{kc}. \quad (2)$$

For the total benefit we add the static benefits b_k multiplied with the fraction of their use θ , subtract the annual fees f_k , and sum over all the K cards in our portfolio:

$$Y(\mathbf{X}, K, \eta, \theta | \mathbf{M}, \mathbf{v}_t, \mathbf{v}_b, \mathbf{b}, \mathbf{f}) = \sum_{k=1}^K (s_k + \theta b_k - f_k). \quad (3)$$

Here it is made explicit that the total benefit Y depends on user-specific variables \mathbf{X} (the budget matrix with spending per card and category), K (the number of cards), η (the fraction of travel redemptions), and θ (the fraction of benefits used), and the card-specific parameters \mathbf{M} (a matrix with the multipliers per card and category), \mathbf{v}_t (a vector with travel point values), \mathbf{v}_b (a vector with base values), \mathbf{b} (a vector with benefits), and \mathbf{f} (a vector with annual fees). One of the goals of this project is to study Y as a function of its variables, given the parameters. The other goal is to return the list of credit cards (including a list of which card to use for which category) that maximizes Y ,

after a user inputs the variables \mathbf{X} , K , η , and θ .

5 Data

The empirical specification from Sect. 4 (Eq. 3) implies that we need data on rewards credit cards point multipliers, point values, benefits, and fees, as well as budget data for our credit card users.

5.1 Credit Card Data

There are many websites writing about rewards credit cards, their perks, and points structure. I mainly used the website [allcards.com](https://www.allcards.com)⁷ to manually scrape the information on point multipliers and their caps (limits) from the most popular rewards credit cards from the seven major banks (American Express, Chase, Bank of America, Citi, Capital One, US Bank, and Wells Fargo). The bank websites were used to find some missing information, if needed. This resulted in a list of 28 unique credit cards. However, some of the cards have a “custom” bonus reward category that can be chosen by the user (e.g. 5x points on groceries, or gas, or home improvement up to \$6000 per year). These cards were treated as a separate credit card for each bonus category, bringing our total number of cards to 38.

I ignored credit cards that are dedicated to users of very specific stores, hotels, or airlines, as well as cards with rather exotic, or variable, reward categories (such as “3x points on all mobile wallet payments”, or “5x points on quarterly changing categories”), since these categories are impossible to map to our budget categories that will be discussed in the next section. The results on total benefits should therefore be considered *lower limits*, as there are many cards that can bring additional benefits to loyal users of certain brands.

A list of 18 common card spending categories was taken from the website [cardpointers.com](https://www.cardpointers.com),⁸ with the modification of removing “warehouse clubs” and “ride sharing”, and adding “streaming” and “travel (other)”, to facilitate a mapping of all the items on our budget to a spending category (see Sect. 5.2). The final list of all the categories can be found in the first column of Table 3 in the Data Appendix.

⁷<https://www.allcards.com> (accessed between June 3–9, 2024).

⁸<https://www.cardpointers.com/app/> (accessed between June 3–9, 2024).

For the base and travel values of the credit card points, I used the table from the [Nerdwallet \(2024\)](#) website. These values are shown in Table 2 in the Data Appendix and will be used for the sensitivity analysis and Monte Carlo simulation. Point values can be subjective, however, so for the R Shiny recommendation app, I foresee an update in the near future, making the point values adjustable by the user.

Annual fees were also collected for all the cards in the dataset, as well as an estimate for the static benefits. We refer the reader to the Data Appendix on page 32 for a more detailed description of how the static benefits were handled. The appendix also shows an excerpt of the final credit card dataset in Fig. 19.

5.2 User Data

For our credit card users we need a vector x with the amount of spend for each of the 18 spending categories. In the credit card portfolio recommendation app, the user can specify a personal yearly budget. For our analysis of how the optimal portfolios behave for different user preferences and budgets, however, we need realistic estimates of the average spending of Americans, since regulations and privacy concerns make obtaining actual spending data practically impossible for researchers outside of the banking sector. I will therefore use the 2022 Consumer Expenditure Survey (CES) from the Bureau of Labor Statistics ([BLS, 2023](#)).

According to the CES, the average annual expenditures in 2022 were \$72,967 (after savings), from an average income before taxes of \$94,003. To approximate the amount that can be spend on credit cards, I subtracted from the \$72,967 the expenses on shelter (such as mortgage, rent, property taxes), vehicle purchases, health insurance, education, cash contributions, personal insurance, and pensions. The resulting budget totals to \$38,576 or 41 percent of the average gross income. The remaining CES items were then mapped to the credit card spending categories according to Table 3 in the Data Appendix. Note that some items were split 50/50 over multiple credit card categories, to make sure that all the categories were populated with some reasonable spend. For example, I assumed that the CES item “Apparel and services” is split 50/50 between the “Online shopping” and “Department store” credit card categories.

The CES also provides expenditures separated by nine income levels. I have repeated the

above construction of the average budget also for these nine income bins separately, and saved the results in the file `BudgetIncome.csv`. The percentages spend per credit card category are also shown in Table 4 in the Data Appendix. For a given user’s income (before taxes), the corresponding income bin from this table is used to calculate the average budget, namely by multiplying the gross income by the percentage spend in each category. Our algorithm will then predict the total credit cards benefit as a function of income, as well as η , θ , and the number of credit cards K (see Sect. 4).

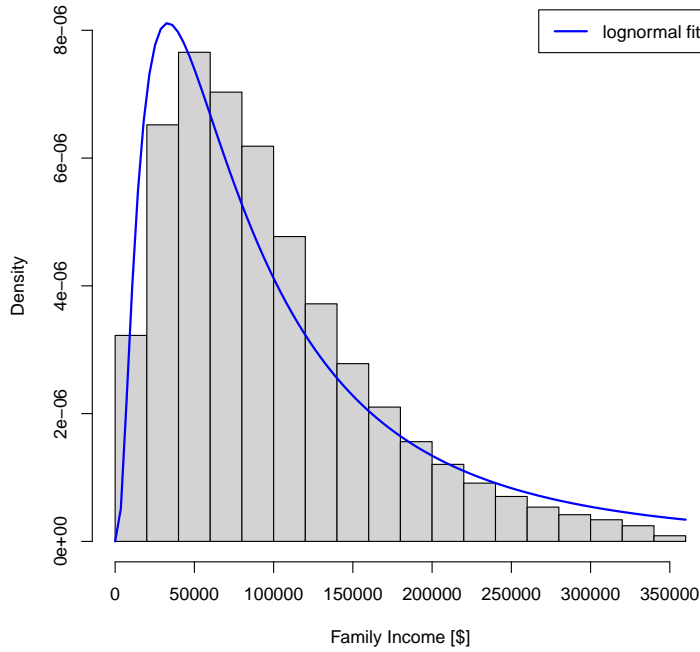


Figure 4: 2022 Florida income distribution from the BLS ACS (microdata sample of 57,005 observations, truncated at \$350,000 for this visualization).

A Monte Carlo simulation should give a reasonable estimate of the average net benefit and standard deviation for Americans who use an optimized credit card portfolio. For this simulation we need to sample (with replacement) income, η , θ , and K from realistic distributions. For the income distribution I used 2022 “Family income” (variable FINCP) from the BLS American Community Survey (ACS) Microdata tables, selecting all counties in Florida (resulting in 57,005 individual incomes).⁹ This income distribution is shown in Fig. 4 and approximates a lognormal

⁹<https://data.census.gov/mdat/#/search?ds=ACSPUMS1Y2022> (accessed June 25, 2024).

distribution. User preferences η and θ are sampled from a uniform distribution between 0–1, and K is sampled from a uniform distribution between 1–8.

6 Results

In this section, I first present the algorithm I developed, to return the optimal credit card portfolio, given a set of user variables and card parameters. I then describe the results of a sensitivity analysis and Monte Carlo simulation that show how the return on spend and marginal benefit depend on the user variables. Finally, I will describe the interactive R Shiny app that is available online for people to play with and to get a recommended credit card portfolio tailored to their personal preferences.

6.1 Portfolio Recommendation Algorithm

The model described in Sections 3 and 4 is coded in R using two functions that are saved in the file `CCPortfolioFunctions.R`. The function `get_budget(income, budget_data)` is a helper function that returns the budget as a named vector, given a gross income, corresponding to the 18 credit card categories from Sect. 5.1 and using the BLS CES income-dependent data described in Sect. 5.2. This budget is then used as input for the second function, `get_portfolio(K, eta, theta, cards_data, budget, verbose = TRUE)`, which returns a named list with the optimized portfolio of credit cards, as well as some other useful metrics (the total net benefit in dollars, percentage return on spend, the marginal benefit in dollars per card, the total spend, and which card to use for each spending category). The full code of both these functions is presented in the Code Appendix on page 38.

Figure 5 summarizes some of the algorithm’s essential output in a user-friendly format, for two different user preferences. The figure shows that the algorithm is successful in recommending different portfolios for different users. The more cashback-oriented user (top panel) gets recommended more cashback focused cards, while the more travel-oriented user (bottom panel) gets recommended a few more travel focused cards, which results in a higher total benefit (the orange bar on the right). Note that the marginal benefit plotted per card is the *additional* benefit that the card delivers, *on top* of the cards already selected to its left. In other words, this marginal benefit

includes the opportunity cost of taking spend categories away from the cards to its left, in order to put that spend on the additional card.

Figure 5 also shows that the marginal benefit diminishes with more cards, which is by design of the greedy algorithm. The algorithm will always recommend cards with the highest marginal benefit first. It is up to the user of the recommendation app (see Sect. 6.4) to decide at which marginal benefit it is no longer worth it to add additional cards, but for the remainder of this paper I assume that \$50 is probably a reasonable amount for most people to stop adding to the portfolio. This allows me to determine the “sweet spot” for the credit card portfolio size, as a function of the variables. I discuss this sweet spot, along with some other metrics of the recommended portfolios, in the next section.

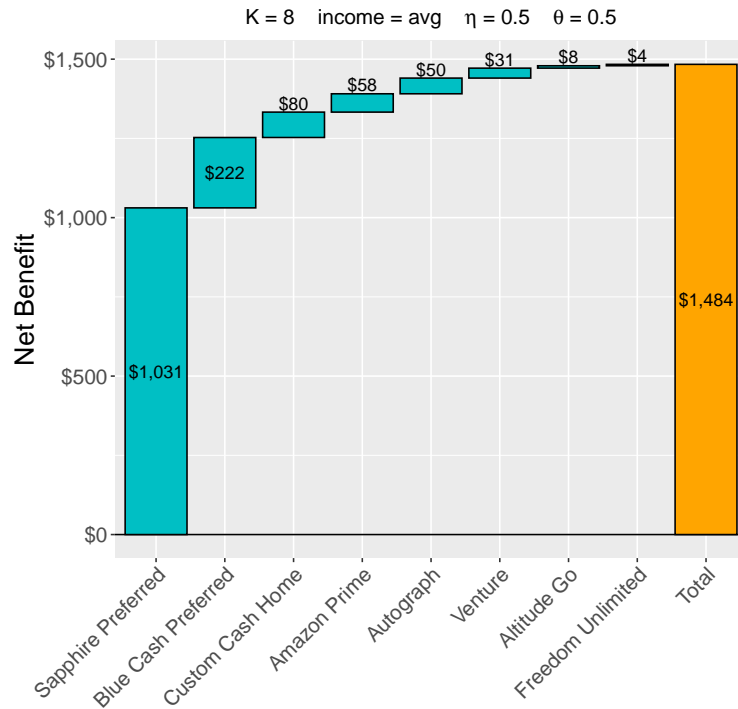
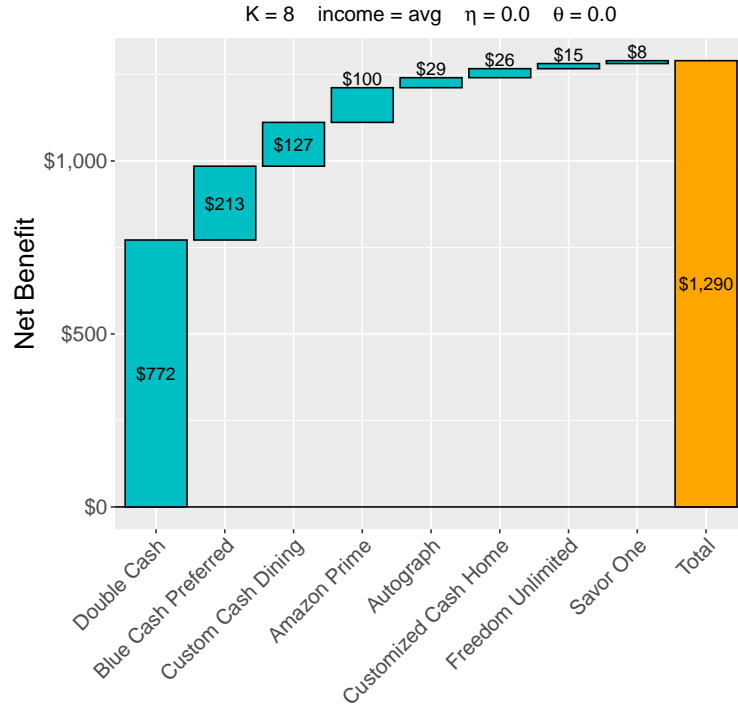


Figure 5: Waterfall-style summary of two recommended portfolios, for different user preferences. Both figures returned the best 8 cards ($K = 8$) for the budget of the average consumer (income = avg). The first card is on the left and the bars indicate the marginal benefit of adding each additional card. The top panel user has a cashback preference and uses no static benefits (η and θ both equal 0). The bottom panel user transfers half the points to travel partners ($\eta = 0.5$) and uses half of each card's benefits ($\theta = 0.5$).

6.2 Sensitivity Analysis

Figures 6, 7, and 8 show the marginal benefit as a function of the number of cards, for different incomes, and for (η, θ) equal to $(0,0)$, $(0.5,0.5)$, and $(1,1)$, respectively. The income levels for these figures are chosen to represent the average incomes of the nine income levels of the BLS CES data (Table 4 in the Data Appendix), while the thick black dashed line represents the average budget of all consumers in the survey (with an average income of \$94k). These figures show that for most people, regardless of preferences or income, four or five is the sweet spot for the number of credit cards to hold, before the additional benefit drops below \$50. There seems to be some evidence that the highest three income levels (above \$100,000) maintain a higher marginal benefit with more cards (up to six or seven, before the benefit drops below \$50).

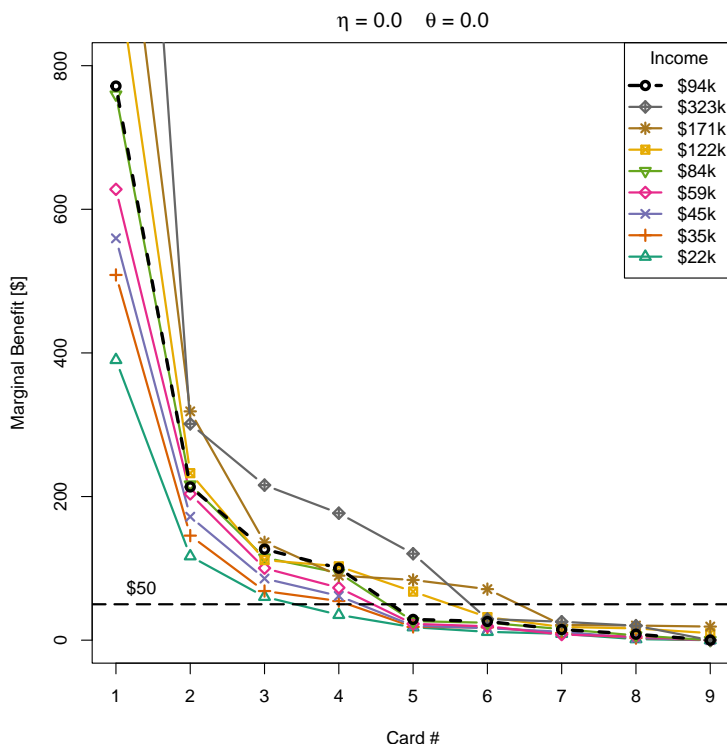


Figure 6: Marginal benefit versus number of cards, assuming $(\eta, \theta) = (0, 0)$.

Figures 9, 10, and 11 show the total return on spend versus the number of cards, for the same choice of (η, θ) as the previous three figures. These figures show that ROS converges to ~ 3.2 percent for cashback-oriented users, regardless of income level (Fig. 9). Once travel transfer partners and benefits are being used, Figs. 10 and 11 show that there is more variation among different

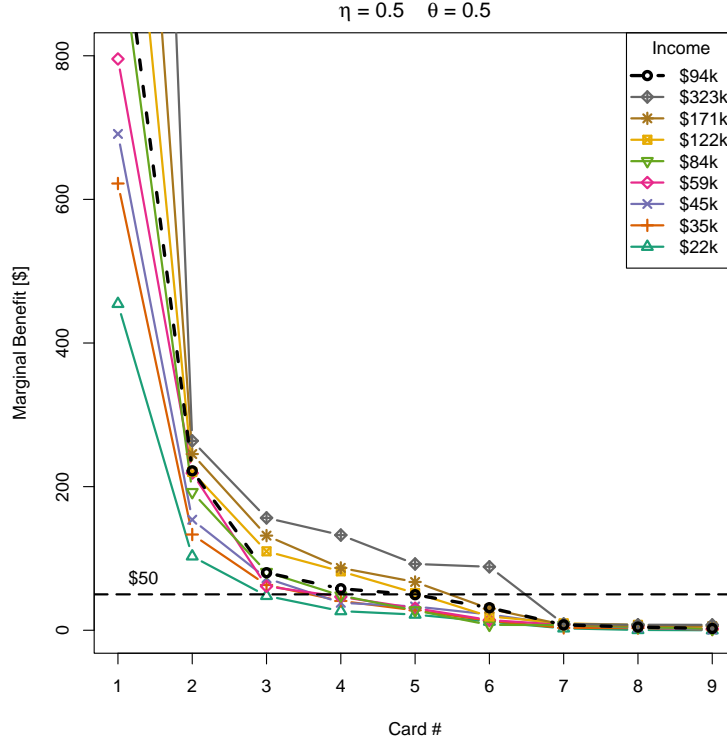


Figure 7: Marginal benefit versus number of cards, assuming $(\eta, \theta) = (0.5, 0.5)$.

income levels. As expected, travel and benefits return additional value, and increase ROS for all incomes. Once all benefits are being used, however ($\theta = 1$, Fig. 11), we see an interesting reversal in the effect for different incomes: the lowest income levels reach a ROS comparable, or even higher, than the highest income levels (e.g., note the cyan triangles above the gray diamonds). Consumers at the lowest income levels can spend much less on credit cards compared to high-income consumers, but Fig. 11 reminds us that low-income consumers can benefit more, relative to their spend, as long as they make full use of the static credit card benefits. Figure 11 also shows that 5.5–6.5 percent is roughly the maximum ROS people can expect, when making full use of an optimized portfolio of 4–5 credit cards.

Figure 12 shows the range in total net benefit (in dollars) versus income. This range lies between $\sim \$700$ – $\$1200$ for the lowest incomes, and increases all the way up to $\sim \$2000$ – $\$5000$ for incomes above $\$300k$.

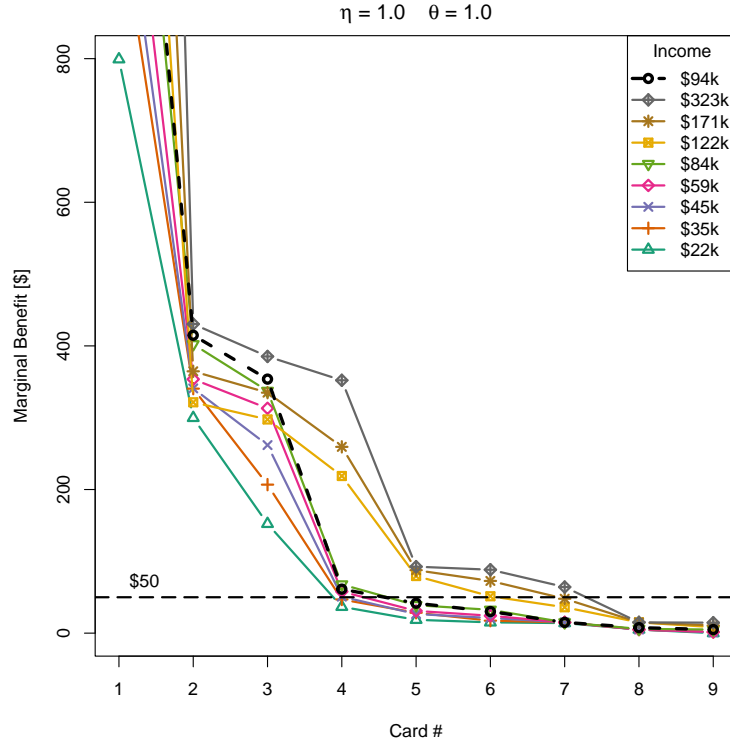


Figure 8: Marginal benefit versus number of cards, assuming $(\eta, \theta) = (1, 1)$.

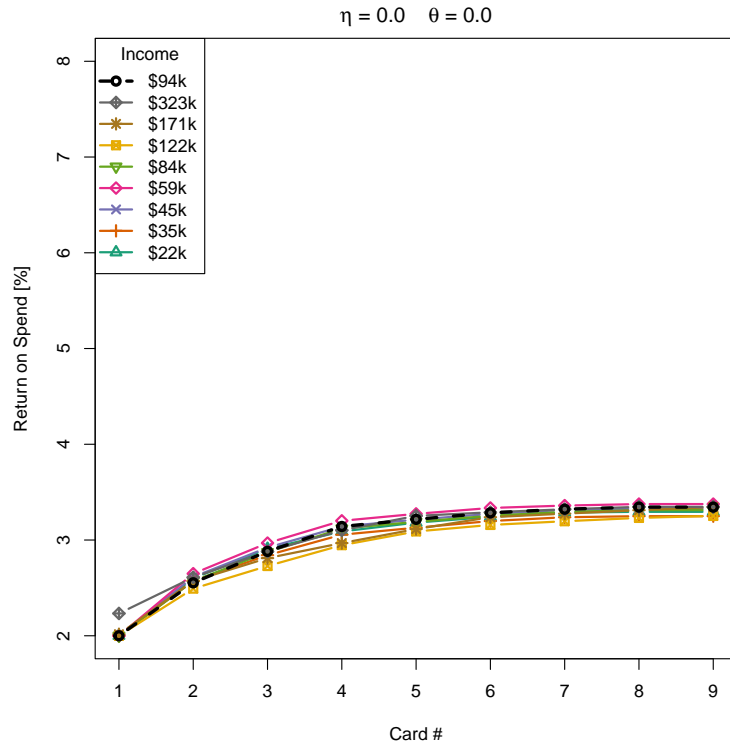


Figure 9: Total return on spend versus the number of cards, assuming $(\eta, \theta) = (0, 0)$.

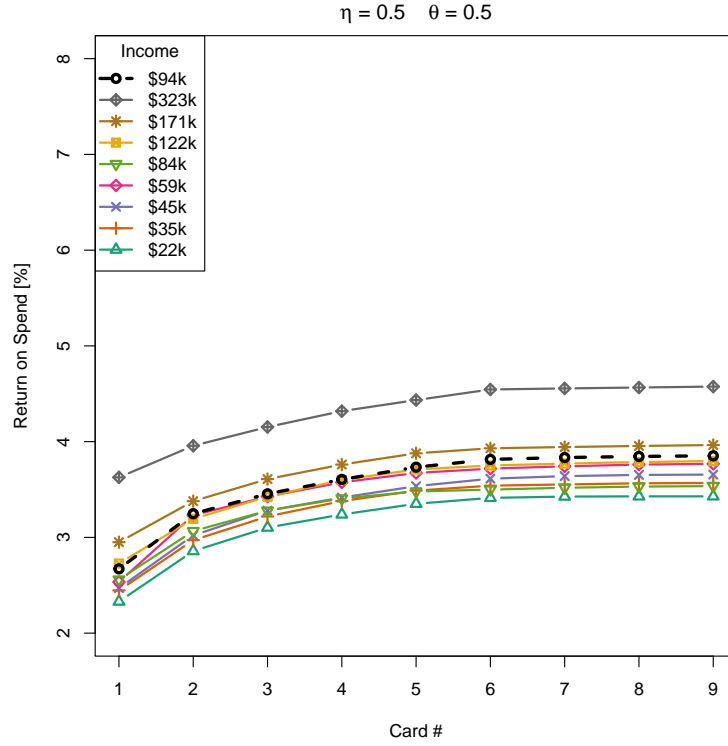


Figure 10: Total return on spend versus the number of cards, assuming $(\eta, \theta) = (0.5, 0.5)$.

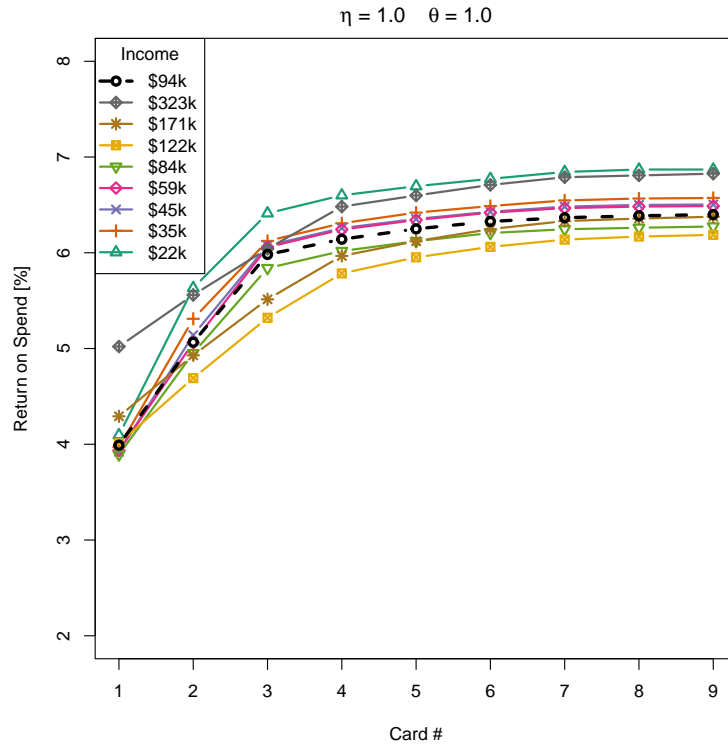


Figure 11: Total return on spend versus the number of cards, assuming $(\eta, \theta) = (1, 1)$.

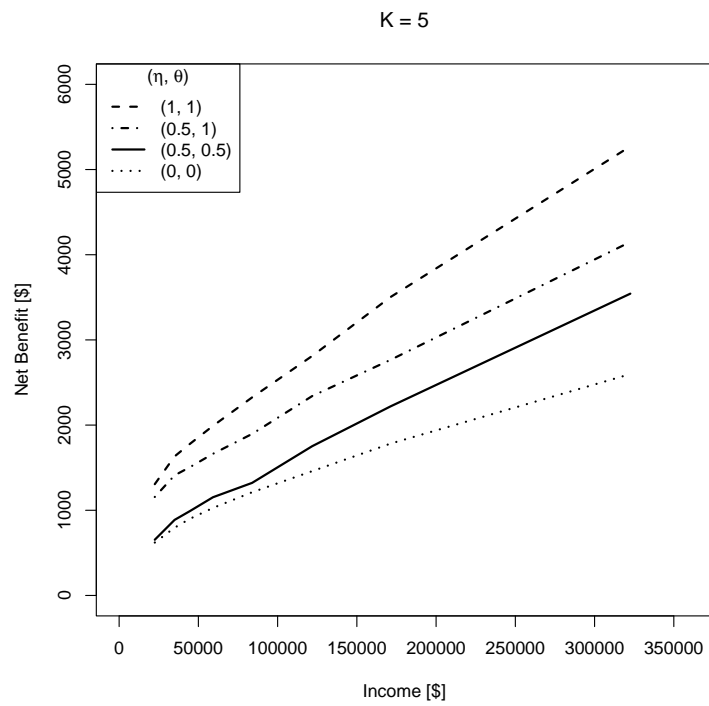


Figure 12: The total net benefit versus income for different combinations of η and θ .

6.3 Monte Carlo Simulation

Above we have seen how income and user preferences impact the expected return on spend from rewards credit cards. Income affects these results in two ways, as can be seen from Table 4 in the Data Appendix. First of all, total spend scales with income; higher incomes can spend more on credit cards, although the total spend as a fraction of income goes down (i.e., higher incomes can put more money toward savings and housing, which does not end up on credit cards). Second, with changing incomes there are also shifts in the spending budgets; essential goods like groceries and gas form a much larger part of the budget for lower incomes, compared to higher incomes.

To represent more accurately how these patterns affect the benefits of rewards credit cards for a large group of users, I performed a Monte Carlo simulation using 100,000 samples. The family income distribution of 57,005 Floridians was used, as described in Sect. 5 and presented in Fig. 4. This distribution was sampled, with replacement, 100,000 times, and for each income the average budget from the corresponding income bin from the BLS CES data was assumed. Each income was then combined with a random value for K (between 1–8), η (0–1), and θ (0–1), which were sampled from uniform distributions. The optimal portfolio was determined for each sample, and results regarding net benefits, return on spend, and a count of each selected credit card was saved to a .csv file for further analysis. This simulation took roughly two hours to process on a 2023 MacBook Pro.

Figure 13 shows the ROS distribution resulting from the Monte Carlo simulation. The distribution approximates a lognormal with a mode at 3.55 percent, while the mean ROS is 3.80 percent with a standard deviation of 0.80.

Figure 14 shows a boxplot of the return on spend versus the number of cards, which confirms the earlier observations that marginal benefit diminishes strongly after four or five cards.

Using a logarithmic scale for income, Fig. 15 shows a smoothed density scatterplot of all 100,000 simulated portfolios. This shows where the highest density of our “observations” are, as well as how ROS varies with income. Per income, the vertical spread in density is caused by the random sampling of K (number of cards), η (use of transfer partners), and θ (use of benefits). The spread in the horizontal direction is caused by the observed Floridian income distribution. A “loess” curve (short for “locally estimated scatterplot smoothing”) is overplotted in red, and

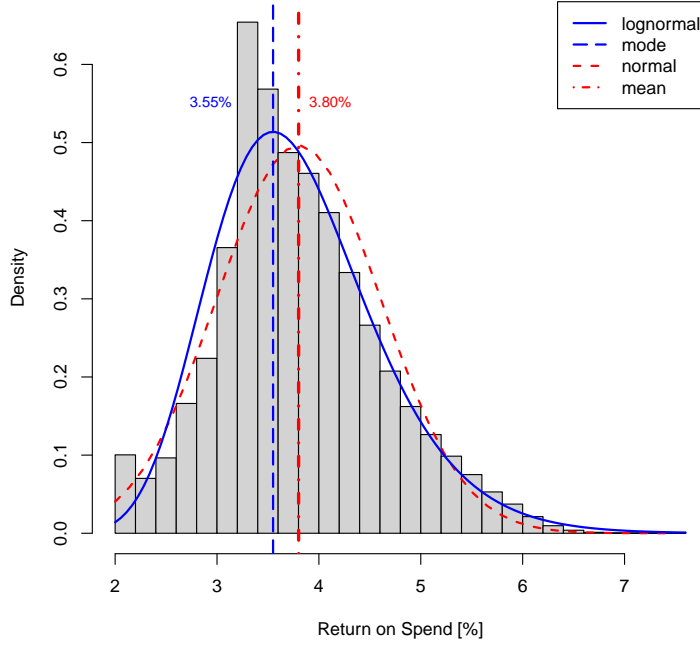


Figure 13: Distribution of return on spend, resulting from the Monte Carlo simulation.

highlights a trend where the mean ROS is increasing for incomes above $\sim \$120,000$. This upward trend must be attributed to shifts in spending patterns with income. Looking at the fraction of spend going toward the travel, dining, and groceries categories in Fig. 16, we see that essentials like groceries becomes a less important part of the budget as income increases. Travel, on the other hand, takes up a larger fraction of the budget for higher incomes, while dining (away from home) shows this effect to a smaller degree. Not surprisingly, in our credit cards dataset we find some of the highest point multipliers on the travel categories, usually on cards with the highest annual fees, such as the American Express Platinum and the Chase Sapphire Reserve. High-income individuals can more easily justify these high annual fees, especially since they also have the spending behavior that matches well with the rewards categories of these cards, leading to above-average benefits.

Finally, the results of the Monte Carlo simulation can be used to show how often a specific credit card ends up in a recommended portfolio. Figure 17 shows the top 10 most recommended cards. The Amazon Prime card and the American Express Blue Cash Preferred are the two most

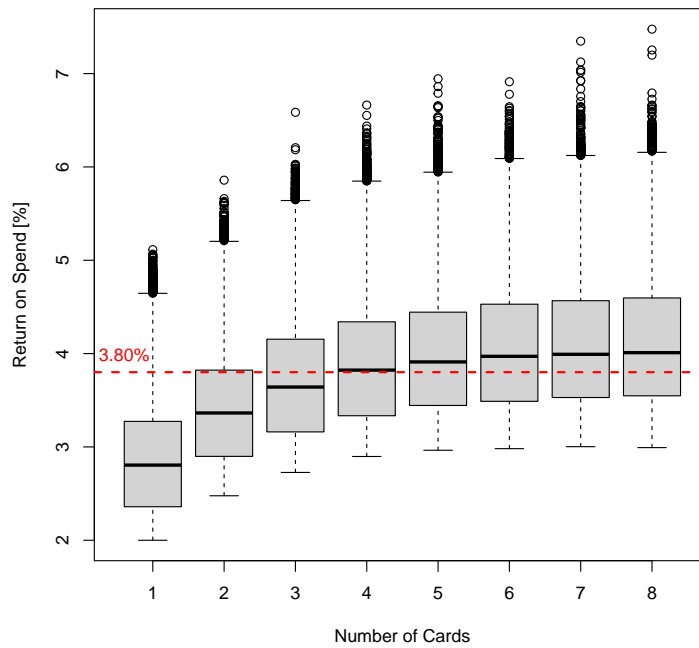


Figure 14: Boxplot of the return on spend versus the number of cards, resulting from the Monte Carlo simulation. The horizontal dashed line indicates the mean ROS of 3.80 percent, which is very close to the mean ROS of consumers using four credit cards.

recommended cards, followed closely by the Wells Fargo Autograph.

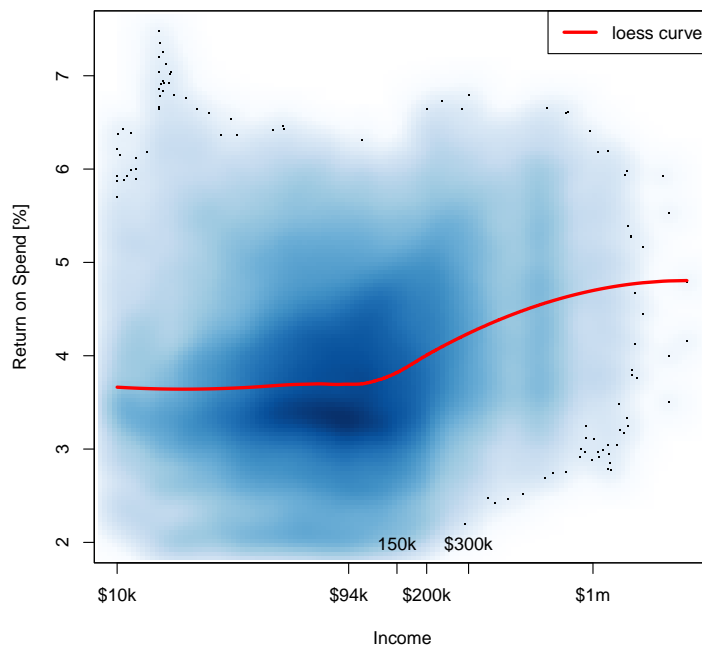


Figure 15: Smoothed density scatterplot of return on spend versus income, for all 100,000 simulated portfolios.

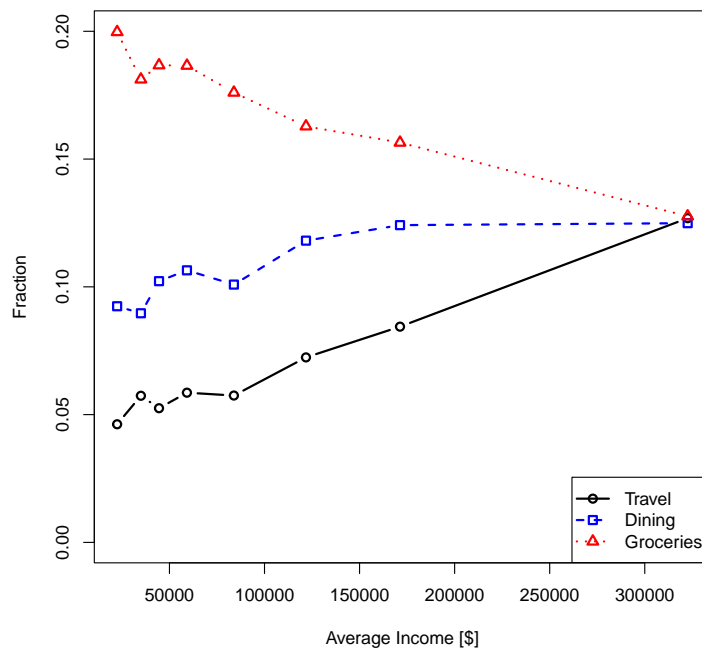


Figure 16: Fraction of total credit card spend going toward travel, dining, and groceries, for the average incomes from the BLS CES data (mapped to the credit card spending categories).

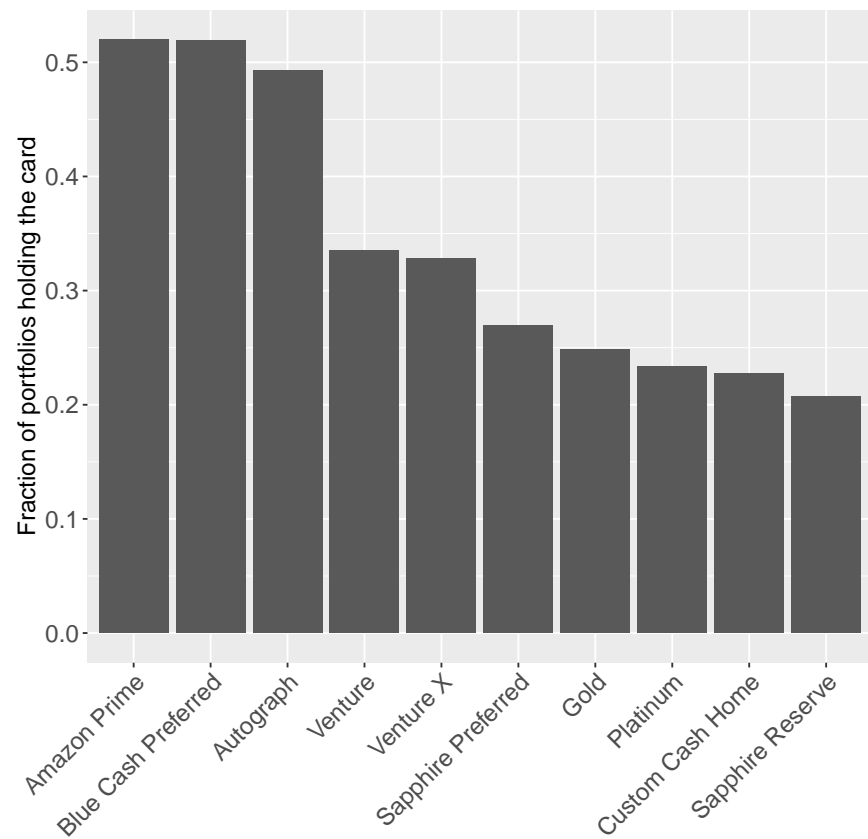


Figure 17: Top 10 most recommended credit cards from the Monte Carlo simulation.

6.4 Shiny App

No credit card user is the same. Above we have seen how credit card rewards can benefit users with average income levels and average budgets, but what we really need is a way to make *personal* recommendations based on someones individual spending behavior, point valuations, and other preferences. This is why I have also developed an R Shiny app that will return the optimal portfolio for a user's personal preferences and individual budget.

A screenshot of the current version of the app is shown in Fig. 18, and the app can be accessed online at <https://remcoscheepmaker.shinyapps.io/ReMCCO/>.¹⁰ Every Shiny app consists of a user interface and a server function. The server function makes use of exactly the same `get_budget()` and `get_portfolio()` functions as the rest of this project (presented in the Code Appendix), so the underlying algorithm works the same (although the app rounds the budget to the nearest dollar, which explains small differences between Figs. 5 and 18).

In the left column of the user interface, a user can specify preferences such as number of cards (K), use of transfer partners (η), use of benefits (θ), and which banks to include. In a second column, the spending budget is pre-filled to match an average budget based on income (using Table 4), but the user can adjust this budget freely, either by adjusting the income, or by changing the amount for each category individually, making the budget completely personal. The waterfall plot in the top right will automatically adjust after each change (using Shiny's "reactive" programming functionality). This waterfall plot shows the recommended credit card portfolio, including the marginal and total benefits, just like Fig. 5. Underneath the figure some summary information is printed, and the card assignments are presented in a table (to show which card to use for each spending category).

In its current form, the app already has more functionality than what was captured by the variables and parameters that I used in the analysis above, since the user can also filter for banks and personalize the budget. The app could be further improved, however, by also making the valuations of points and benefits user-adjustable. This way, the user can specify that it values points or benefits from certain banks differently, for example because one bank has a transfer partner that the user really values, or because the user is a preferred client at one bank, which

¹⁰The working title of the app is "ReMCCO", which stands for Recommend Me Credit Cards Optimally.

might come with additional benefits.¹¹ A future update of the app is planned to include these improvements.

¹¹For example, Bank of America's *Preferred Rewards* program offers 25, 50, or 75 percent higher point values for clients with \$20,000, \$50,000, or \$100,000 in combined assets at Bank of America / Merrill.

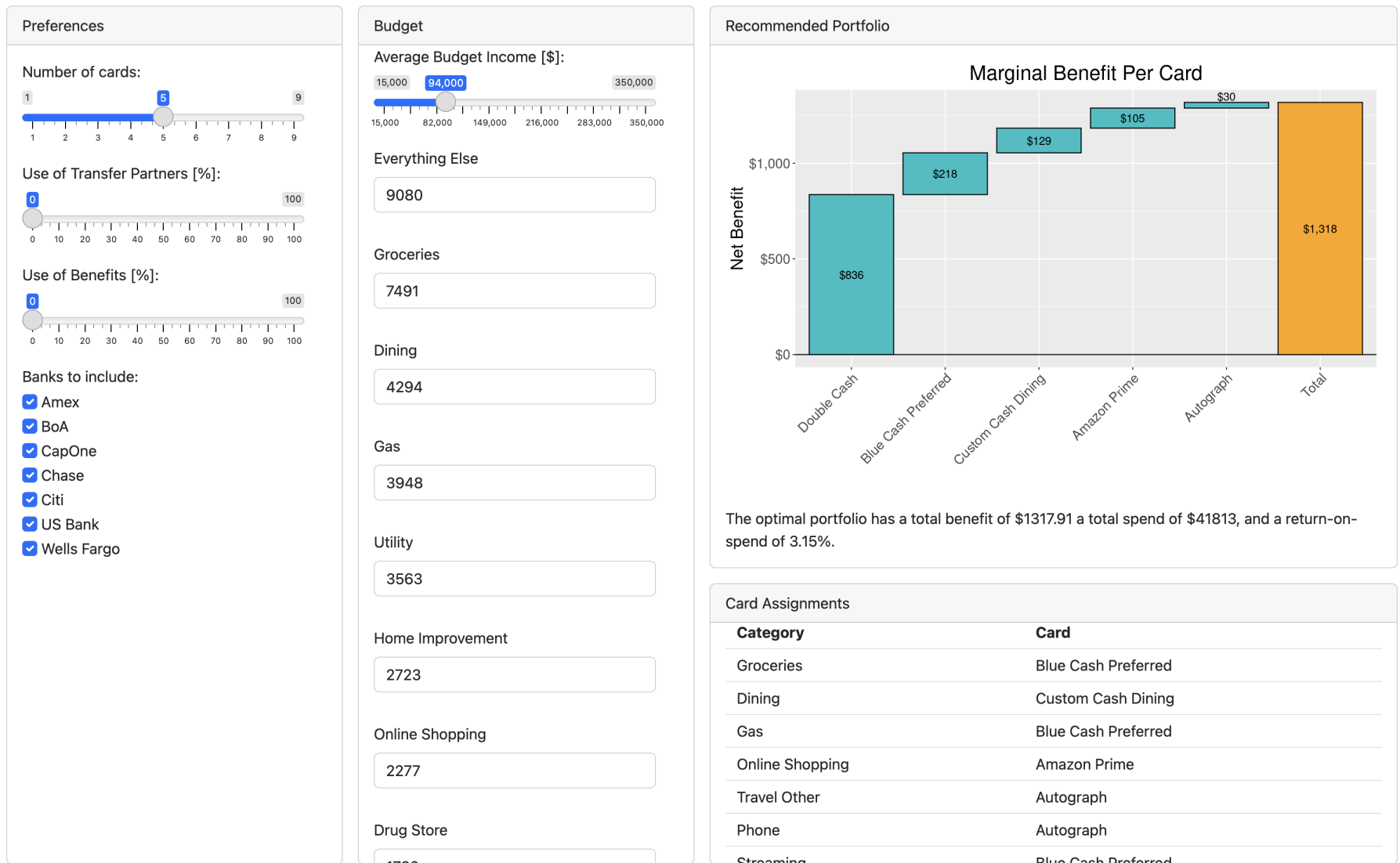


Figure 18: Screenshot of the online R Shiny user interface at <https://remcoscheepmaker.shinyapps.io/ReMCCO/>.

7 Conclusions

We have seen that credit cards can be very rewarding for those (financially sophisticated) people who have the discipline to pay them off in full each month. Due to inefficiencies in the the rewards programs, it is mostly people who use cash, or the (financially naïve) people who use credit cards in a nonoptimal way, who are paying for the rewards of the financially sophisticated, through inflated prices, interest, and fees. Also the people who use credit cards optimally, however, pay for some fraction of their rewards through the same inflated prices, that merchants most likely have raised to recoup some of the interchange fees set by the credit card networks.

By developing a credit card recommendation algorithm, I have shown that people can expect an ROS of up to ~ 6.5 percent, by putting all their spend on up to 4–6 credit cards. A Monte Carlo simulation showed that the mean expected ROS is 3.80 ± 0.80 percent, and the ROS is higher for people who are using travel transfer partners for their points, and who make use of all the static benefits that credit cards might offer. These numbers assume that people are willing to put in some time and research into their personal finances, in order to estimate their yearly spending patterns and apply for those cards that best fit this spending budget. To help with this, I have developed an online app that takes a user’s budget and preferences into account, in order to recommend the optimal credit card portfolio.

My results also show a trend of ROS increasing with income, once a threshold income of roughly \$120,000 is reached. This is most likely explained by a shift in spending patterns, as people with higher incomes spend an increasing fraction of their budget on travel, which is a category that offers higher rewards, without any meaningful spending limits (caps). Categories like groceries, or custom categories with elevated 5x multipliers, on the other hand, usually come with caps on the amount of rewards that can be earned, limiting the ROS on these categories for big spenders with high incomes. The observation that ROS increases with income seems consistent with the conclusion from [Agarwal et al. \(2023\)](#) that reward credit cards are widening existing disparities between the poor and the rich.

Since this project was planned, executed, and presented in a total timespan of about 11 weeks, I had to ignore certain details, and make some reasonable assumptions. For example, I have ignored foreign transaction fees, which are typically 3 percent for non-travel oriented credit cards.

Travel credit cards usually do not have foreign transaction fees, but this still means that frequent international travelers have to leave certain other cards from their portfolio at home, complicating the optimal portfolio. With 119 million merchants accepting American Express internationally, compared to 130 million accepting Visa/Mastercard ([Thrifty Traveler, 2024](#)), the acceptance rate for American Express is still significantly lower abroad. This might also complicate someone's decision to choose credit cards, and is part of the reason why I included a bank filter in the Shiny app.

One of the assumptions I made is that all users qualify for all the cards in the dataset, because I assumed that users are financially sophisticated, with high FICO scores, before they even want to optimize their credit card portfolio. One could argue indeed that this is not very realistic, and even with a prime-plus or super-prime credit score (i.e., a FICO score above 720, see Table 1 on page 4), the income of a user might not be high enough to get approved for certain cards that come with high credit lines. Banks, however, do not openly share their income and credit score requirements for their credit cards, and they probably also use other, non-disclosed, information to make approval decisions. So, even if a user would be asked to provide their income and FICO score in the app, an algorithm could only guess for which cards the user would be approved. Due to the time constraints, I have not included such an approval estimate in my algorithm, but it is a feature to consider for a future update. Until then, users of the app should understand that the recommended portfolio is something to aim for, even if certain cards seem currently unfeasible, based on the user's current income and credit score.

Despite all these assumptions, I hope my app delivered some utility to the reader and my fellow classmates, and that this project has led to an increased awareness, and understanding, of the economics of credit card rewards, and the benefits for their financially sophisticated users.

Acknowledgements

This capstone project would not have been possible without Dr. Harry Paarsch. I would like to thank him for all his teachings and guidance throughout the last year, and for his book with Dr. Konstantin Golyaev ([Paarsch and Golyaev, 2016](#)). He made me realize that economics is much closer to a real (natural) science than I always feared. I would also like to thank all the other

professors that made this first cohort of the Business Analytics program possible: Dr. Lealand Morin, Dr. Alexander Mantzaris, and Dr. Majid Mahzoon for their software and data visualization classes. Dr. Michael Tseng for his math class and Dr. Nizam Uddin for his class on probability and statistics. I thank Dr. Morgan Wang for introducing me to SAS, and I would like to thank Mr. Joshua Eubanks for teaching me my first microeconomics class. I would also like to thank Ms. Meredith Smart for enrolling me exactly on schedule, every Friday before the start of the semester.

Finally, I would like to thank my fellow Business Analytics students from the 2023–2024 cohort, who made going back to school after 15 years a lot more fun. Malak, Kelly, Cody, Jared, Manoj, Nili, Frank, Danielle, Yasmeen, Anabella, and Avery: thanks for all the discussions, and for putting up with me and my Dutch directness. I'm wishing you all the best in your future careers, wherever those may bring you. Go Knights, and stay curious!

A. Data Appendix

In this appendix, I present an excerpt of the rewards credit cards dataset, and describe in more detail how the static benefits were handled. I also present tables of the spending budgets and income levels that were used for the sensitivity analysis and Monte Carlo simulation.

Figure 19 shows an excerpt of the credit cards dataset. Note that the columns “id” contains duplicate numbers for:

- cards that are mutually exclusive (Chase Sapphire Preferred and Reserve);
- cards that have custom reward categories, but the user can only hold one of these cards at a time (Citi Custom Cash);
- cards that have identical rewards, but are from different banks (Citi Custom Cash and Wells Fargo Active Cash).

Having identical “id” numbers for these cards makes it easier to deal with these situations in the algorithm, after one of these cards is selected for the portfolio (namely by setting all values to zero of cards with the same “id”).

If a card has a travel or food credit, this credit was taken as a benefit at full value (column “benefit_credits”). Benefits that are not easily combined between different credit cards, such as a “Global Entry / TSA Precheck” credit, airport lounge access, and CLEAR membership credit, received a “1” flag in the columns “benefit_globalentry”, “benefit_lounge”, and “benefit_clear”, respectively (or “0” if not available). This allows for an easy solution in the code to avoid counting these benefits multiple times, when in practice they can only be used once. The \$100 “Global Entry / TSA Precheck” credit every five years is assumed as a \$20 yearly benefit. Airport lounge access is valued at \$40, which is approximately the value of two free meals with drinks. The CLEAR membership credit is a \$189 value. Once a card is added to the portfolio in the recommendation algorithm, all these values are multiplied by 1 – the flag of the selected card, which means the benefit is “turned off” after its first use.

A boolean column “cash_only” was added to indicate if the card is cashback only (TRUE), or if the card allows for higher-valued travel redemptions (FALSE). This could potentially be used

to filter for certain user preferences. The columns “base_value” and “travel_value” were populated with the point valuations from [Nerdwallet \(2024\)](#), which are also shown in Table 2. The final 36 columns (two for each spending category) contain the point multipliers and their limits (“..._cap”), sourced from the website <https://www.allcards.com> as well as the bank’s websites. All these credit card data were manually recorded in a Google Sheet and finally saved to the file `CreditCards.csv`.

id	bank	name	fee	benefit_globalentry	benefit_lounge	benefit_clear	benefit_credits	cash_only	base_value	travel_value	travel_other	travel_other_cap	airline_portal	airline_portal_cap	hotel_portal	hotel_portal_cap	car_portal
1	Amex	Blue Cash Preferred	95	0	0	0	84	TRUE	0.01	0.01	1	0	1	0	1	0	1
2	Amex	Blue Cash Everyday	0	0	0	0	0	TRUE	0.01	0.01	1	0	1	0	1	0	1
3	Amex	Everyday Preferred	95	0	0	0	0	FALSE	0.01	0.02	1	0	2	0	2	0	2
4	Amex	Everyday	0	0	0	0	0	FALSE	0.01	0.02	1	0	1	0	1	0	1
5	Amex	Green	150	0	1	1	0	FALSE	0.01	0.02	3	0	3	0	3	0	3
6	Amex	Gold	250	0	0	0	140	FALSE	0.01	0.02	3	0	3	0	1	0	1
7	Amex	Platinum	695	1	1	1	795	FALSE	0.01	0.02	5	500000	5	500000	5	0	1
8	Chase	Freedom Unlimited	0	0	0	0	0	TRUE	0.01	0.01	1.5	0	5	0	5	0	5
9	Chase	Sapphire Preferred	95	0	0	0	50	FALSE	0.0125	0.026	2.1	0	5.1	0	5.1	0	5.1
9	Chase	Sapphire Reserve	550	1	1	0	300	FALSE	0.015	0.027	3	0	10	0	10	0	10
10	Chase	Amazon Prime	0	0	0	0	0	TRUE	0.01	0.01	1	0	5	0	5	0	5
11	BoA	Premium Rewards	95	1	0	0	100	TRUE	0.01	0.01	2	0	2	0	2	0	2
12	BoA	Travel Rewards	0	0	0	0	0	FALSE	0.01	0.01	1.5	0	3	0	3	0	3
13	BoA	Unlimited Rewards	0	0	0	0	0	TRUE	0.01	0.01	1.5	0	1.5	0	1.5	0	1.5
14	BoA	Customized Cash Dining	0	0	0	0	0	TRUE	0.01	0.01	1	0	1	0	1	0	1
15	BoA	Customized Cash Online	0	0	0	0	0	TRUE	0.01	0.01	1	0	1	0	1	0	1
16	BoA	Customized Cash Gas	0	0	0	0	0	TRUE	0.01	0.01	1	0	1	0	1	0	1
17	BoA	Customized Cash Home	0	0	0	0	0	TRUE	0.01	0.01	1	0	1	0	1	0	1
18	BoA	Customized Cash Drug	0	0	0	0	0	TRUE	0.01	0.01	1	0	1	0	1	0	1
19	Citi	Double Cash	0	0	0	0	0	TRUE	0.01	0.01	2	0	2	0	2	0	2
19	Wells Fargo	Active Cash	0	0	0	0	0	TRUE	0.01	0.01	2	0	2	0	2	0	2
20	Citi	Strata Premier	95	0	0	0	0	FALSE	0.01	0.015	3	0	10	0	10	0	10
21	Citi	Custom Cash Gas	0	0	0	0	0	TRUE	0.01	0.01	1	0	1	0	1	0	1
21	Citi	Custom Cash Groceries	0	0	0	0	0	TRUE	0.01	0.01	1	0	1	0	1	0	1
21	Citi	Custom Cash Dining	0	0	0	0	0	TRUE	0.01	0.01	1	0	1	0	1	0	1
21	Citi	Custom Cash Streaming	0	0	0	0	0	TRUE	0.01	0.01	1	0	1	0	1	0	1
21	Citi	Custom Cash Drugs	0	0	0	0	0	TRUE	0.01	0.01	1	0	1	0	1	0	1
21	Citi	Custom Cash Home	0	0	0	0	0	TRUE	0.01	0.01	1	0	1	0	1	0	1
21	Citi	Custom Cash Entertainment	0	0	0	0	0	TRUE	0.01	0.01	1	0	1	0	1	0	1
22	CapOne	Venture X	395	1	1	0	400	FALSE	0.01	0.017	2	0	5	0	10	0	10
23	CapOne	Venture	95	1	0	0	0	FALSE	0.01	0.017	2	0	2	0	5	0	5
24	CapOne	Venture One	0	0	0	0	0	FALSE	0.01	0.017	1.25	0	1.25	0	5	0	5
25	CapOne	Savor	95	0	0	0	0	TRUE	0.01	0.01	1	0	1	0	1	0	1
26	CapOne	Savor One	0	0	0	0	0	TRUE	0.01	0.01	1	0	1	0	1	0	1
27	US Bank	Altitude Connect	95	1	1	0	30	TRUE	0.01	0.01	4	0	4	0	5	0	5
28	US Bank	Altitude Go	0	0	0	0	15	TRUE	0.01	0.01	1	0	1	0	1	0	1
29	Wells Fargo	Autograph Journey	95	0	0	0	50	FALSE	0.01	0.015	3.5	0	4	0	5	0	3
30	Wells Fargo	Autograph	0	0	0	0	0	FALSE	0.01	0.015	3	0	3	0	3	0	3

Figure 19: Exerpt of the file `CreditCards.csv` with all the 38 credit cards shown, but only 4 of the 18 spending categories with multipliers and their corresponding caps.

	Base Value v_b [\$]	Travel Value v_t [\$]
Amex Cashback Cards	0.01	0.01
Amex Membership Rewards Cards	0.01	0.02
Chase Cashback Cards	0.01	0.01
Chase Sapphire Preferred	0.0125	0.026
Chase Sapphire Reserve	0.015	0.027
Bank of America (all cards)	0.01	0.01
Citi Strata Premier	0.01	0.015
Citi Cashback Cards	0.01	0.01
Capital One Miles (Venture Cards)	0.01	0.017
Capital One Cashback (Savor Cards)	0.01	0.01
US Bank (all cards)	0.01	0.015
Wells Fargo (all cards)	0.01	0.015

Table 2: Base and travel point values for various types of rewards credit cards. Note the elevated base values for the Chase Sapphire travel cards, which assumes redeeming the points in Chase’s travel portal. Source: [Nerdwallet \(2024\)](#).

Table 3 shows how the mapping between credit card categories and the items from the CES was performed, for the average consumer in 2022 with an income (before taxes) of \$94,003. Finally, Table 4 shows the credit card spending budgets (as percentages of gross income), separated by income bins.

Category	Expenditure [\$]	%	Consumer Expenditure Survey Items
Everything else	7,786	20.18	Household operations, Vehicle finance charges, Maintenance and repairs, Vehicle insurance, Medical services and supplies, Reading, Tobacco, Miscellaneous
Groceries	6,362	16.49	Food at home, Laundry and cleaning supplies, Other household products
Dining	4,222	10.94	Food away from home, Alcoholic beverages
Gas	3,120	8.09	Gasoline, other fuels, and motor oil
Utility	3,117	8.08	Utilities, fuels and public services
Home improvement	2,606	6.76	Household furnishings and equipment
Online shopping	1,881	4.87	50% Apparel and services, Pets, toys, hobbies, and playground equipment
Drug store	1,481	3.84	Drugs, Personal care products and services
Travel (other)	1,460	3.78	50% Other lodging, 50% Vehicle rental, leases, licenses and other charges, 50% Public and other transportation
Phone	1,431	3.71	Telephone services
Streaming	1,020	2.64	Audio and visual equipment and services
Department store	973	2.52	50% Apparel and services
Entertainment	833	2.16	Fees and admissions
Cable internet	698	1.81	Other entertainment supplies, equipment, and services
Hotel (portal)	644	1.67	50% Other lodging
Airline (portal)	423	1.10	50% Public and other transportation
Car rental (portal)	394	1.02	50% Vehicle rental, leases, licenses and other charges
Office supplies	128	0.33	Postage and stationery
Total	38,576	100.00	

Table 3: The average credit card budget as derived from the 2022 BLS Consumer Expenditure Survey. The corresponding mean income (before taxes) is \$94,003, showing that, with this budget, 41 percent of the gross income can be spend on credit cards. The first column lists the 18 credit card categories that are used throughout this project, while the last column shows how the expenditure items were mapped to each of the credit card categories.

	Expenditures as percentage of gross income									
	All consumer units	Less than \$15,000	\$15,000 to \$29,999	\$30,000 to \$39,999	\$40,000 to \$49,999	\$50,000 to \$69,999	\$70,000 to \$99,999	\$100,000 to \$149,999	\$150,000 to \$199,999	\$200,000 and more
Everything else	8.28	46.26	16.61	15.49	12.55	11.05	9.66	7.92	6.50	4.86
Groceries	6.77	55.85	17.33	13.17	11.73	9.89	7.97	6.31	5.23	3.16
Dining	4.49	22.10	8.02	6.52	6.42	5.65	4.57	4.58	4.15	3.09
Gas	3.32	19.86	7.71	6.61	5.66	5.17	4.20	3.37	2.62	1.46
Utility	3.32	24.30	10.08	7.72	6.25	4.96	3.79	2.94	2.37	1.50
Home improvement	2.77	12.41	5.16	5.07	3.74	3.27	2.90	2.64	2.65	1.96
Online shopping	2.00	5.61	3.89	3.11	2.96	2.56	2.42	1.85	1.75	1.24
Drug store	1.58	7.92	3.95	2.88	2.63	1.32	1.84	1.35	1.24	0.80
Travel (other)	1.55	3.27	2.01	2.09	1.65	1.55	1.30	1.40	1.41	1.57
Phone	1.52	9.85	4.14	3.20	2.88	2.29	1.85	1.51	1.19	0.62
Streaming	1.09	6.81	2.98	2.37	1.84	1.43	1.29	1.06	0.85	0.53
Department store	1.03	0.00	1.77	1.29	1.62	1.29	1.24	1.00	0.84	0.68
Entertainment	0.89	1.90	0.65	0.82	0.69	0.84	0.79	0.80	1.10	0.91
Cable internet	0.74	0.94	0.47	0.00	0.33	0.00	0.00	0.52	0.00	0.74
Hotel (portal)	0.68	0.00	0.72	0.81	0.63	0.60	0.52	0.58	0.60	0.79
Airline (portal)	0.45	1.90	0.61	0.64	0.50	0.44	0.37	0.40	0.39	0.46
Car rental (portal)	0.42	1.38	0.68	0.64	0.52	0.52	0.41	0.43	0.42	0.32
Office supplies	0.14	1.15	0.00	0.26	0.24	0.21	0.15	0.12	0.11	0.06
Total	41.04	221.52	86.78	72.68	62.85	53.02	45.28	38.78	33.42	24.77

Table 4: The average credit card budget for different income bins, as derived from the 2022 BLS Consumer Expenditure Survey. Note that consumers with an income less than \$15,000 spend more than 200 percent of their income, which is likely explained by receiving gifts or loans. This income bin will therefore be ignored in the analysis.

B. Code Appendix

In this appendix, I present the R code of the two functions (`get_buget()` and `get_portfolio()`) that represent the core of this project, and that were needed by all the analysis presented in Sect. 6.

Both functions are saved in the file `CCPortfolioFunctions.R`.

```
get_budget <- function (income = "avg", budget_data) {  
  # Given an income (before taxes), this function returns a named  
  # vector with the spend per category.  
  # "avg" is accepted as a special input argument to return the  
  # average budget from the Bureau of Labor Statistic's  
  # Consumer Expenditure Survey (2022), corresponding to an average  
  # income of $94,003. If a dollar amount is given as input instead,  
  # the function returns the average budget corresponding to the  
  # closest of 9 income bins.  
  
  if (income == "avg") {  
    budget <- t(budget_data[2:19, 2])  
  } else if (income < 15000) {  
    budget <- t(budget_data[2:19, 5] * income)  
  } else if (income >= 15000 & income < 30000) {  
    budget <- t(budget_data[2:19, 7] * income)  
  } else if (income >= 30000 & income < 40000) {  
    budget <- t(budget_data[2:19, 9] * income)  
  } else if (income >= 40000 & income < 50000) {  
    budget <- t(budget_data[2:19, 11] * income)  
  } else if (income >= 50000 & income < 70000) {  
    budget <- t(budget_data[2:19, 13] * income)  
  } else if (income >= 70000 & income < 100000) {  
    budget <- t(budget_data[2:19, 15] * income)  
  } else if (income >= 100000 & income < 150000) {  
    budget <- t(budget_data[2:19, 17] * income)  
  } else if (income >= 150000 & income < 200000) {  
    budget <- t(budget_data[2:19, 19] * income)  
  } else if (income >= 200000) {  
    budget <- t(budget_data[2:19, 21] * income)  
  }  
  
  budget <- as.numeric(gsub('[$,]', '', budget))  
  names(budget) <- budget_data[2:19,1]  
  return(budget)  
}
```

```

get_portfolio <- function (K = 4, eta = 0, theta = 0, cards_data,
  budget, verbose = TRUE) {
  # Given K (number of cards), eta (fractional use of travel point value),
  # theta (fractional use of benefits), the cards data (CreditCards.csv
  # in the form of a dataframe), and a budget (from get_budget()), this
  # function returns a named list, containing:
  # - the optimal credit card portfolio
  # - the total net benefit [$]
  # - the marginal benefit per card
  # - the return on spend after each additional card [%]
  # - the card assignments for each spend category
  # - the total spend [$]

  #####
  # Initialization
  #####

  total_spend <- sum(budget)
  # Get the category names (are sorted by highest average spend)
  categories <- names(budget)

  # Specify benefits that should only be valued once (do not include
  # 'benefit_credits', which will be valued per card and are assumed
  # to be stackable)
  single_use_benefits <- c("benefit_globalentry", "benefit_lounge",
    "benefit_clear")
  # Benefit values
  ge_value <- 20      # Global Entry / TSA pre value in annual dollars
  lounge_value <- 40  # Lounge Access in dollars per year
  clear_value <- 189  # Clear Credit in dollars per year
  single_use_benefits_values <- c(ge_value, lounge_value, clear_value)
  benefit_value <- sweep(cards_data[, single_use_benefits], 2,
    single_use_benefits_values, '*')

  # Initialize the category spend value matrix
  cat_value <- matrix(0, nrow = length(cards_data[,1]),
    ncol = length(categories))

  # The portfolio that contains the selected cards in order
  cards <- c()
  # The net and marginal benefit after adding each card to the portfolio
  net_benefit <- c()
  marginal_benefit <- c()
  # Named vector that tracks which card to use for every category
  card_assignments <- c()

  #####
  # Calculate Value of Spend per Card and Category (matrix)

```

```
#####
# K is the number of cards in the portfolio , while
# N is the number of cards in the dataset.

for (n in 1:length(cards_data[,1])) {
  for (c in 1:length(categories)) {
    cat <- categories[c]
    cap <- paste0(cat, "_cap")

    # weighted average of point value , using eta
    point_value <- eta * cards_data[n, "travel_value"] + (1 - eta) *
      cards_data[n, "base_value"]

    if (cards_data[n, cap] == 0) {
      # No spending cap, multiply spend * multiplier * value as usual:
      cat_value[n, c] <- budget[[cat]] * cards_data[n, cat] *
        point_value
    } else if (budget[[cat]] <= cards_data[n, cap]) {
      # There is a cap, but we stay below it:
      cat_value[n, c] <- budget[[cat]] * cards_data[n, cat] *
        point_value
    } else {
      # We spend more than the cap, so fall back to 1x multiplier
      # beyond the cap
      cat_value[n, c] <- (cards_data[n, cap] * cards_data[n, cat] +
        (budget[[cat]] - cards_data[n, cap]) ) *
        point_value
    }
  }
}

#####
# Select the Optimal Portfolio
#####

for (k in 1:K) {
  # Net benefit per card
  net_benefit_per_card <- rowSums(cat_value) +
    theta * (rowSums(benefit_value) + cards_data[, "benefit_credits"]) -
    cards_data[, "fee"]

  # Pick the card with max net_benefit_per_card:
  max_ind <- which.max(net_benefit_per_card)
  # Look up that card's name
  cardname <- cards_data[ max_ind , "name"]
  # Add the selected card to the portfolio
  cards <- append(cards, cardname )
  # Store the cumulative net benefit and marginal benefit
}
```



```

if (k == 1) {
net_benefit <- append(net_benefit, net_benefit_per_card[ max_ind ])
} else {
net_benefit <- append(net_benefit, tail(net_benefit, n=1) +
net_benefit_per_card[ max_ind ])
}
marginal_benefit <- append(marginal_benefit,
net_benefit_per_card[ max_ind ])

# Assign spending categories with additional value to the card:
card_assignments[categories[cat_value[ max_ind ,] > 0] ] <- cardname

# Subtract the selected card's values from the value matrix
cat_value <- sweep(cat_value, 2, cat_value[max_ind, ])
# And set negative categories to zero,
cat_value[cat_value < 0] <- 0
# Now we're left with positive categories that have remaining value
# that we can retrieve with additional cards in the next iteration.

# Multiply all single-use benefits with (1 - single_use_benefits)
# of max_ind, to set them to zero after the first time they
# have been selected
benefit_value <- sweep(benefit_value, 2,
unname(unlist(1 - cards_data[max_ind,
single_use_benefits])), "*")

# If the card can only be held once (Amex Platinum, Citi Custom
# Cash, etc.) we need to make sure it will not be selected again
# for just the benefits alone, or additional rewards in other
# custom categories. This is accomplished by setting all the values
# for the card, including benefits, to zero after selection.
# We use ID to find duplicate cards that can only be held once.
# If cards can be held multiple times (e.g. BoA Customized Cash),
# they will have different ID numbers, if not, the duplicates will
# have the same ID.
same_ind <- which(cards_data[, "id"] == cards_data[max_ind, "id"])
cards_data[same_ind, "benefit_credits"] <- 0
cat_value[same_ind, ] <- 0
benefit_value[same_ind, ] <- 0
}

if (verbose == TRUE) {
# Print output to the screen
print(sprintf("The_optimal_portfolio_with_a_total_benefit_of
$%.2f_is:", tail(net_benefit, n=1)))
print(cards)
print(sprintf("The_marginal_benefits_are:"))
print(marginal_benefit)

```

```

print(sprintf("The_total_spend_is:_$%.0f", total_spend))
print(sprintf("The_return_on_spend_is:_%.2f%%",
100 * tail(net_benefit, n=1) / total_spend))
print(sprintf("Use_the_following_card_assignments:"))
print(card_assignments)
}

# Return a named list
return(list("cards" = cards, "net_benefit" = net_benefit,
"marginal_benefit" = marginal_benefit,
"return_on_spend" = net_benefit / total_spend,
"card_assignments" = card_assignments,
"total_spend" = total_spend))
}

```

References

- Agarwal, S., A. Presbitero, A. F. Silva, and C. Wix (2023). Who pays for your rewards? Redistribution in the credit card market. *Finance and Economics Discussion Series* (2023-007).
- BLS (2023). Consumer expenditure surveys. Technical report, U.S. Bureau of Labor Statistics.
- CFPB (2023). The consumer credit card market. Technical report, Consumer Financial Protection Bureau.
- Hayashi, F. (2009). Do U.S. consumers really benefit from payment card rewards? *Economic Review* 94(Q I), 37–63.
- Nerdwallet (2024). How much are travel points and miles worth in 2024? [nerdwallet.com. https://www.nerdwallet.com/article/travel/airline-miles-and-hotel-points-valuations](https://www.nerdwallet.com/article/travel/airline-miles-and-hotel-points-valuations) (accessed June 8, 2024).
- Paarsch, H. J. and K. Golyaev (2016). *A Gentle Introduction to Effective Computing in Quantitative Research*. Cambridge MA: The MIT Press.
- Schuh, S., O. Shy, and J. Stavins (2010). Who gains and who loses from credit card payments? Theory and calibrations. *Public Policy Discussion Papers* (10-3).
- The Wall Street Journal (2010). Credit cards take from poor, give to the rich. [WSJ.com. https://www.wsj.com/articles/BL-REB-11033](https://www.wsj.com/articles/BL-REB-11033) (accessed June 1, 2024).
- Thrifty Traveler (2024). Is Amex still a bad option when traveling abroad? [thriftytraveler.com. https://thriftytraveler.com/news/credit-card/amex-card-acceptance-abroad/](https://thriftytraveler.com/news/credit-card/amex-card-acceptance-abroad/) (accessed July 24, 2024).