FundRight

Consumer models for applications in factoring

Advisor: Dr. Peter Fader

Team 7

Cristina Amusategui (ESE)

Braden Fineberg (ESE)

Alec Gelfenbein (ESE)

Raouf Tawfik (ESE)

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Nonprofits rely on unpredictable and sporadic donations. As a result, they often have to accept soaring interest rates of up to 50% to obtain a loan. Since the 2008 financial crisis, alternative financing methods such as factoring have become more prevalent. FundRight's vision is for nonprofits, and other organizations with unpredictable revenue streams, to obtain an immediate loan in exchange for their revenue for a predetermined period of time. In the non-profit space, FundRight uses factoring to model cash flows and output the number of days that would be required to re-capture that amount.

FundRight created a stochastic consumer spending application based off of a dataset containing donor specific information in addition to donor frequency and amount. After iterating through several different models, including the pure Negative Binomial Distribution (NBD), the Pareto NBD, the Beta Gamma Beta Binomial (BG/BB) Model and the Honeymoon Period Effect, a final model was created. The selected model is a variation of a pure NBD model with the implementation of various cohorts, specifically time joined, and segments, specifically number of cards.

On a forty loan portfolio, FundRight was accurate 87% of the time, meaning that 87% of the simulated initial amounts given to nonprofits were fully captured back within the predicted time frame. The key to FundRight's accurate predictions is the ability to minimize inter-run variance. FundRight was able to tighten the confidence bounds around its loan-modelling by simulating a variance of 12%.

FundRight's model is not only accurate, but also fills a big gap in the market in terms of factoring in the non-profit space. This project aims to change how banks measure nonprofit risk and enable nonprofits to have confidence in their financial future. By adding time dependency to customer lifetime value, FundRight's model is able to project forward revenue over time.

Overview of Project

Just like for profit organizations, nonprofits are businesses with assets and liabilities. Unlike the majority of for profits however, nonprofits have unpredictable revenue streams. The lumpy, uncertain donation streams of nonprofits is a result of an extremely high churn of donors - only 1% of those who give become repeat donors - and external donation motivations such as tax credits or corporate matching programs. In addition, it's extremely hard for a nonprofit to identify which supporters will become donors and which supporters will remain just volunteers.

It is exactly because of this that nonprofits have been wary of obtaining bank debt. Though they benefit from an increased capacity to finance various projects around the world, levered nonprofits are not only burdened with interest rates that are upwards of 30% per annum, but are also restricted by the strictest of bank covenants. Unsurprisingly, this structural misalignment between lender and borrower has strongly disincentivized nonprofits in need of financing from borrowing in the first place.

Ever since the financial crisis of '08, there has been a drive towards exploring alternative financing methods including factoring, the process through which a lender provides a borrower with the requested financing amount in exchange for capturing the latter's revenue for a pre-specified period of time. In an attempt to fill the current market gap, FundRight leverages this framework to model a given nonprofit's cash inflows over time, outputting the expected number of days to re-capture that given amount.

By partnering with Flourish, an organization that allows individuals to donate to the nonprofit of their choice through either round up or fixed amount transactions, FundRight has access to a micro-donation database containing valuable information about thousands of users and their behavior. We then aggregate this data – splitting it by certain commonalities and trends across users – to determine nonprofit specific trends, including repayment patterns. Though our models are currently proprietary to Flourish, our results are applicable to - and can be used by -other players in the factoring industry too.

Business Analysis

Problem and Need

Nonprofits in the US are very skeptical of bank loans. Primarily, this is attributed to their inability to predict donation revenue to repay the loan that they have taken. In addition, banks are highly skeptical when giving loans to small nonprofits for this very reason. As a result, the most common fundraising method of nonprofits is to request large sums from a large donor. While this works relatively well, the large funds often contain spending limitation or clauses that make it difficult to deploy the money where it is needed most. In the last 6 months, Flourish Technologies, has collected a donation dataset that may contain the solution to this problem. Flourish is a micro-donation platform processing thousands of recurring donations per day with an average size of just under $0.57. These small, frequent transactions would, in theory, make predicting donor revenue much easier. If donor revenue could be predicted, then making loans would be a lot less risky because its likelihood of being repaid is supported by a revenue model.

From our knowledge in the nonprofit and consumer spending model space, a model like ours has never been built before. We are very fortunate to have the data to enable us to create such a model. We believe that our model will enable interest-free loans to nonprofits, will has the potential to create a big shift in the nonprofit space, since it will remove all risk which previously fell on them, onto the intermediary platform, Flourish. Therefore, we believe that our models are of interest to both nonprofits and Flourish. Furthermore, our hope is that in the future, a bank could factor organizations' donations for a time period comparable to receiving low-cost funding.

Value Proposition

FundRight allows for non-profits without stable cash flows to receive low-cost financing and avoid prohibitively high interest rates from banks by using consumer behavior to create an accurate risk profile for these organizations. FundRight will capture an organization's donation stream for a given period of time in return for money up-front. All risk is transferred to FundRight, allowing the nonprofits to focus on achieving their missions to improve the world.

Stakeholders

The biggest stakeholder for FundRight are small nonprofits. They are our target segment and will receive great benefit from our funding strategy. It will reduce their reliance on high interest bank loans. Our hope would be for nonprofits to no longer have to take on high interest rates, but rather give their donations for a certain period of time to an intermediary platform or bank in exchange for an immediate, interest free loan.

Banks are also stakeholders, potentially losing business from nonprofits. If banks are unable to decrease their interest rates, then nonprofits will no longer turn to them for loans. However banks are so large that they will likely not feel a large effect on losing this segment of their business. That being said, we believe that banks could also benefit from our model and use it in order to give out loans in exchange for collecting a nonprofit's donations for a period of time. Our model could provide banks with a new way to give out loans, one where them and the nonprofit win.

Market Size and Growth

The majority of donations are given by individuals, summing to around $290B[1], which although large, are very unreliable. Nonprofits retain only 43%[2] of those who give become repeat donors, increases the riskiness of relying on donations. Younger generations are donating even more spontaneously, however although donations are small and random, 84% of Millennials donated[3] to charity in 2014. WIth the emergence of micro donation platforms and the millennial generation going into the workforce in the next few years, there is a gap in the market for predicting small and spontaneous donations. Our model is aimed towards facilitating the loan process for nonprofits. A nonprofit will be able to use this model to predict how long it will take them to collect the sum needed in terms of donations. This could be used independently as a way to budget for the future, or by banks as a way to give out loans in return for collecting a nonprofit's donations for a period of time.

Nonprofits Quarterly estimates that nonprofits pay around $16B per year in interest on $46B in short-term loans. This is extremely expensive for small nonprofits that do not benefit from endowments, long credit histories or reliable projections. A bank will often make up to 40% on interest. By using our model, we believe we can offer cheaper rates to the nonprofits market. Our method of factoring donations instead of offering loans, removes signficicant risk from the nonprofit. According to the National Philanthropic Trust, the donation market is a total of $410B. This market is made up of individuals, foundations, and corporate organizations who donate. The sum of individual donations is around $290B. If we assume that in the coming years there will be 90M millennials who donate in smaller and more random sums than older generations. Using historical donation propensities and lifetime salary projections, we believe that the individual donations market will grow by up to $22B reaching the 1.5M nonprofits.

Our models are designed to target a very specific segment of the overall nonprofit market. We believe that our ideal nonprofits have an annual operating budget of less than $5MM, with little or no endowment, employing 2-8 full time employees, and have grants or large gifts to cover annual fixed costs. This description represents about 66% of the nonprofits across the country, including approximately 900k organizations.

---

[1] https://givingusa.org/giving-usa-2018-americans-gave-410-02-billion-to-charity-in-2017-crossing-the-400-billion-mark-for-the-first-time/

[2] https://www.causevox.com/blog/donor-retention-statistics/

[3] https://www.thestreet.com/story/14445741/1/why-millennials-are-more-charitable.html

Customer Segment

We believe that our key customer segments are small nonprofits who have small and irregular donation streams. These nonprofits are at the most risk when taking out a loan because their credit rating is on the line and probability of default is often high. Our model will enable nonprofits to predict donation streams and also take out risk free loans in return for capturing their donation stream for a number of days.

Additionally, we believe that banks could use our model in order to give out loans to nonprofits. Rather than using the method of giving a traditional loan with a required interest and collateral, they could use our model to give a loan to a nonprofit in return for securing their donations for a period of time.

Competition

Currently, there are no models or companies on the market that predict donations with the goal of giving out loans in return for collecting donations for a certain number of days. Our model is unique in the sense that it does not require any amount to be paid back for a loan, removing interest rates, possibility of default of any collateral that could place financial stress upon a nonprofit. We take on all of the risk of giving out a loan and transfer it away from a nonprofit. The only competition that could detract nonprofits from using our model would be banks who give money via loans. However, due to the removal of risk in our model away from the nonprofit, we do not anticipate banks being a huge threat or taking market share away from us due to our differing business models.

Revenue Model

We would seek to target a SaaS revenue model requiring monthly or yearly payments in exchange for using our software. Most likely, we would target banks or other businesses equipped to loan money. The bank would then be able to offer a new service to non-profits to factor their donations instead of receiving a loan. We would pursue this model instead of launching FundRight as a standalone business because of the upfront capital required to give out these sums of money to nonprofits in high volume. The customer base that would pay for such a software is relatively small in quantity, thus the price of our service would be quite high to produce returns. However, there is high potential for customers to create significant revenues using the FundRight software due to the size of the nonprofit market and lack of opportunity for nonprofits to stabilize their cash flows via traditional methods.

TECHNICAL DESCRIPTION

The problem in the nonprofit, short-term debt market exists because large banks use for-profit models to price non-profit risk. This leads to incredibly expensive interest rates. Fundamentally, mis-pricing nonprofit risk is a result of incorrect data; therefore, we started from first-principles. The core values of statistics indicate that to increase certainty, one needs to increase the number of samples to find the underlying pattern. As data increases, we are able to correctly understand and model these patterns.

DATA

To begin, we looked at how nonprofits currently raise money and what data they have. Starting with a dataset from DonorsChoose.org provided on Kaggle that contained information including donation amount, donor IDs, and time as well as several explanatory variables such as donor city, state, and zip. We very quickly realized that this is a fairly common dataset and contains information that doesn't truly impact one's ability to predict future donations (revenue), a key tenet of pricing a loan. After looking at the responses from the Kaggle competition, it appears that while one is able to determine whether a known repeat donor will give again, there is no discernible pattern as to whether any randomly pulled individual will give. Fundamentally, we needed better data.

When looking for the correct information to price risk, we spent a considerable amount of time studying the market to ensure that we are obtaining the right amount of data in the first place. Using first principles in statistics, we looked for a dataset that contained highly repetitive small donations that were reliable. Today, Flourish uses round-up donations to raise money for organizations. A round-up is when a user goes to starbucks, for example, and pays $3.50 for a latte; the transaction is "rounded up" to $4, and the $0.50 is given to the donors chosen organization. This increase in donation frequency, and decrease in donation size, allows for much more predictable patterns. Ultimately, we are able to model consumer behavior, not donations. This slight change in thinking and modeling allowed us to build models with unparalleled accuracy. We are building nonprofit models to price nonprofit risk.

SPECIFICATIONS AND REQUIREMENTS

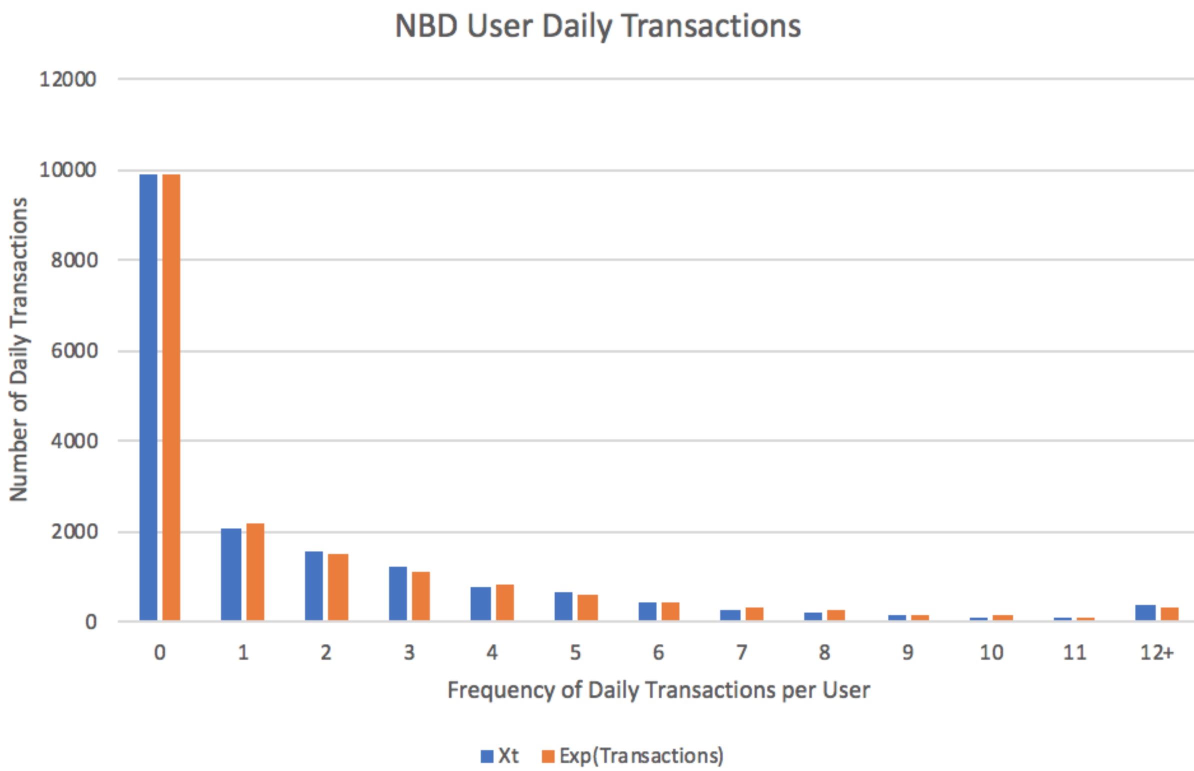
As mentioned earlier, constraining inter-run variance is key to FundRight's success. As such, our initial goal was to cap this variance number to less than 10%. Additionally, our model incorporates a minimum ROI per loan of 10%. While we expect to dump this constraint in the future, we incorporated this in the current iteration as a way to hedge the likelihood of losing money. This is because we expect the initial loan portfolio size to be small. As more nonprofits begin adapting this financing method, the number of loans in our portfolio will increase.

Given that our project has a budget of $0, we did not have to specify any cost constraints.

ITERATIONS CONSIDERED

Negative Binomial Distribution

Our conversations with Professor Fader drove us to initially look at the pure negative binomial distribution model as a starting point. Interestingly, the pure NBD ended up being the final model structure FundRight used. This is because alternative models explored did not tend to fit our data just as well.



Pareto NBD

More specifically, models which allow for non-stationarity, such as the Pareto NBD, did not work because they added a time dependency element to counting. While our data did in fact exhibit seasonality of spending patterns over time, the frequency of spending changes did not change over time for everyone in aggregate. *See Appendix A for Pareto NBD model output.*

Beta Gamma Beta Binomial (BG/BB) Model

The BG/BB model also failed to accurately predict donor behavior over time due to its heavy emphasis on the death process, which in this case, cannot be accurately estimated due to data characteristics (4 churns in Q1 of 2019, for example). The "death process" models the likelihood that any given donation is a donor's last. In this application, it is a stochastic process that is a function of the recency and amount of the donor's last donation.

Honeymoon Period Effect

We also tried to implements a honeymoon period in the model. This effectively shifts the death process from the beginning of a customer's lifetime by a certain period. The logic is that a user is less likely to churn in their first or second usage of a product, before they begin to think about cancelling a subscription. We quickly realized that the Honeymoon period effect was not applicable to our model due to the "set it and forget it" attitude most users tend to have regarding their round-up donations. The likelihood of someone opting-out of donating at any point was simply too low.

Societal and Economic Considerations

The key societal consideration that influenced our design is the constantly rising interest rates that banks give to nonprofits. We believed it was not correct for a bank to be making millions from giving out loans for beneficial causes to nonprofits. Our model, although it does provide some form of profit to the lender, is far less extreme. In order to avoid our model enabling someone to make a large ROI on a loan, we plan on putting a cap on each loan that is given out. For example, if the number of days to recapture the loan is 120 but at day 100 the person has already made 20% ROI, we will cap it and stop collecting donations.

Furthermore, the volatile and random economic environment of donations is what made us curious to research the donation space and see if there was any way that we could understand and predict the randomness of donations. By having access to a random, micro donations dataset, we were able to further understand the randomness of donations and interpret it in such a way that makes it slightly less random.

Technical Description and Approach

As hinted at, our model is essentially a pure NBD model with various cohorts and segments applied. While cohorts are formed by pooling groups of users based on time joined, segments are groups of individuals who display the same behavior. In this case, behavior was proxied by the number and type of cards a given user holds.

Cohorts

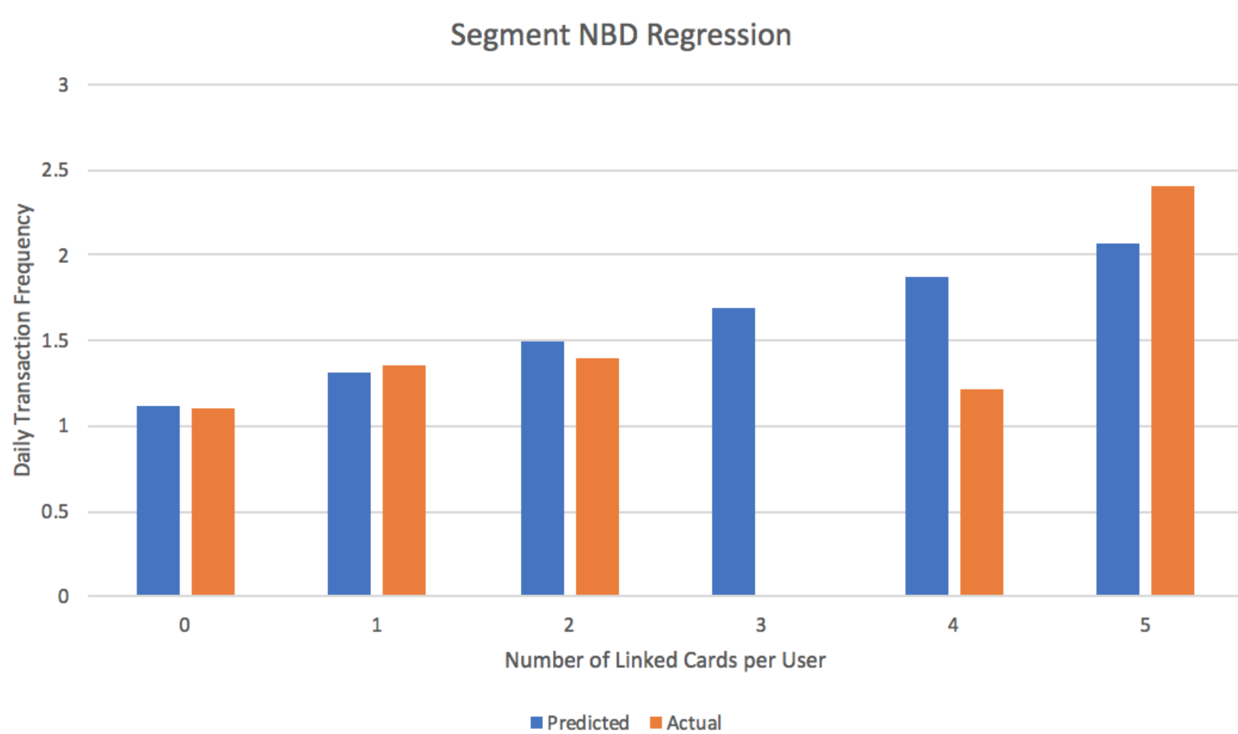
The assumption behind grouping people in different cohorts (a total of four) is that data from older cohorts better (and more accurately) reflects users' behaviors relative to that from newer cohorts. Overweighting older cohorts in our model therefore discounts the information from more recent cohorts simply because we don't know these users' behavior; the more recent the cohort, the greater the discount rate. Conversely, we purposely overestimate the presence of older, more predictable users as a means to extend and apply their behavior to users who have just joined. See *appendix B* for NBD cohort parameters.

Segments

On the other hand, pooling users in different segments (a total of five) relies on the premise that similar users have similar spending patterns and vice versa. The goal then becomes to find a common similarity between different users that is mutually exclusive yet comprehensively exhaustive, meaning that each user fits in one and only one bucket. Initially, we thought age would be a good fit. We shortly realized that an individual's age does not influence their average number of transactions over a period of time. The number of cards an individual holds, however, strongly matters as shown in *appendix C*. Generally speaking, the frequency of transactions increases with the number of cards.

NBD Regression

To confirm and test the validity of our segments, we ran an NBD regression on a table containing users grouped by the number of credit cards. The NBD regression was a natural choice as we came to realize that our data had an amazingly long tail, indicative of over-dispersed count data. Please note the regression excludes individuals without any tracked cards. With a theta of 0.243, the regression equation takes the form of y = 1.12 + 0.19x. This indicates that for every additional card, a user's transaction increase by 0.19 per card they linked to their donation patterns. This means that the average number of transactions per user is 1.3 transactions a day. However, we have several users with 5 tracked credit cards, indicating a transaction count of just over 2, a 53% increase over the mean.



Covariates

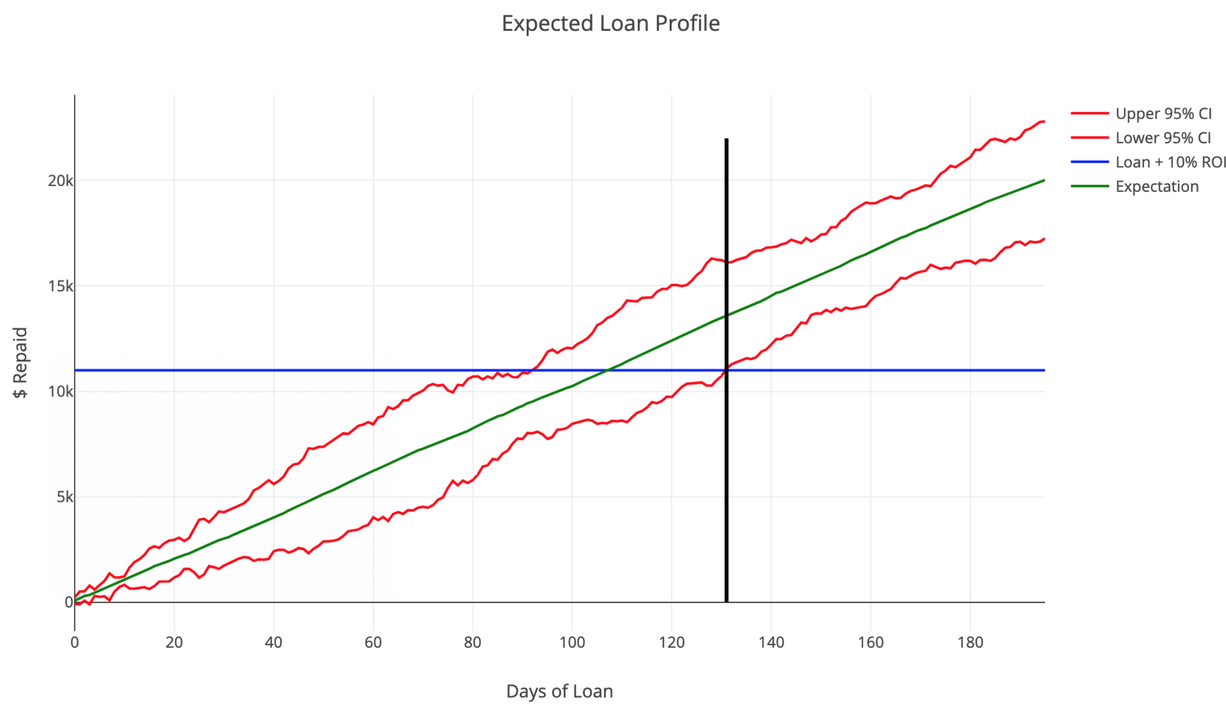
Round-up donations are considered passive in the sense that a user “sets it and forgets it.” Therefore, the covariates we used tended to be externalities to individual user behavior. Instead, we focused on gender, age, location (urban, rural), and date as our covariates. We found that gender and location were crucial in determining the number of roundups a user has in one day. Qualitatively, this can be explained as a result of transaction opportunity. Living in a big city, as a young working man, you likely swipe your cards more than 2.5 times per day, on average; the first time is a coffee in the morning, before getting lunch, and maybe drinks after the workday. In a rural setting, there aren’t as many opportunities to swipe a card. While we found that this covariate pair had good explanatory value, seasonality was much more important. People roundup less frequently during the week simply because they spend more time in the office. Weekends, holidays, and between Thanksgiving and New Years is when the greatest number of roundups are processed. Combined, these covariates were able to assist in predicting future donation revenues with a high degree of confidence.

Confidence Intervals

As we are factoring donations, confidence is much more important than expectation. There is a large difference between offering to factor donations for 100 days ± 20 days with 95% confidence and factoring for 100 days ± 5 days with 95% confidence. Therefore, to consider our model successful, we aimed to have our 95% confidence interval contain less than 10% of the total loan period. While this was a good method in theory, we realized when running practical simulations, there are uncontrollable events for which we cannot account such as large giving campaigns, or donor fatigue from unrelated solicitations. In practice our confidence interval was closer to 12% of factor period instead of 10%. In order to further have confidence in loan repayment, we used the lower 95% confidence interval in deciding for how long to offer the factor. In our final design, shown below, the blue and black lines intersect when the loan amount plus ROI ($10,000+10%ROI=$11,000) intersects the lower 95% confidence interval. This is done to further hedge the risk of non-repayment from the nonprofit. We found that using these methods generated a 87% success rate on a simulated 40 loan portfolio.

Final Status of the Project and Test Results

On a simulated 40 loan portfolio ($10k/loan), our model was accurate 87% of the time, meaning that FundRight captured back the entirety of the loan over the specified period. During the simulated 12 year loan portfolio, our $400k investment became $453,782.90, generating an IRR of 13.45%. From our perspective, this represents a successful proof of concept.



As of now, we have delivered this model to Flourish. From what we know, they have begun to dissect the models and methods we used to predict donation revenues. We believe it is their intent to use these models as a basis for factoring donations within the next 12 months.

Self-Learning

This project is a huge synthesis of various skills. Two members of our group had taken STAT 476 (Applied Probability Models in Marketing) with Professor Fader, however the other two had a significant learning curve to overcome in order to understand the models we would be dealing with in this project. As a group we went over several of Professor Fader's papers to become familiar with various of his models: Negative Binomial Distribution[4] (NBD), Pareto NBD[5], BG/BB[6] and other models for consumer base analysis[7]. As well as using material from Professor Fader's class, we all took Professor Vohra's class, ESE 204 (Decision Models) and used various of the concepts learned in our project. To build our transactions table and cash flow projections we used a random walk process and modeled it in the @Risk software. This enabled us to build and validate our expected confidence intervals successfully before finalizing the model in python. Our background in mathematics, particularly statistics through Professor Fader's class, was very helpful. We also had an array of computer programming skills, particularly in Python, amongst our group, which also facilitated the building of the backend of our project.

We were fortunate to have all taken courses that were relevant to our project, however there were definitely aspects that we had to teach ourselves. Before beginning the modeling aspect of our project there was a learning curve in cleaning the data. We had to be very deliberate in which aspects of the data we decided to use for our project and how we would display the data. When we started the modelling, the first obstacle we encountered was building the optimization model in python. Our group had experience building API's and models in python, however none of us had learned how to implement optimization models in python. We spent a significant amount of time getting familiar with CVXPy[8], the convex optimization framework, and SciPy's optimization tool[9]. After a learning curve of a few weeks, the optimization frameworks were implements which then enabled us to move forward with the rest of the project.

The second major learning curve was deployment for production. We initially considered Docker[10] containers and virtual machines as deployment options. Ultimately, we used Google App Engines (virtual machines built out of Dockers) as the final solution due to the simplicity and ease of use. Our priority was to ensure our model could update itself and continue to learn over time so we spent time understanding how we could implement a reinforcement framework. As our dataset continued to grow, it was becoming even more important to find a way to update the existing model rather then starting from scratch each day due to the increasing training time of the growing model. Therefore, to overcome this barrier we build a webapp that ingests and auto-updates the models on a 12 hour interval. These were built using python and django.

---

[4] https://faculty.wharton.upenn.edu/wp-content/uploads/2013/08/JCGS-Bayes-NBD-2002.pdf
[5] http://brucehardie.com/papers/018/fader_et_al_mksc_05.pdf
[6] https://repository.upenn.edu/cgi/viewcontent.cgi?article=1056&context=wharton_research_scholars
[7] http://www.brucehardie.com/talks/ho_cba_tut_art_12.pdf
[8] https://www.cvxpy.org/
[9] https://docs.scipy.org/doc/scipy/reference/optimize.html
[10] https://www.docker.com/

ETHICAL AND PROFESSIONAL RESPONSIBILITIES

Our project is based in the non-profit space, giving it significant potential for positive economic and societal impact. There is a 22 billion dollar market from Millennial spending each year. In the US, there were 14.5 billion credit card transactions in the first months of 2015 alone, totalling over $1.4 trillion in purchase volume. In 2017, there was about $410 billion donated to charitable causes, just a fraction of the US credit card purchase volume in only 6 months. It is clear that while Americans are spending at high rates, charitable donations are being left behind.

Despite high levels of general spending in the US, nonprofits expect to lose 12 billion each year. The current donation structure for charities makes it incredibly difficult to stay in consumers' minds. Donors have to be aware of exactly what charity they want to give to, and take time out of their day to give a donation. Flourish is changing the donation structure for charities by allowing consumers to automatically give each time they swipe a credit card. Even with this model, however, most of a nonprofit's revenue stream is still coming from less frequent donations. With expected losses and unpredictable cash flows, nonprofits have difficulty borrowing money. This leaves them struggling to accomplish their mission.

Our project is seeking to impact the funding structure for nonprofits to enable to achieve their goals. By removing risk and adding stability to their cash flows, nonprofits can more carefully and accurately plan their spending habits. If the 1.5 million nonprofits in the US are able to reduce their financial woes and instead focus on their missions, the impact could be massive. Banks have not focused on finding applicable financing methods for the non-profit sector. This sector has high societal and global impact with the sole goal of improving the world, unlike many other business industries with a focus on a bottom line. FundRight can enable the non-profit sector to increase their positive impact on the world.

Along with our solution comes some risk for ethical issues. We are offering a product that captures donations of a nonprofit for a given period of time. There are a variety of potential ethical issues surrounding this donation structure. The first is that we must accurately model the payoff for offering our product. It would be unethical if the party offering FundRight were to make a huge profit on the donation stream. To prevent this from occurring, our models need to provide ample information to allow the offering party to understand what they are providing for non-profits and how to use FundRight to properly predict donation streams. nonprofits are focused on a goal outside of making money. Thus, their attention is not necessarily always on the bottom line as in other businesses. This could make it easier for FundRight's product to take advantage of the non profits by capturing an unfair quantity of their revenues. It is critical that if implemented, FundRight would be highly transparent in working with nonprofits to ensure that they were not being overcharged.

The second potential ethical barrier is consumer-facing. It is critical that the donation data being processed by FundRight is encrypted to ensure privacy is met for donors. This was accomplished by encoding usernames so that an individual user could be tracked over time without being able to trace that user back to their true identity.

MEETINGS

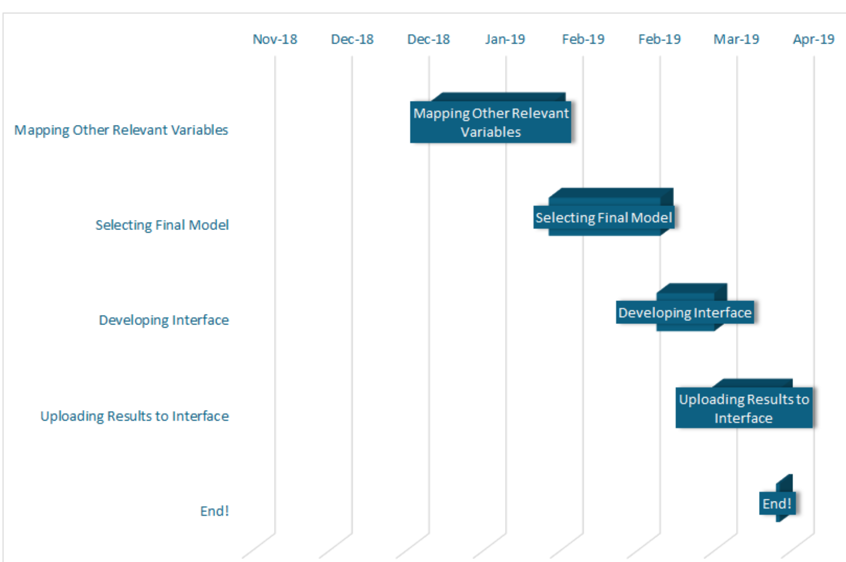
This semester, the team met consistently and routinely. By setting aside a total of four hours every week, we were able to not only meet the expectations we had set for ourselves at the beginning of the semester, but also echo ideas off of each other. This is in contrast to last semester where work was divided up amongst us, and we each ended up working on our individual sections.

As well as meeting as a team, we kept an open communication channel with Professor Fader, making sure that he is fully aware of what it is we are working on at any given moment. Our group met with Professor Fader around twice per month and also informed him of key updates via email. Furthermore, we met with Professor Vohra when we encountered an obstacle during our simulation implementation.

SCHEDULE WITH MILESTONES

As originally planned, the majority of this semester was spent building and testing different models. After reaching a consensus about using the pure NBD model in March, the team started building the UI interface as well as running various different simulations with test data. All of the expectations we had set out for ourselves at the beginning of the year were successfully met ten days before the final competition, giving us more than enough time to work on the pitch and on final information delivery.

Although there are no unmet milestones, with more time, we would have liked to test our model on a wider array of loan portfolios as well as spoken to more nonprofit organizations in order to gauge their perspective on our model. Furthermore, our model is unable to account for low probability high impact events, Black Swans. Although it is not vital to account for these right now, if given more time, we would have liked to see the inclusion of Black Swans on the accuracy of our model.



Discussion of Teamwork

Overall our team has been highly collaborative in every aspect of the project. Our frequent team meetings allow us to all work on each part of the project to gain broad experience in building a project. Although we are all SSE, we have some differences in skill set that have brought us to concentrate more heavily on specific areas of the project. Braden and Cristina both have some prior experience with consumer models after taking MKTG 776 with Dr. Fader. Alec and Raouf both have summer internship experiences in finance. Leveraging the varying skill sets within our group has led to the following breakdown in responsibilities:

Individual Responsibilities:

Braden:

- Primary correspondent with Dr. Fader
- Data collection
- Testing the NBD and refining it with available data.
- Implementing final model in Python
- Building final web applications

Alec:

- Non-profit & market research
- Debt/loan research
- Note taking at meetings
- NBD Regression

Cristina:

- Researching different consumer spending models
- Testing the Pareto NBD refining it according to our data.
- Weekly email update with reminders

Raouf:

- Financial viability research
- Legal research
- Weekly Meeting Organization
- NBD Regression

BUDGET AND JUSTIFICATION

Our project entirely relies on software we already have. We did not spend any funds.

STANDARDS AND COMPLIANCE

In dealing with any financial data, US GAAP standards automatically are in affect. GAAP is an accounting standard that dictates how money is handled and represented in the system. For example, it prohibits publishing/reporting individual transactions. Instead, we must only report monetary aggregates, which in turn, ensure user confidentiality. Given that model only outputs (i) aggregate amounts that can be given to particular non-profits and (ii) the amount of time required to recapture initial money back, we have ensured that the individual transactions only appear as part of the inputs to our model (and not the output). Additionally, we do not anticipate to have to store any data (such as transaction number or user ID) that would enable us to link a particular monetary amount with a specific user or transaction.

In addition, PCI compliance is a set of security standards designed to ensure that all companies that store or process credit card information maintain a secure environment. We do not anticipate PCI compliance to be a concern as our data is encrypted and cannot be traced back to a specific user, though we can track a particular user across time. Similarly, we must keep GDPR standards in mind. Though donation data is not currently incoming from outside of the US, this could become the case at any time. If this occurred, we would be required to comply with GDPR rulings with privacy. Again, we handle this by anonymizing our data.

WORK DONE SINCE LAST SEMESTER

As a group, we are extremely proud of the progress we have made since last semester. As detailed in the technical description, we implemented an array of segments, cohorts, covariates, confidence intervals as well as considering a variety of different models. By testing out all these different functionalities we were able to come up with our final NBD model. Once our model had been finalized we trained and tested it on a simulation of 40 loan portfolios.

Aside from modelling, we spent a significant amount of time carrying out market research in order to further confirm the viability of our model. We took time to consider the tax implications, which as of now are nonexistent. We also delved deeper into the ethical considerations our model could present. Market research enabled us to assess product viability and understand the need for a tool such as ours. Initially, we had planned for our model to be used in a very specific scenario with a specific dataset, that of Flourish. However, upon completing more market research we realized a wide variety of applications it could have. For example, a bank could use our model to provide financing in factoring applications. Therefore, the significant amount of time we spent on market research enabled us to see the widening gap in the market for a took such as ours.

Furthermore, a large portion of our time was used for UI implementation. We created a website which enables a user to request a loan amount, and outputs the number of days that a bank or intermediary platform would need to recapture that amount.

Discussion and Conclusion

On a forty loan portfolio, FundRight was accurate 87% of the time, meaning that 87% of the simulated initial amounts given to nonprofits were fully captured back within the predicted time frame. An average loan within that portfolio had an ROI of 13%. To put this into perspective, that is an inherent interest rate 17% lower than that a typical nonprofit would have received from a bank.The key to FundRight's success is being able to minimize its inter-run variance. That is because while the expectation measure determines the inherent interest rate the organization receives on any given amount, the variance measure is a direct gauge of the likelihood of losing money. With a simulated variance of 12%, FundRight was able to tighten the confidence bounds around the financing it provides.

The narrowness of these bounds is a direct consequence of incorporating seasonality as well as effective user segmentation. People spend more (and more frequently) during the holiday month of December than they do in March. On one hand, having a model that is able to replicate this sort of seasonal spending behavior enables our simulations to be consistently more aligned with reality.

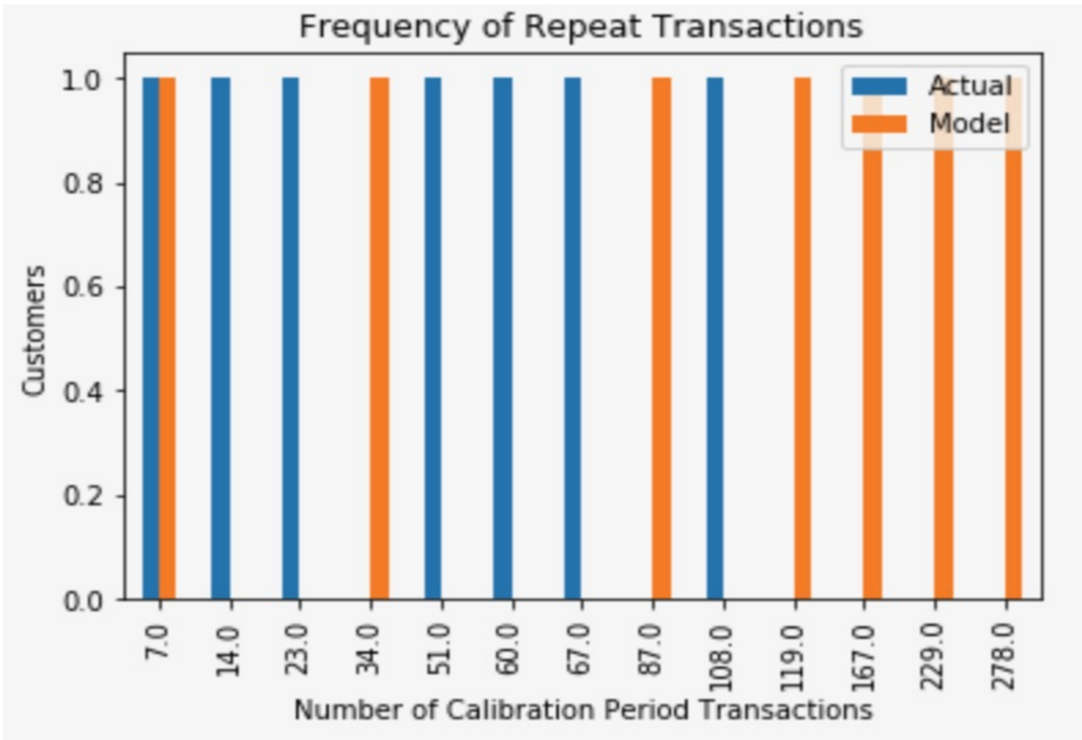
On the other hand, the segmentation process – performed by pooling users with the same number and type of credit cards – enables us to assess the different user composition of each organization. While one would have initially expected a uniform distribution of segments across nonprofits, it turns out that different groups of people have different tendencies. Conversely, similar groups of people have similar tendencies, meaning that more often than not, there is a cluster of users who donate to the same organization from a particular segment.

Though our model is trained to predict highly recurring behavior, it is unable to account for low probability high impact events, colloquially known as Black Swans. These occurrences are semi-impossible to predict, despite having an enormous influence on the final output. (Think of a donor winning the lottery and choosing to give $1 million to a nonprofit, for instance.) Luckily however, integrating these sorts of predictions is not vital at the moment as they mostly give the lender increased upside (and not downside) exposure. In essence, Black Swans are probabilistically more likely to decrease than increase time to repayment. Looking forward, it would be beneficial to model the extent to which these Black Swans affect the accuracy of our model.

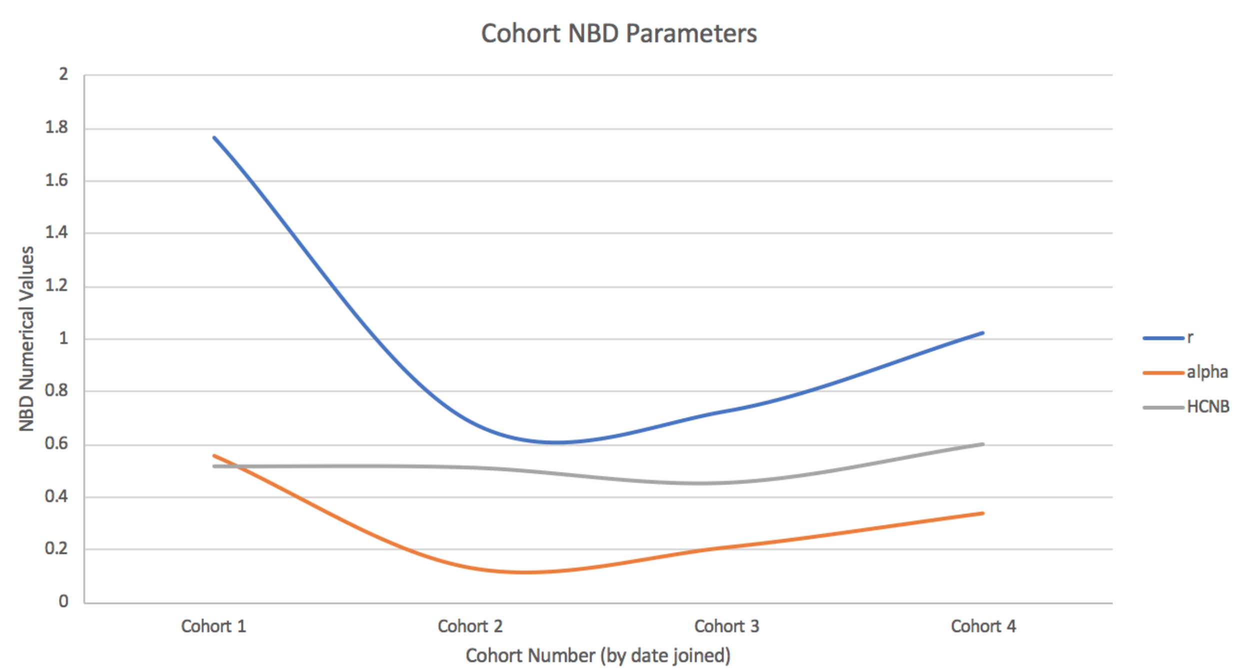
Overall, this project was a success not just in terms of producing a model that spits out an accurate output, but also in terms of being able to identify and fill a sizeable market hole. FundRight's presence enables nonprofits to view debt in a different light and value their customer base not just in monetary amounts, but also in terms of their retention over time.

APPENDICES

Appendix A - Pareto NBD Model



Appendix B - NBD Cohort Parameters



Appendix C - NBD Segment Parameters