How Does A Firm Adapt in A Changing World? The Case of Prosper Marketplace

Xinlong Li and Andrew T. Ching*

First draft: May 31st, 2019

This draft: March 18th, 2022

Abstract

We propose a *generalized revealed preference approach* to infer how a firm adapts in a changing environment. We apply this approach to Prosper, which is a peer-to-peer lending platform. To implement our approach, we develop a structural model, in which Prosper uses an adaptive learning/data selection algorithm to continuously update its training sample and re-estimate its predictive models about borrowers' and lenders' behavior, and use these updated models to conduct risk assessment for loan applications over time. To infer which adaptive learning algorithm Prosper may use, we consider a set of algorithms motivated by the machine learning literature. For each algorithm, we use Prosper's loan risk rating classification decisions to estimate the structural parameters of its objective function. By comparing the goodness-of-fit of these algorithm-specific models, we find evidence that Prosper most likely uses an ensemble algorithm that selects past data points based their economic conditions. We also find evidence that Prosper values both short-term expected revenue and long-term reputation when assigning ratings. We use our model to conduct counterfactual experiments to demonstrate how Prosper would change its rating choices if it (i) ignored short-term revenue, (ii) ignored long-term reputation, (iii) failed to adaptively learn, and (iv) knew the true data generating process.

Keywords: Adaptive Learning, Generalized Revealed Preference, Concept Drift, Machine Learning, Peer-to-peer Lending, Fintech

*Xinlong Li (xinlong.li@ntu.edu.sg) is Assistant Professor of Marketing at Nanyang Business School, Nanyang Technological University. Andrew T. Ching (andrew.ching@jhu.edu) is Professor of Business, Economics and Public Health at Carey Business School, Johns Hopkins University.

1 Introduction

What is the hottest item one day later may become just a fad. Due to constant technological innovations and an ever-changing economic environment, individual's preferences and behavior change rapidly these day. Therefore, treating all historical data to be equally informative when building an analytical model to guide business decisions could be quite misleading. More formally, the problem can be framed as follows. Suppose the independent variables are denoted by X_t (e.g., consumer characteristics) and the dependent variable is denoted by Y_t (e.g., consumer choice), where t indexes time. The relationship between X_t and Y_t can be formulated as $Y_t = F(X_t, \epsilon_t; \theta_t)$, where ϵ_t is the stochastic element, and θ_t is the parameter vector at time t. If we fail to take into account that θ_t may change over time and assume it is a constant, even if the model F is correct, the estimated $\widehat{\theta}$ can be very far from the true θ_t, resulting in poor predictive power of the estimated model. Researchers in statistics and machine learning call this the *concept drift* problem.

The concept drift problem has received more attention lately because the recent technology allows many companies to collect customer data that arrive in a stream; hereafter, we refer to this type of data as *streaming data*. Some common examples of streaming data include online reviews, website clicks, mobile phone apps, credit card transactions, E-commerce purchases, and social networks. Businesses are harnessing streaming data to develop deeper evidence-based insights about their customers to change the way they run their businesses. However, streaming data is likely subject to the pitfall of concept drifts. Traditional estimation approaches usually pool most data together and treat them equally when conducting estimation. But the predictive performance of models estimated using older data could quickly deteriorate over time. Therefore, inferences drawn from streaming data analysis will be questionable if the concept drift problem is not accounted for. Acknowledging the concept drift problem, the machine learning literature has taken advantage of the very large number of observations from streaming data and introduced ways to handle it (Tsymbal, 2004).

Companies have been recruiting a large number of researchers in Data Science, Statistics and Computer Science to help them analyze data and predict customer behavior in this fast-changing world. It is conceivable that tech companies are using some methods from the machine learning literature to address the concept drift problem. From the market intelligence viewpoint, if a firm can learn how its competitors use their data to make business decisions (e.g., setting prices, evaluating risks, etc.), it can gain valuable competitive advantages. In this paper, we propose a *generalized revealed preference* approach to infer how a firm uses its data (i.e., its adaptive learning/data selection algorithm) to make decisions. In particular, we use a peer-to-peer (hereafter P2P) lending platform to illustrate our approach.

The platform we study is Prosper Marketplace Inc. (hereafter Prosper), which is one of the largest online P2P lending platforms in the U.S.. One of the key services that Prosper provides is to evaluate each loan application's risk level and assign it to one of the seven ratings: AA, A, B, C, D, E and HR, where AA represents the lowest risk and HR represents the highest risk. Our key insight is that Prosper's decision on classifying risk categories will not only reveal the structural parameter values of its objective function, but also the way it selects the past data to adaptively learn and update its predictive models about borrowers' and lenders' behavior over time. This is the main contribution of our paper.

Our research also makes a substantive contribution of inferring how Prosper weigh different factors in its objective function. On the one hand, Prosper's rating should reflect a loan's true risk accurately in order to maintain Prosper's long-term reputation. On the other hand, Prosper only earns revenue when a loan is successfully funded, and this creates an incentive for Prosper to choose a rating to increase a loan's likelihood of getting funded. Our structural estimation results will reveal the relative importance of these two factors.

To capture this insight, we develop a structural model to capture Prosper's decision process. We assume that in each period, as more data come in, Prosper selectively uses the past data to update and re-estimate its borrower and lender models, and then uses them in its objective function to assign a rating to each loan application. Because the estimates of borrower and lender model parameters are a function of the way that Prosper selects and weights the past data, Prosper's loan rating classification is also a function of its adaptive learning algorithm. Hence, by observing Prosper's loan rating choices, it should be possible to infer how Prosper uses the past data (i.e., its adaptive learning/data selection algorithm). We call this the *generalized revealed preference approach*.

To implement our approach, we face one major challenge. There are infinitely many non-nested algorithms that a firm could use to weight the past data. It is impossible for us to consider all of them. Hence, we restrict our attention to a set of machine learning methods in handling the concept drift problem. For each adaptive learning algorithm, we re-estimate our structural model. Then we compare their goodness-of-fit. The adaptive learning algorithm which leads to the best model fit will be regarded as the most plausible algorithm adopted by Prosper within our consideration set.

Our data set consists of 31,807 unsecured personal loan applications from Dec 2010 to Dec 2012 provided by Prosper. Using our generalized revealed preference approach, we find that, among the set of data selection algorithms that we consider, an ensemble algorithm, which selects past data points based their economic conditions, best describe Prosper's data selection process. Moreover, we find that Prosper weigh both revealing the true risk and expected revenue into account. We conduct counterfactual experiments to investigate how Prosper will change its rating choices if (i) it does not care about short-term revenue; (ii) it does not care about long-term reputation; (iii) it does not adapt at all; (iv) it knows the true data generating process. We find that Prosper's short-term revenue would decrease by around 4.25% in case (i), decrease by 0.40% in case (ii), decrease by 3.06% in case (iii) and increase by 8.04% in case (iv).

The rest of the paper is organized as follows. Section 2 gives an overview of related literature. In section 3, we discuss the P2P lending industry and the data in detail, and provide some reduced form evidence that concept drifts exist in our context. The model framework is summarized in section 4. We introduce different algorithms of using historical data in section 5. We discuss the identification issues in section 6. Section 7 presents the estimation results. Section 8 is the conclusion.

2 Related Literature

Our paper is related to the economics literature on firms' adaptive learning behavior. This literature assumes that agents proceed like an econometrician and use the available data to estimate a model of the economy and update their beliefs about the model parameters as new data arrive. The pioneer works include Sargent et al. (1993), Evans et al. (2013), and more recent works include Doraszelski et al. (2018), Kozlowski et al. (2020). Aguirregabiria and Jeon (2020) provide a comprehensive review about how firms learn about demand, costs, or the strategic behavior of other firms in the changing market.[1] Similar to these studies, we also model how a firm revises its belief in an adaptive manner. It should be highlighted that previous works assume the data generating process is stable for an extended period of time, and the timing of structural breaks (if any) are known to agents in the model. Within each stable period, this literature typically assumes an adaptive learner will simply pool all data available up to the current time stamp to re-estimate and update the parameter values of the model that she faces.[2] In contrast to the previous works, we assume that firms recognize that concept drifts may be happening in the data generating process, but they do not know how often and when concept drifts happened. We argue that in order to adapt a fast-changing environment, firms may try to use the data available selectively when updating their models. In this research, we focus on uncovering the adaptive learning algorithm that a firm uses.[3]

Our paper is also related to the literature on P2P lending and more generally, crowdfunding. Wei and Lin (2017) and Liu et al. (2020) studied how the change in pricing mechanism (from auction to posted price) at Prosper affected its participants' behavior. Fu et al. (2021) studied whether machine learning algorithms can

[1]There are also some works which model firms learn about the demand conditions in a Bayesian manner. Hitsch (2006) models firms learn about the parameter values of the reduced form demand curves of their products, Ching (2010) models both firms and consumers are uncertain about product quality, and social learning allows both sides to learn its true quality over time.

[2]For instance, Kozlowski et al. (2020) assumes consumers use data available to update their estimate of the distribution of macroeconomic shocks to the economy.

[3]The marketing literature has also found evidence that consumer preferences evolves over time, e.g., Sriram et al. (2006), Liechty et al. (2005), Dew et al. (2020), etc. Preference evolution can be one source of concept drifts. More generally, concept drifts can also be due to change in the macro environment, competitor's policies, etc. Another source of concept drift could be cross-sided network externality, which we do not explicitly model. Some of our adaptive learning algorithms should be able to address it because the parameter estimates at each time stamp will be specific to the data being selected for estimation.

beat the crowds of investors in predicting loan defaults and lead to greater welfare for both investors and borrowers. Lin and Viswanathan (2016), Lin et al. (2013) and Iyer et al. (2016) investigated if soft information, such as home bias, online friendships of borrowers, pictures and text descriptions acts as signals of credit quality in P2P lending market.[4] Zhang and Liu (2012) studied lenders' rational herding decisions. Freedman and Jin (2011) investigated the role of learning-by-doing in alleviating the information asymmetry about borrowers. Kawai et al. (2021) examined how signalling can mitigate welfare losses that result from adverse selection. All of these papers, except Wei and Lin (2017) and Liu et al. (2020), study the early period when Prosper uses the auction mechanism to fund loans (similar to eBay auction).

None of the papers above considers the possibility that firms may face a concept drift problem. However, in the machine learning and statistics literature, many studies have shown the prevalence of concept drift and how it causes problems in predictive modeling (Hoens et al., 2012). Schlimmer and Granger (1986) first point out concept drift may affect model's prediction performance. Kelly et al. (1999) show that the concept drift problem exists in credit card default detections. Crespo and Weber (2005) show that adaptive data mining methods that update the model continuously outperform static models in customer segmentation analysis. In the famous Netflix Prize Competition, one of the lessons learned by the winning team is that taking temporal dynamics into account substantially contributes to building accurate models (Koren, 2009a). Their method allows users' average rating to change over time to capture their drifting preference. Our paper is the first to show that by using generalized revealed preference, we could infer whether a firm is actively dealing with the concept drift problem. Moreover, by running counterfactual analysis, we show that taking care of the concept drift problem and adaptively learning about the changing market can help the firm to increase revenue.

3 Institutional Background and Data

3.1 Institutional Background

Prosper is the first P2P lending platform in the U.S. It facilitated over $18 billion of unsecured loans in 2020, and it is still growing. Over time, P2P lending has gained more acceptance, and become an alternative way to borrow in the U.S.. The value of global P2P lending is expected to rise to one trillion U.S. dollars by 2050.[5]

It should be highlighted that in the early years of Prosper, it used an eBay style auction system that allowed lenders and borrowers to determine the interest rates. This business model earned Prosper the name "eBay of

[4] Another relevant paper is Ni and Xin (2020). They study the existence of local bias in one of the largest online crowdfunding marketplaces in China and quantify the importance of information asymmetry and preference toward local projects on inducing local biases using a structural model.

[5] https://www.statista.com/statistics/325902/global-p2p-lending/

Loans." However, Prosper was not very successful during this early phase. The credit score of the borrowers were relatively low, and the average investor returns were negative.[6] On Dec 19, 2010, the platform switched to a post-price model with pre-set interest rate for each loan application. The post-price model marked a new era for Prosper. Since its launch, Prosper has been growing rapidly. This provides us a good opportunity to study how a firm adaptively learns about the market when it switches to a new business model. In this study, we focus on loan applications initiated during the two-year window, Dec 19, 2010 to Dec 31, 2012, right after the regime change.

To register as a borrower on Prosper, an applicant needs to provide some basic identity information including name, Social Security number, address, telephone number, etc. Borrowers can request anywhere from $2,000 to $25,000 [7] per loan on Prosper and choose to repay over a 36- or 60-month amortization periods. In our analysis, we only consider loans with a fixed loan length of 36 months because this is what most borrowers choose, and the interest rate for 60-month loans is determined quite differently.

Before a borrower's loan application is approved, Prosper pulls the borrower's credit history from its credit reporting partner Experian. Some key credit history variables (e.g., the borrower's credit score, number of delinquencies in the last seven years, total number of inquiries, bankcard utilization rate, etc.) help Prosper to assess loan risk by mitigating asymmetric information, and determine eligibility (Chan et al., 2020). Loan applications will not be approved unless the borrower's credit score is above 600. If a loan application is approved, Prosper then uses its risk assessment algorithm to assess its risk and assign it a rating. Prosper states that their risk assessment algorithm is trained based on the historical performance of Prosper loans with similar characteristics, without providing any details. Prosper assigns each loan application a rating using a system with seven grades: AA, A, B, C, D, E and HR, where AA is the lowest risk and HR stands for the highest risk. Each rating corresponds to a certain interest rate. Once the loan application is approved, it is open for investment for 14 days. Moreover, for each loan application, Prosper posts Loss Rate (hereafter, we refer it to as "posted loss rate"), which is a proxy for expected loss rate. It is important to highlight that posted loss rate only depends on its risk rating. So, similar to loan interest rates, posted loss rates are not loan application specific, instead, they are rating specific.

Prosper charges lenders a 1% annual loan service fee based on the current outstanding loan principal lenders hold. For each funded loan, Prosper also charges borrowers an origination fee, which is a percentage of the loan amount and varies across Prosper's rating. Note that Prosper only earns revenue when a loan application is successfully funded. Table 1 shows the average interest rate, average annual return, origination fee rate and posted loss rate for each rating level. It is worth noting that, on average, lower rated loans have higher annual returns, and E loans have the highest average annual return. This is consistent with the general finance principle that high risk investment (i.e., high variance) carries higher expected return.

[6] https:www.lendacademy.comprosper-review, accessed on June 4th, 2021.

[7] The amount limit only applies to our sample period. Nowadays borrowers can request as much as $40,000 per loan.

Borrowers make repayments of equal amount on a monthly basis. If a borrower misses four repayments in a row, the loan is marked as "Defaulted."[8] Lenders lose their outstanding principal in defaulted loans. Defaulted loans hurt Prosper in three ways: (i) Prosper loses the annual service fee on the unpaid principal of defaulted loans; (ii) Prosper needs to incur the operating cost of loan collections; (iii) the default rate is one of the main factors that affects lenders' investment decisions, and it is widely discussed online (Cunningham, 2014). Hence, the credit screening and risk assessment processes are very important to Prosper.

Lenders are allowed to contribute as little as $25 to a loan and as much as the full amount requested by the borrower. Lenders have access to all the financial history information of borrowers when they make their investment decisions. Prosper provides full historical loan performance data on its website.

3.2 Data Description

Our entire data set spans from Feb 2007 to Dec 2012. However, because we focus on studying Prosper's decision process after it has implemented the posted price business model, we only describe the data on loans initiated after Dec 19, 2010, when Prosper switched to the posted price business model. We use data prior to Dec 19, 2010 to construct initial conditions for our models.

Our data set consists of 31,807 loan applications. Among them, 22,277 (70.04%) loan applications were funded, 5,825 (18.31%) were withdrawn by the borrowers, 3,705 (11.65%) were expired (i.e., they were not funded), and 3,529 (15.84%) were defaulted. Table 2 provides summary statistics of loan applications' outcome based on their Prosper ratings. Figure 1 presents the percentage distribution among seven ratings over time. The rating distribution changes over time, and this can be driven by changes in borrowers' characteristics and/or Prosper adaptive learning over time In our empirical exercise, we will take both factors into account.

Figure 2 shows how the conversion rate (i.e., the rate at which loan application become funded) and default rate change over time, and they show large variation. These variations reflect lenders' and borrowers' changing investing and borrowing behaviors, which can help identify Prosper adaptive learning/data selection algorithm.

For each loan application, we observe the borrower's credit history variables (e.g., his/her FICO score range, credit lines, current delinquencies, etc.), Prosper rating, interest rate, monthly loan payment amount, and whether this loan application is funded, withdrawn, or expired. We also observe the borrower's full repayment history, including the percentage of principal loss in the case of default. Table 3 provides a full list of the variables we can observe in the data. Table 4 presents their summary statistics. The average amount requested by each borrower is $6,829 and the average interest rate a borrower needs to pay is 23.34%. The

[8] For instance, if borrower A did not make repayment for four months in a row starting month 2, then she will be coded as default in month 2.

average interest rate is relatively high, which reflects that it is harder for borrowers on this platform to borrower from traditional channels (e.g., banks). The high interest rate should also attract potential lenders to invest in this platform.

On average, Prosper charges 4.36% of the loan amount as the origination fee. The average lower bound of FICO score across all loan applications is 697.47, which is much higher than the minimum eligibility requirement of 600 on Prosper. The mean annual income of borrowers was $68,004, which was more than two times higher than the per capita annual income of $29,173 in the U.S in 2011.[9] Borrowers' average monthly debt is $873, which is 15.40% of their average monthly income. Borrowers average bankcard utilization rate is 51.50%, which is much higher than the average 30% credit utilization rate.[10] About half of the borrowers own a home and the average mortgage balance is $107,599. For comparison, the average home ownership rate in 2011 is 66.1% in the U.S. Hence, a typical borrower in Prosper is relatively young and middle class.

3.3 Reduced Form Evidence of Concept Drifts

In this section, we use the test proposed by Gama et al. (2004) to provide evidence of the concept drift problem. The intuition of this test relies on detecting a significant deterioration in model prediction using data points just beyond the training set. The test makes use of a growing sample of observations to keep revising the prediction error rates of the model. If there is no concept drift, we expect that the prediction error rate should remain fairly stable. But if a major concept drift happened, it would lead to a significant increase in the prediction error rate. The basic mechanism of the test is as follows: it starts with a base window of sample period, estimates the model and computes its prediction error in the period just beyond the base window. If the prediction error rate is below a threshold, we will expand the base window, and repeat this procedure until the prediction error rate is above the threshold. This is when we detect at least one concept drift has happened. As soon as we detect a positive test outcome, we will start a new data window by discarding the old data.

We will use the following example to illustrate how this test works. Consider a set of loan applications, in the form of pairs (X_i, y_i), where X_i represents borrower i's characteristics, and y_i takes value 1 if the loan application is funded and 0 otherwise. Suppose that we use a logistic regression model to predict each loan application's funding probability. Let $\hat{y}_l$ denotes the predicted outcome for loan application l, where subscript l denotes observations in the period just beyond the base window. The event of false prediction, $\hat{y}_l \neq y_l$, is a random variable from Bernoulli trials.[11] For a sequence of n_t observations, the number of false

[9] https://www.deptofnumbers.com/income/us/

[10] https://www.creditcards.com/credit-card-news/credit-card-use-availability-statistics-1276.php

[11] We set $\hat{y}_l = 1$ if our logistic model predicts loan application l's funding probability is larger than 0.5, and $\hat{y}_l = 0$ otherwise. Note that we model X_i's to enter the model linearly.

prediction events follows a Binomial distribution. Let p_t denote the prediction error rate (i.e., the percentage of false predictions) in period t. If lenders' investing behavior is stationary, p_t should remain stable over time. A significant change in the error rate suggests a change in lenders' investing behavior. We describe the details of the test in Appendix A, including how we define the detection threshold and update the data window. In Figure 3(a), we show the prediction error rate on whether loan applications are funded in each time stamp. The red dots represent when concept drifts are detected by the test. For lenders' investing behavior, the test is able to detect concept drift 17 times during our sample period. It is important to note that because the test relies on the "cumulative" effects of concept drift on the prediction error, when the test result is positive, one or more concept drifts may have already happened for some time. Hence, the number of times detected should be interpreted as the lower bound for the number of concept drifts happened in the sample period.

In addition to observing whether loan applications are funded or not, Prosper also observe whether borrowers withdrew their applications, and whether they defaulted their loans. We test the existence of concept drift for these two outcomes, using a logit model with the same set of explanatory variables as in the loan funding model. Figures 3(b) and 3(c) show the test results. We are able to detect concept drifts 7 and 5 times for the withdrawal and default decisions, respectively.

4 Model

We model Prosper as an adaptive decision maker. At the end of period t, we assume Prosper updates and re-estimates its borrower side and lender side models by incorporating the new information it receives in that period, and then uses these models with newly updated parameter values to assign rating to loan applications in period $t + 1$.

Our model has four parts: (i) the borrower side model; (ii) the lender side model; (iii) the risk assessment model; (iv) Prosper's utility function, which consists of the expected revenue from each loan application, and a penalty on assigning a rating deviated from the loan's true risk. We consider five different adaptive learning/data selection algorithms that Prosper may utilize to estimate parts (i)-(iii) in each period, and then incorporate them in the utility function. In each period, Prosper chooses loan rating to maximizes its utility. For each adaptive learning algorithm, we estimate the structural parameters of the utility function. We then compare the goodness-of-fit of these five models distinguished by adaptive learning algorithm, and infer which adaptive learning algorithm is most likely used by Prosper,

When developing the components for parts (i)-(iii), we are thinking from the viewpoint of Prosper. Prosper's goal is to use these models to predict how likely each new loan application will be funded as well as its risk. Prosper also needs to update these predictive models very frequently in each period. Hence, each predictive model must be simple enough so that Prosper can update it quickly.

The adaptive learning algorithms under consideration here are motivated by some common approaches proposed in the machine learning literature (e.g., gradual forgetting, moving window, etc.). We wait until section 5 to discuss the details of these adaptive learning algorithms.

4.1 Borrower Side

We assume that borrowers arrive exogenously, submit their loan applications and then find out their rating. They then decide whether to withdraw their loan applications or not.

Let X_i denote the observed characteristics of the borrower for loan application i. It consists of the borrower's financial history information, such as credit score, monthly income, number of delinquencies, number of prior Prosper loans, etc.[12] Let Z_{il} denote Prosper-determined variables associated with rating l, including interest rate, posted loss rate, origination fee rate and monthly repayment.

In each period t, borrower i's utility of staying with Prosper depends on the rating, l, assigned by Prosper:

$$\begin{aligned} U^b_{ilt}(X_i, Z_{il}; \boldsymbol{\gamma_t}) &= \gamma_{lt} + \gamma_{1t} \cdot O_l \cdot M_i + \gamma_{2t} \cdot MP_{il} + \gamma_{3t} \cdot M_i + \epsilon_{i1t}, \\ &= f_1(X_i, Z_{il}; \boldsymbol{\gamma_t}) + \epsilon_{i1t}, \end{aligned} \tag{1}$$

where $\boldsymbol{\gamma_t}$ represents the vector of borrower's side coefficients in period t; O_l is the origination fee rate for a rating l loan; M_i indicates the amount of loan i; MP_{il} is borrower i's monthly payment. Note that MP_{il} is a function of the interest rate (which has 1-1 mapping with rating l) and M_i. This is why borrower i's utility is a function of rating l.

For a 36-month loan, monthly payment is calculated as follows:

$$MP_{il} = \frac{R^*_{il} \cdot M_i}{1 - (1 + R^*_{il})^{-36}},$$

where $R^*_{il} = \frac{R_l}{12}$; R_l is the annual interest rate for loans assigned with rating l.

Note that if the interest rate set by Prosper is too high, a borrower may find it more attractive to apply for a loan somewhere else and withdraw her loan application. Moreover, the macro environment can also affect an individual's likelihood to participate in Prosper. Therefore, we model a borrower's outside option as a function of her characteristics and the macro environment,

$$U^b_{i0t} = f_0(X_i, E_t; \boldsymbol{\gamma_t}) + \epsilon_{i0t}, \tag{2}$$

[12] Each borrower can only have one submitted loan application at any point of time. So we use loan application i and borrower i interchangeably. A borrower can submit another loan application once her previous loan application is funded, expired or withdrawn. A repeated borrower will receive a new index value, and the prior Prosper loans variable will capture whether she is a repeated borrower.

where E_t represents macroeconomic variables including S&P 500 closing quotes, the TED spread, the U.S. 30-year mortgage rate. Assuming the idiosyncratic errors, ϵ_{i0t} and ϵ_{i1t}, follow type-I extreme value distribution, the withdrawal probability of loan application i with rating l is:

$$W_{ilt} = Pr(\text{Withdraw} = 1|X_i, Z_{il}, E_t; \boldsymbol{\gamma_t}) = \frac{1}{1 + \exp(f_1(X_i, Z_{il}; \boldsymbol{\gamma_t}) - f_0(X_i, E_t; \boldsymbol{\gamma_t}))}. \tag{3}$$

The full set of variables included in X_i, Z_{il} and E_t are listed in Table 3. We should highlight that the parameter vector, γ_t, depends on t. This is because we allow Prosper to use new information available in each period to update its model and adapt to an non-stationary environment with concept drifts.

In each period, Prosper observes which loan applications are funded, withdrawn or defaulted, and uses the new information to update its models accordingly. Suppose that some loan applications were withdrawn in period t. At the end of period t, Prosper labels these loan applications as withdrawn and updates the model that predicts a loan application's withdrawal probability by taking these newly withdrawn loan applications into account. For all the loan applications arrive in period $t+1$, Prosper uses the updated model to predict their withdrawal probabilities. Prosper updates its lender side and risk assessment models in the same way.

4.2 Lender Side

Here, we use a simple logit model to capture funding probability. In each period t, the probability that loan application i getting funded (hereafter we refer it to "funding probability") is modeled as follows:

$$F_{ilt} = Pr(\text{Funded} = 1|X_i, Z_{il}, E_t; \boldsymbol{\beta_t}) = \frac{\exp(\beta_{lt} + X_i\beta_{1t} + Z_{il}\beta_{2t} + E_t\beta_{3t})}{1 + \exp(\beta_{lt} + X_i\beta_{1t} + Z_{il}\beta_{2t} + E_t\beta_{3t})}, \tag{4}$$

where the definitions of X_i, Z_{il} and E_t are the same as those in the borrower side model; $\boldsymbol{\beta_t}$ represents the parameter vector on lender's side. Again, the parameter values is a function of t because we assume Prosper updates its lender side model in each period using its adaptive learning algorithm.

The simplicity of our funding probability model is in line with our intention of using a predictive model which can be updated quickly. Note that Prosper does not release individual lender's investment decision data after switching to the posted price business model. Even if one wants to model individual lender's investment decisions, there is no publicly available data to do so.

4.3 Risk Assessment Model: Naive Bayes Classifier

In addition to updating the borrower side and lender side models, Prosper also needs to re-evaluate its risk assessment model in each period. In this section, we explain the model for evaluating each loan's default probability and loss rate given default (hereafter LGD).

We will use Naive Bayes classifier in our study to assess the risk of each loan application. In practice, a Naive Bayes classifier performs very well compared with other machine learning algorithms in predicting financial risk. Viaene et al. (2002) compared different classification methods in the context of expert automobile insurance claim fraud detection and find that the Naive Bayes method outperformed other methods. Wang et al. (2003) showed that Naive Bayes outperformed decision trees in dealing with the concept drift problem. Humpherys et al. (2011), Domingos and Pazzani (1997), Friedman et al. (1997) also provided evidence to support the superior performance of Naive Bayes classifier in other contexts. In addition, Naive Bayes model can be estimated very quickly. This is important here because Prosper needs to evaluate the risk for a large number of loan applications in each period, after incorporating new data points.

Let us explain how to use the Naive Bayes model in our context. Suppose that we have a labeled data set $\{(X_i, y_i), i = 1, 2, ..., n\}$, where X_i is a vector of characteristics for loan application i; and $y_i \in \{1, 2, ..., K\}$ represents loan application i's label (e.g., the label can indicate whether loan i defaults or not). Assume each loan application has m characteristics, which can be represented by $X_i = (X_{i1}, X_{i2}, ..., X_{im})$. Let $p(y_i = k)$ be the prior belief about the label. According to the Bayes's rule, the posterior belief is:

$$p(y_i = k|X_i) = \frac{p(y_i = k) \cdot p(X_i|y_i = k)}{p(X_i)} = \frac{\omega_k \cdot p(X_i|y_i = k)}{p(X_i)}, \tag{5}$$

where $\omega_k = p(y_i = k)$. To approximate the initial prior belief on a loan's default probability, we use the empirical default rate for all loans prior to the period under consideration. To apply equation (5), one challenge is to obtain $p(X_i|y_i = k)$. Estimating this joint conditional distribution is very data-demanding when X_i consists of many variables.[13] The Naive Bayes estimator uses several simplifying assumptions to by-pass this hurdle.

First, given the category y_i, it assumes that loan application characteristic X_{ij} are conditionally independent of each other. That is,

$$p(X_i|y_i = k) = \prod_{j=1}^{m} p(X_{ij}|y_i = k).$$

Therefore, we have,

$$\begin{aligned} p(y_i = k|X_i) &\propto \omega_k \cdot p(X_i|y_i = k) \\ &\propto \omega_k \cdot \prod_{j=1}^{m} p(X_{ij}|y_i = k). \end{aligned}$$

Second, it assumes the conditional distributions $p(X_{ij}|y_i = k)$ follows a multinomial distribution with the number of trials equal to 1. For example, suppose that the jth loan application characteristic is income and it takes four levels: 1, 2, 3 and 4. If borrower i has income level 2, then $X_{ij} = (X_{ij1}, X_{ij2}, X_{ij3}, X_{ij4}) = (0, 1, 0, 0)$.

Let the vector $P_{kj} = (p_{kj,1}, ..., p_{kj,n_j})$ denote the probability distribution of characteristic j given $y_i = k$, where n_j represents the number of values or levels that characteristic j can take. We have,

[13] If X_i consists of 10 variables, and each variable takes 5 values, there will be 9.7 million cells. The true distribution could be very sparse for many cells, and that will require a very large sample size to obtain a precise estimate.

$$p(X_{ij}|y_i = k) \propto \prod_{h=1}^{n_j} p_{kj,h}^{X_{ijh}}.$$

Moreover, we have

$$P(X_i|y_i = k) = \prod_{j=1}^{m} p(X_{ij}|y_i = k) = \prod_{j=1}^{m} \prod_{h=1}^{n_j} p_{kj,h}^{X_{ijh}}. \tag{6}$$

To obtain an estimate for P_{kj}, we use the crude frequency estimator. That is, for each characteristic j, we use the empirical frequency of its values conditional on the data with label k, up until the current period. Note that in contrast to estimating the conditional joint distribution of $(X_{i1}, ..., X_{iJ})$, it is much less data-demanding to estimate the conditional marginal distribution of each characteristic precisely; even a sample size with just a few hundred observations can give us a fairly precise estimate because it only takes a handful number of values.

In the actual implementation, we keep updating ω_k and P_k over time. We use $\boldsymbol{\Omega_t} = \{\omega_{1t}, \omega_{2t}, ..., \omega_{Kt}\}$ to represent prior belief at time t and $\boldsymbol{P}_{kt} = (P_{k1t}, P_{k2t}, ..., P_{kmt})$ to represent the joint distribution of all m characteristics in class k at time period t. The details on how we update $\boldsymbol{\Omega_t}$ and $\boldsymbol{P}_{kt}$ can be found in Appendix B.

We apply the naive Bayes classifier to estimate each loan application's expected default probability (ED) and expected loss rate given default ($ELGD$). In particular, in period t, we use ED_{ilt} and $ELGD_{ilt}$ to represent loan application i's expected default probability and expected loss rate given default, respectively. Note that both ED_{ilt} and $ELGD_{ilt}$ are functions of rating. The reason is different ratings correspond to different interest rates and therefore a borrower needs to pay different monthly payment to the lenders, which in turn will affect the borrower's default probability and LGD. When estimating a loan application's LGD, we discretize the LGD into four groups using the 25%, 50% and 75% quantiles, and apply the standard naive Bayes classifier to the discretized data to get each loan application's probabilities of belonging to each group. At the same time, we compute the average LGD, which is denoted as $\overline{ELGD}_{kt}$, in each loss group k at time period t, $k = 1, 2, 3, 4$. Then, the predicted LGD for loan application i with rating l, $ELGD_{ilt}$, is defined as $ELGD_{ilt} = \sum_{k=1}^{4} p_{ilk}^{LGD} \cdot \overline{ELGD_{kt}},\ k = 1, 2, 3, 4$, where p_{ilk}^{LGD} is the probability that loan application i belongs to LGD group k if it is assigned rating l. For default prediction, We use $\boldsymbol{\overline{\Omega}_{dt}}$ and $\boldsymbol{\overline{P}_{dt}}$ to denote all the priors and posteriors beliefs, respectively. For LGD, we use $\boldsymbol{\overline{\Omega}_{lt}}$ and $\boldsymbol{\overline{P}_{lt}}$ to denote all the priors and posteriors beliefs, respectively.

We should point out that at any given period, some loans will be "on-going," and we do not know whether they will eventually default or not, and if so, what their LGDs are. Therefore, we follow Prosper and define a loan as defaulted if a borrower misses 4 repayments in a row.

4.4 Firm's Objective Function

Prosper considers each loan application i separately, and assigns it a risk rating l to maximize the expected utility from this loan application. The decision is based on the state vector $S_t = (\boldsymbol{\gamma_t}, \boldsymbol{\beta_t}, \overline{\boldsymbol{\Omega}}_{\boldsymbol{t}}, \overline{\boldsymbol{P}}_{\boldsymbol{t}})$, where $\boldsymbol{\gamma_t}$ and $\boldsymbol{\beta_t}$ are borrower and lender side parameters explained in sections 4.1 and 4.2, respectively; $\overline{\boldsymbol{\Omega}}_{\boldsymbol{t}} = (\overline{\boldsymbol{\Omega}}_{\boldsymbol{dt}}, \overline{\boldsymbol{\Omega}}_{\boldsymbol{lt}})$, and $\overline{\boldsymbol{P}}_{\boldsymbol{t}} = (\overline{\boldsymbol{P}}_{\boldsymbol{dt}}, \overline{\boldsymbol{P}}_{\boldsymbol{lt}})$ denote the the prior and posterior belief parameters for the default and loss rate given default prediction models explained in section 4.3. Recall that interest rates and posted loss rates each have a 1-1 mapping with Prosper's loan rating. We focus on modeling Prosper's rating assignment decisions and take its mapping to interest rates and posted loss rates as given. Prosper's utility of assigning rating l to loan i posted in t is:

$$U^P_{ilt}(X_i, Z_{il}|S_t, AL, \boldsymbol{\alpha}, \delta_1, \delta_2) = \alpha_l + \delta_1 \cdot (O_l + L_{ilt}) \cdot M_i \cdot F_{ilt} \cdot (1 - W_{ilt}) + \delta_2 \cdot |PLR_l - ED_{ilt} \cdot ELGD_{ilt}| + \epsilon_{ilt}, \tag{7}$$

where O_l is the origination fee rate for loan applications with rating l; L_{ilt} is the expected annual loan service fee rate that Prosper charges the lenders;[14] M_i is the loan amount requested; F_{ilt} and W_{ilt} are the predicted funding and withdrawal probability, respectively; PLR_l is the posted loss rate; ED_{ilt} and $ELGD_{ilt}$ are the expected default probability and expected loss rate given default, respectively, and $(ED_{ilt} \cdot ELGD_{ilt})$ is the expected loss rate for loan application i given rating l; δ_1 is the utility weight on expected revenue; δ_2 captures the cost of misreporting the loan application's risk (e.g., it may damage Prosper's long-term reputation, and affects its future revenue); α_l's are the alternative specific intercepts; AL denotes a specific adaptive learning algorithm used in estimation step 1 to obtain $F_{ilt}, W_{ilt}, ED_{ilt}$ and $ELGD_{ilt}$; ϵ_{ilt} is the random utility term which follows the type-I extreme value distribution.

Note that O_l and PLR_l do not depend on i because they are rating specific. In addition, it is almost surely that $|PLR_l - ED_{ilt} \cdot ELGD_{ilt}| \neq 0$ because PLR_l is rating specific, and can only take one of seven values, but $ED_{ilt} \cdot ELGD_{ilt}$ is continuous. Our intuition suggests that $\delta_1 > 0$ and $\delta_2 < 0$ because expected revenue should bring Prosper positive utility and misreporting a loan's risk should bring Prosper negative utility.

[14] Notice that the annual service fee is charged monthly based on the current outstanding loan principal a lender has. So if a borrower defaults on her loan, lenders who invested in this particular loan will lose part of their principal, and Prosper cannot collect a service fee from lost principal. Therefore, the expected service fee rate for loan application i is a function of the annual service fee rate and when the loan will default. The later the loan defaults, the more service fee Prosper can charge from the lenders. In our data sample, the service fee rate is fixed at 1%. We approximate the number of repayments borrower i has made as $PT_{ilt} = \textbf{round}((\text{1-ELGD}_{ilt})/\frac{1}{36}) = \textbf{round}((\text{1-ELGD}_{ilt}) \cdot 36)$. The monthly service fee rate is $r_s = 0.01/12$. Since the service fee is charged monthly on the current outstanding loan principal, L_{ilt} is calculated as

$$L_{ilt} = \begin{cases} 0, & PT_{ilt} = 0; \\ ED_{ilt} \cdot r_s \cdot \sum_{j=1}^{PT_{ilt}} (37 - j)/36 & PT_{ilt} = 1, 2, ..., 36. \end{cases}$$

Hence, L_{ilt} is a function of ED_{ilt} and $ELGD_{ilt}$.

As we report our results in the next section, the estimated signs of these two structural parameters confirm our intuition. Like standard discrete choice models, $\frac{\delta_1}{\delta_2}$ will determine how Prosper makes the trade-off between earning revenue and reporting the true risk of a loan application. If the estimated $\frac{\delta_1}{\delta_2}$ is very close to zero, then the data reveals that Prosper is practically only interested in using the rating to reflect a loan application's true risk, and vice versa.

We assume Prosper takes a two-step approach to determine loan ratings. First, it uses an adaptive learning algorithm to estimate the predictive models in sections 4.1 and 4.2 to obtain $F_{ilt}, W_{ilt}, ED_{ilt}$ and $ELGD_{ilt}$. Second, it takes the predicted $F_{ilt}, W_{ilt}, ED_{ilt}, ELGD_{ilt}$ as given, and chooses a loan rating to maximize its objective function for each loan application. Let $\theta = (\alpha_1, ..., \alpha_7, \delta_1, \delta_2)$, we have

$$l^* = arg \max_l \left[U^P_{ilt}(X_i, Z_{il}|S_t, AL, \theta)\right]. \tag{8}$$

We can then write the likelihood function in the following closed form:

$$L(X, Z|\mathbf{S}, AL, \theta) = \prod_{t=1}^{T}\prod_{i=1}^{I_t}\prod_{l=1}^{7}\left(\frac{\exp(U^P_{ilt}(X_i, Z_{il}|S_t, AL, \theta))}{\sum_{j=1}^{7}\exp(U^P_{ij}(X_i, Z_{ij}|S_t, AL, \theta))}\right)^{\mathbb{1}\{\text{Rating}_i^o = l\}}, \tag{9}$$

where $\mathbf{S}$ denote the state variables in all periods; T represents the total number of time periods; I_t represents the number of loan applications in period t; Rating_i^o is loan application i's observed Prosper rating. We estimate the structural parameters of Prosper's objective function $\theta = (\alpha_1, ..., \alpha_7, \delta_1, \delta_2)$ using maximum likelihood.[15] It is important to highlight that the likelihood function is conditional on an adaptive learning algorithm, AL, and so does Prosper's policy function. Hence, when we repeat the two-step estimation exercise for each adaptive learning algorithm that we consider, we can compare the goodness-of-fit of this set of policy functions and find out which explains the data the best. This is essentially our inference strategy.[16] Figure 4 illustrates our model framework. We now turn to discuss the set of adaptive learning algorithms that we consider.

5 Adaptive Learning Algorithms

According to Tsymbal (2004), there are three general approaches of handling the concept drift problem: (1) observation selection; (2) observation weighting; and (3) ensemble learning. The goal of the observation selection approach is to select observations relevant to the current situation to estimate the model. Observation weighting is about assigning a weight to each observation depending on its relevance to the current situation. Ensemble learning method maintains multiple individual models and use them to

[15]Note that one of the utility intercept needs to normalized for identification reason. We normalize $\alpha_7 = 0$, where alternative 7 corresponds to the HR rating. Similarly, alternative 1 corresponds to AA rating.

[16]Similar two-step estimation approach has been used in the structural estimation literature (e.g., Bajari et al., 2007; Ching, 2010; Erdem et al., 2003; Yao et al., 2012, etc.)

make a combined prediction. Each individual model relies on its own observation selection and weighting approach. We consider five different adaptive learning algorithms, which represent these three approaches well. In particular, algorithms (1)-(4) rely on observation selection and weighting approaches, and algorithm (5) is an ensemble learning approach.

Of course, there are many other adaptive learning algorithms that Prosper can use. Unfortunately, it is impossible to examine every algorithm. Our approach follows McCullagh and Nelder (2019), who argue that "all models are wrong; some, though, are better than others and we can search for the better ones." As a proof-of-concept study, our goal is to demonstrate that our framework allows us to compare different adaptive learning algorithms and see which one is closest to what Prosper actually uses. If researchers are interested in other adaptive learning algorithms, they can apply our framework to evaluate them as well.

We emphasize that the model framework remains the same for all algorithms discussed below. Each algorithm leads to its own set of parameter values in the borrower, lender and risk assessment models (the withdrawal, funding, default probabilities and LGD), which go into the firm's objective function as we explained in section 4.

5.1 Equal Weight Method

The first algorithm that we consider is the Equal Weight Method (EWM). This is by far the most commonly used method when companies or researchers face limited sample size. It simply makes use of all the data available and weight them equally. Its underlying assumption is that there is no concept drift and the data generating process remains unchanged over time. We treat this algorithm as the benchmark case. For EWM, the length of a period is a day. That is, Prosper updates its model every day.

5.2 Moving Window Method (MWM)

The second method we consider is the Moving Window Method (MWM), which belongs to the observation selection approach. MWM is one of the most widely used observation selection approaches to deal with the concept drift problem (Pechenizkiy et al., 2010; Forman, 2006). At each time step, the algorithm re-estimates the parameters of the model using the data from the new training window. We consider the most recent N days as a fixed window size, and Prosper updates its model every day. We use grid search to determine the optimal window size. The selection is done by using the receiver operating characteristic (ROC) curve as the measurement metric for funded, withdrawal and default probabilities predictions, and mean square error (MSE) for LGD prediction. We provide more details about ROC curve in Appendix C.

The optimal window size is 210 days (roughly 7 months) for the withdrawal probability model; and the optimal window sizes are 90, 1080 and 1440 days (roughly equivalent to 3, 36 and 48 months) for funding probability, default probability and LGD models, respectively. Notice that the optimal window sizes for default prediction and LGD prediction are much larger. This is because only 70.04% of the loan applications get funded, and among the funded loan applications, only 15.84% of them defaulted. Hence, the number of defaulted loans is much smaller than the number of total loan applications. The default prediction and LGD prediction models need larger window sizes to have enough observations to reach better prediction performance. Table 6 presents the prediction performances based on these window sizes.

5.3 Recession Probability Method (RPM)

The next adaptive learning algorithm we consider is Recession Probability Method (RPM), which is a moving window based method with varying window sizes. This method selects observations into the window based on the similarities of economic conditions. The intuition is that borrowers' and lenders' behavior should depend on the economic environment. For example, during recession periods, investors tend to invest in less economically sensitive sectors and more borrowers tend to default. Therefore, we can use historical data generated from an economic environment which is similar to the current period to help us make inferences about the current market.

To define similar economic environments, we make use of the Smoothed U.S. Recession Probabilities released by the Federal Reserve Bank. U.S. Recession Probabilities are the smoothed probabilities of a recession in the U.S., which are calculated from a dynamic-factor Markov-switching (DFMS) model on non-farm payroll employment, industrial production index, real personal income, real manufacturing, and trade sales. The DFMS model was originally developed by Chauvet (1995) and has been used to study stock market volatility business cycle turning points in many studies including Sornette (2004), Kim and Nelson (1998), Hamilton and Lin (1996), etc. Chauvet and Piger (2008) analyzes the performance of the model for identifying economic turning points defined by NBER and show the model can give satisfactory prediction results. Compared with single economic index like GDP growth rate or unemployment rate, smoothed recession probability is a more comprehensive measure. We provide more details about the Smoothed U.S. Recession Probabilities in Appendix D.

We need to determine the optimal data window based on economic conditions for each of borrower side, lender side, default prediction and LGD prediction models. Take the borrower side model for example, the optimal similar economic environment should enable the model to make best predictions for borrower's withdrawal decisions. Let RP_t denote month t's recession probability. We find that a model estimated using data generated from months with recession probability within [RP_t-0.4%, RP_t+0.4%] performs the best. The corresponding optimal bandwidth recession probabilities for lender side model, default prediction model and LGD prediction model are RP_t +/- 0.3%, 3.1% and 1.1%, respectively.

We use the data from Feb 01, 2007 to Dec 19, 2010 as our initial data set. That is, we select similar economic environment from Feb 01, 2007 to Dec 19, 2010 when we begin our model estimation. Therefore, we use

71 months data in total for our estimation. For RPM, since the recession probability is a monthly measure, we update our model every month. Even though RPM is updated each month, we still use the model to make daily decisions. For instance, we use the model updated on June 30th, 2012 to make rating assignment decisions for every day in July, 2012.

Unlike MWM, RPM does not discard older information in a mechanical way. Instead, it utilizes older information in a more complicated way. The rationale behind RPM is that consumer behavior is highly correlated with the economic environment. Therefore, taking advantage of historical consumer behavior data that is generated from similar economic environments as today should be beneficial for the current period's prediction problems.

5.4 Gradual Forgetting Method (GFM)

As Koren (2009b) pointed out, observation selection may be reasonable when the drift is abrupt, but less so when the drift is gradual. One alternative way to deal with graduate concept drift is observation weighting. Gradual Forgetting Method (GFM) is the most widely used observation weighting algorithm. To implement GFM, we follow Klinkenberg (2004) and weight observations according to their age using an exponential aging function. To be specific, let $\lambda \in (0, 1)$ denote the forgetting factor. Then all the observations in period t will be assigned weight λ^t. The older an observation is, the less weight it carries. To estimate the logistic regression models for the borrower side and lender side, we employ the estimation method proposed by Balakrishnan et al. (2008). This method is based on a quadratic Taylor approximation to the log-likelihood. A forgetting factor can be easily incorporated into this estimation scheme and the model can be recursively estimated, which significantly reduces the computational burden. We provide the estimation details of this method in Appendix E.1. With respect to the prediction about default and LGD, we incorporate a forgetting factor to introduce temporal adaptivity in the naive Bayes model as well, and we explain the details in Appendix E.2. For this algorithm, we still assume the length of a period is a day and that Prosper updates its model every day.

For each of the four models (funding probability, withdrawal probability, default probability, and LGD), we employ a grid search over the forgetting factor. As in section 5.2, we compare their out of sample prediction performance with different forgetting factors based on ROC and MSE. We find that the best forgetting factor for the funding probability, withdrawal probability, default probability, and LGD models are 0.997, 0.990, 0,995 and 0.997, respectively. The full set prediction performances are presented in Table 6. Note that these are daily forgetting factors. Take the default prediction model as an example. An observation which is one year "old" only carries weight of $0.995^{365} = 0.16$.

5.5 Ensemble Recession Probability Method (E-RPM)

The last algorithm that we consider is Ensemble Recession Probability Method (E-RPM), which extends RPM and uses an ensemble learning approach.[17] In the first four algorithms, they all rely on one model to deal with concept drifts. In the machine learning literature, researchers find that it can advantageous if we combine several predictive models together. This ensemble modeling approach captures the idea of "the wisdom of the crowd." Even if each individual model might be weak, the aggregated model can perform very well in prediction.

In E-RPM, we applies this ensemble approach to RPM. To train a model which specializes in a specific economic environment, we first use the recession probability index to divide the whole data set into different groups. For instance, data generated from recession period is assigned to the recession group, while data generated from economic booms is assigned to the booming group. We then estimate a model using each data group separately. We take a weighted average of each individual model to get our ensemble model. Each model's weight is determined dynamically by its predictive performance in $t-1$. The better its predictive performance is in $t-1$, the higher weight it carries in t. In appendix F, we provide the details about how we set the weights and update the model over time. Note that each sub-model has the same model framework, as shown in Figure 4. The only difference is that they are estimated using data from different economic environments.

6 Identification

In this section, we provide some intuitions about our model identification.

Recall that for each adaptive learning/data selection algorithm, we take a two-step approach to estimate different parts of our model. Step 1 is to obtain the lender side, borrower side and risk assessment model parameters, which are time-specific. In each period, the funding probability and withdrawal probability model parameters can be identified from the variations in loan characteristics, and the corresponding lenders' investment decisions and borrowers' withdrawal decisions. Similarly, parameters in the risk assessment model can be identified from variations in borrowers' default behavior, conditional on loan characteristics. Clearly, the parameter estimates change with adaptive learning algorithms because each algorithm has its own way to select the available data for estimation.

Step 2 is to estimate the structural parameters in Prosper's objective function ($\alpha_1, ..., \alpha_6, \delta_1, \delta_2$). Unlike the predictive models for funding probability, withdrawal probability, default probability and LGD, which we assume Prosper needs to adaptively learn about their parameter values, we assume Prosper's objective function remains unchanged during our sample period. The intercept terms, $\alpha_1, ..., \alpha_6$, are pinned down by

[17] Cogley and Sargent (2005) also use an ensemble modeling approach to model adaptive learning.

the average shares of Prosper rating in the sample period. The parameters, δ_1, δ_2, determine how Prosper makes the trade-off between short-term expected revenue and reporting the true risk.

The identification of the structural parameters in Prosper's objective function is straightforward once we obtain $W, F, ED, ELGD$ from step 1. More importantly, we need to identify which adaptive learning algorithm Prosper uses. We will use a simplified and hypothetical example to show the intuition. Suppose the average default rate for borrowers who work in the restaurant business is 50% from Jan 2020 to Dec 2021 because of the COVID shock. As the economy reopened, the restaurant workers' average default rate drops to 6% from July 2021 to Dec 2021. The default rate changes are shown in Figure 5. At the very beginning of 2022, Prosper needs to determine what rating to assign to a borrower who is a restaurant worker. If Prosper uses the data from Jan 2020 to Dec 2021 to calibrate its lenders, borrowers and default behavior predictive models, it would predict that a restaurant worker to be very risky, and assign a low rating to the borrower. On the contrary, if Prosper only uses the most recent 6 months data to calibrate its predictive models, restaurant workers would receive much better ratings. Hence, different adaptive learning algorithms can generate very different rating choices.

One potential concern is that the changes in the competitive or macro environment may lead to changes in the composition of borrowers and lenders who come to Prosper over time. Hence, the change in the observed distribution of the dependent variables (withdrawal, funding, default and LGD) may reflect shifts in the covariates instead of concept drift. We believe that this should not be a serious issue. First, we have controlled for covariate shift explicitly over time (see Table 3). Those covariates are included in the lender side, borrower side and risk assessment models. Second, it is possible that we do not observe all relevant characteristics of borrowers, and the distribution of unobserved characteristics may change over time. But if this happens, it will change the data generating process over time, which is what concept drift captures. Recall that we use $Y_t = F(X_t, \epsilon_t, \theta_t)$ to describe the data generating process at the beginning of the Introduction (section 1). The shift in the distribution of unobserved characteristics can affect the relationship between Y_t and X_t if unobserved characteristics interact with X_t, and that would reflect in changes in θ_t over time.

7 Results

7.1 Estimates of Prosper's Objective Function

For each data selection/adaptive learning algorithm, we take the estimated model parts (i)-(iii) described in section 4 as given and plug them in Proper's objective function. Then we estimate the structural parameters of its objective function using maximum likelihood. We will use the model fit w.r.t. Prosper's choice on loans' risk ratings to infer which data selection/adaptive learning algorithm is closest to the one used by Prosper. The estimates of α's, δ_1, δ_2, log-likelihood and BICs are shown in Table 5. The parameter estimates are all statistically significant. It is worth highlighting that even though we did not impose any restrictions on the sign of the estimates, the results show $\delta_1 > 0$ and $\delta_2 < 0$ in all adaptive learning algorithms, as the theory suggests.[18]

The structural parameters (δ_1 and δ_2) are all statistically significant in all five adaptive learning algorithms. In terms of the goodness-of-fit, it is clear that E-RPM gives the best log-likelihood and BIC. Hence, we infer that among the five algorithms, E-RPM is the closest to what Prosper uses. Under E-RPM, the estimates of δ_1 (utility weight on short-term revenue) and δ_2 (utility weight on long-term reputation) are 4.26 and -13.98, respectively.

To have an idea about how Prosper weights these two components, let's consider the following example. Suppose we have a rating A loan which requests \$5,000. Its withdrawal probability is 0 and its funding probability is 1. The part-worth utility from the short-term revenue of this loan application is $\frac{5000\times 3\%}{1000} \times 4.26 = 0.639$, where 3% is the origination fee rate for rating A loans. With respect to the part-worth utility on long-term reputation, 1% deviation from the true estimated loss rate will results in $13.98 \times 1\% = 0.140$ loss in Prosper's utility. Compare 0.140 to 0.639, one can see that to Prosper, the reputation term is still quite important relative to the short-term revenue.

7.2 A Closer Look at E-RPM

The estimation results in the previous section suggest that among the set of adaptive learning algorithms being considered, Prosper is most likely to use E-RPM. We now take a closer look at how well E-RPM predicts customer behavior compared with other algorithms, and how Prosper adjusts the weights on each individual sub-model over time.

To measure prediction performance, we use the Receiver Operating Characteristic (ROC) Curve for funding, withdrawal and default predictions (because their outcomes are all binary) and use MSE for LGD prediction (because it is a continuous variable). Roughly speaking, the larger the area under ROC curve (AUC) is, the better the model's overall prediction performance is. Table 6 summarizes the comparison results, and Figures A1, A2, A3 and A4 in Appendix H visualize the ROC comparisons. For funding probabilities, E-RPM's AUC is 0.854 and it outperforms the other four models; MWM with AUC of 0.847 is the runner-up. For default probabilities, E-RPM also performs the best with AUC = 0.617; however, GFM's AUC is 0.615, which comes very close to E-RPM. For withdrawal probability predictions, MWM performs the best with AUC = 0.618, and E-RPM being the runner-up with AUC = 0.595, which comes very close. In addition, for LGD, E-RPM performs the best in terms of MSE. Overall, the comparison results show that E-RPM is able to help Prosper predict customer behavior better than other four algorithms.

[18]Note that there is no reason for us to expect that the structural parameter estimates of Prosper's objective function to be close across different adaptive learning algorithms because the distribution of (F, W, D, LGD) should be very different across algorithms.

Next, we examine how Prosper adaptively adjusts the weights it puts on each individual sub-model in E-RPM. Figure 6 displays the weights on the individual models that predict loan application's funding outcomes. During 2011, the weights on different models change frequently and significantly. The findings are consistent with the fact that Prosper just changed its business model in Dec 2010. In 2011, both lenders and borrowers need to learn how to participate in this platform and their behavior may change more dramatically in this early stage. Hence, 2011 is characterized with fast changing lender and borrower behavior when facing the new business model. In contrast, 2012 is a relatively stable phase. Adjustments of weights are less frequent and smaller than what we observe in 2011, as shown in Figure 6.

It indicates that lenders and borrowers have adapted the new business model, and hence their behavior becomes more stable.

Figure A5 shows changes in weights on withdrawal prediction models. The patterns are similar to the funding probability prediction models. Figure A6 shows the weights for default probability prediction models. Note that the weights for the default prediction models remain unstable by the end of the second year. This is partly because default happens much less often.

7.3 Counterfactual Analysis

Having estimated our model and inferred that E-RPM is the most likely adaptive learning method used by Prosper, we use the estimated model to investigate several counterfactual scenarios: (i) Prosper does not care about expected short-term revenue, (ii) Prosper does not care about the long-term reputation, (iii) Prosper does not adaptively learn at all, and (iv) Prosper knows the true DGP.

Our counterfactual analysis is done period by period. More precisely, when simulating Prosper's rating choices, we assume that for any given period, all the past data remains unchanged, and Prosper only switches their objective function or adaptive learning algorithm in that period. We will report the aggregate changes in rating assignments and expected revenue.

7.3.1 Distribution of Loan Ratings

Experiment (i) assumes that Prosper does not take the short-term revenue into account. We set $\delta_1 = 0$, and keep everything else unchanged. The simulated rating distribution is reported in Figure 7. Interestingly, the rating distribution now has thicker tails on both ends, i.e., Prosper assigns more AA and HR loans (about 900 more for AA loans, and 300 more for HR loans). This make sense because (i) AA loans because they are actually less profitable due to the lower commission fee; (ii) HR loans because they typically have high withdrawal probability and low funding probability, and it is less likely that they get funded and generate revenue. In other words, both AA and HR loans are the least profitable loans. Hence, in the actual scenario where Prosper cares about short-term revenue, it assigns fewer AA and HR loans.

In experiment (ii), we assume Prosper does not care about its long term reputation and only focuses on short-term revenue. We set $\delta_2 = 0$, and keep everything else unchanged. Figure 8 shows that more loan application would be assigned with D and E ratings and fewer loan applications would be assigned with A and B ratings. These changes can be explained by the differences in origination fee across loan ratings: the origination fee rate of ratings D and E is 50% higher than that of ratings A and B (as shown in Table 1). Hence, when we assume Prosper does not care about its long-term reputation, it will try to maximize its short-term revenue by assigning more loan applications with lower ratings so that they can charge more origination fees. Note that Prosper probably does not want to deviate from the actual scenario too much because such deviation will likely encourage borrowers to withdraw their loan applications when they see the higher interest rate associated with a lower rating. By comparing the simulation outcomes of experiments (i) and (ii), we indeed see that their rating distributions are not too far from each other.

Experiment (iii) considers a counterfactual scenario where Prosper does not adaptively learn. We assume Prosper only uses data in the initial period to estimate the models and predict $(F_{ilt}, W_{ilt}, ED_{ilt}, ELGD_{ilt})$, and then we simulate Prosper rating choices using the structural parameter estimated in step 2. The simulated rating distribution is reported in Figure 9. We see that Prosper would assign many more loans with AA, A and B rating, e.g., the number of AA loans would be roughly doubled, the number of A loans would increase by 1,500. The shift of the loan rating distribution has led to the number of HR loans dropped by around 2,000. The results show that if Prosper always holds her beliefs calibrated by the data during the auction period, she will tend to assign higher ratings to loan applications. The main reason is in the auction period, funding probability is quite low in general, especially for loan applications with poorer ratings. Hence, in this counterfactual, Prosper tends to assign loans with better ratings to increase their funding probability, in an attempt to obtain higher expected revenue.

In experiment (iv), we assume Prosper knows the true values of $(F_{ilt}^{True}, W_{ilt}^{True}, D_{ilt}^{True}, LGD_{ilt}^{True})$. How do we apporximate these true values? Let's highlight the differences between the econometrician's estimate of their true values vs. Prosper's prediction about these values. At any given t, as an adaptive learner, Prosper can only use the past data available up to the end of period $t-1$. In contrast, an econometrician can use all data available to her for research, i.e., she can use data before and after t. Because an econometrician can use data from both sides of the neighborhood of t, we apply standard nearest neighbor kernel and use data in [t - 14 days, t + 14 days] to estimate the true funding and withdrawal probability models, and [t - 30 days, t + 30 days] to estimate the true default probability and LGD models. Figure 10 compares the simulated loan rating distribution from this counterfactual with the data, and show that they are not too far from each other. More precisely, compared it with the previous counterfactual simulation results, we find that the deviation from the data in experiment (iv) falls somewhere between those in experiment (i-ii) and (iii).

7.3.2 Simulated Revenue

In this subsection, we analyze how Prosper's revenue changes under different counterfactual experiments. Before we discuss the results, it is important to stress that it is Prosper's prediction about loan application i's funding, withdrawal, default probabilities and LGD with rating l ($F^P_{ilt}, W^P_{ilt}, ED^P_{ilt}, ELGD^P_{ilt}$) that go in her objective function, and these predictions are not supposed to be estimates of their true values (hereafter $F^{True}_{ilt}, W^{True}_{ilt}, D^{True}_{ilt}, LGD^{True}_{ilt}$). In order to map Prosper's choices to expected revenue, we need to estimate ($F^{True}_{ilt}, W^{True}_{ilt}, D^{True}_{ilt}, LGD^{True}_{ilt}$), which we have already explained how to do it in the previous subsection.

Specifically, given an objective function or adaptive learning method (AL), we calculate Prosper's estimates of ($F^P_{ilt}, W^P_{ilt}, ED^P_{ilt}, ELGD^P_{ilt}$). Then, we use the estimates of structural parameters ($\alpha' s, \delta_1, \delta_2$) to obtain Prosper's rating choice probabilities, $P_{ilt}(AL)$. To map $P_{ilt}(AL)$ to expected revenue, we need to use ($F^{True}_{ilt}, W^{True}_{ilt}, D^{True}_{ilt}, LGD^{True}_{ilt}$):

The expected revenue of loan application i with rating l is:

$$E[Revenue_{it}|AL] = \sum_l (O_l + L^{True}_{ilt}) \cdot M_i \cdot F^{True}_{ilt} \cdot (1 - W^{True}_{ilt}) \cdot P_{ilt}(AL).$$

Definitions of the variables can be found in section 4.4. Notice L^{True}_{ilt} has superscript *True* because it is a function of D^{True}_{ilt} and LGD^{True}_{ilt} (see footnote 14). Then we sum up all the expected revenue across loan applications to obtain the total expected revenue in the two year period. Tables 7 reports the simulated expected revenue. We regard the simulated expected revenue under E-RPM as the baseline since E-RPM is the algorithm that closest to what Prosper is using. We compare the expected revenue from other scenarios with the baseline.

Column (2) of Table 7 shows the result if Prosper ignores the short-term revenue (setting $\delta_1 = 0$). The expected revenue drops by about 4.5%. This is consistent with our intuition. As our estimates suggest, Prosper cares about both truthfully reporting each loan application's risk and maximizing its expected revenue. But this "revenue" factor plays a relatively small role in its decision process. The biases caused by the revenue term mainly affect AA and HR rating groups, as shown in Figure 7.

Column (3) of Table 7 shows the simulation results if Prosper does not care about long-term reputation and assign ratings to only maximize its short-term revenue. In this scenario, Prosper's revenue decreases by 0.40%. This finding shows that only focusing on maximizing short-term revenue could backfire. The reason is if Prosper's rating assignments deviate too much from reflecting the borrower's true risk, it increases the likelihood that a loan will not be funded, because either lenders or borrowers may stop participating. Note that lenders have a general idea about how risky a loan is based on the observed borrower's characteristics (e.g., their occupation, income, credit score, etc.). Hence, if Prosper assigns a very good rating to a risky borrower, lenders will likely realize it and choose to not invest. On the other hand, if Prosper assigns a very bad rating to a good borrower, the borrower will likely choose to withdraw because she has a better outside option.

Column (4) of Table 7 shows that if Prosper does not adaptively learn at all, its expected revenue drops by about 3.06%. Hence, actively adapting to the changing market can indeed help a firm increase its revenue. Finally, column (5) of Table 7 shows that if Prosper knows the true data generating process, it could have achieved the highest expected revenue; it gives 8.04% higher expected revenue compared with the benchmark E-RPM model. It is tempting to draw the conclusion that E-RPM is sub-optimal.

8 Conclusion

This is the first structural econometric modeling paper that proposes a framework to infer which adaptive learning algorithm that a firm is most likely to use in handling concept drifts. We refer to this approach as *generalized revealed preference*. To illustrate our approach, we apply it to study Prosper's adaptive learning behavior. We first provide evidence that consumers' borrowing and lending behavior are changing over time. We then show that by analyzing Prosper's choices and the lens of a structural model, we can uncover not only the parameters in its objective function, but also the way Prosper selects data to make decisions. Among the five adaptive learning algorithms we consider, we find that an ensemble method that relies on macroeconomic conditions of the data (E-RPM) is the most plausible method adopted by Prosper. In the counterfactual exercises, we demonstrate how Prosper's rating assignments and expected revenue change in different scenarios. In particular, we find Prosper's expected revenue will decrease if it does not care about short-term revenue or does not adaptively learn at all. On the other hand, Prosper's expected revenue will increase if it does not care about long-term reputation or knows the true data generating process.

To conclude, we reiterate that the concept drifts could happen quite often. At the same time, it can be very difficult for a firm to figure out how the underlying data generating process changes over time, in particular, when it needs to make real time decisions. Hence, we do not assume that Prosper uses an optimal adaptive learning method. An optimal adaptive learning method would require Prosper to know exactly how and when concept drifts happen. Such an assumption does not seem very plausible here, especially because we are studying the period when Prosper just started to use posted price business model. Hence, we hypothesize that during our sample period, Prosper uses one adaptive learning algorithm to address concept drifts, even though it may not be a perfect solution. Our research does not address how Prosper settled with this algorithm. Also, it is certainly possible that Prosper may change its adaptive learning algorithm from time to time. Studying this problem is beyond the scope of this paper. We will leave these questions for future research.

References

Aguirregabiria, V. and Jeon, J. (2020). Firms' beliefs and learning: Models, identification, and empirical evidence. *Review of Industrial Organization*, 56(2):203–235.

Bajari, P., Benkard, C. L., and Levin, J. (2007). Estimating dynamic models of imperfect competition. *Econometrica*, 75(5):1331–1370.

Balakrishnan, S., Madigan, D. B., and Bartlett, P. (2008). Algorithms for sparse linear classifiers in the massive data setting.

Chan, T., Hamdi, N., Hui, X., and Jiang, Z. (2020). Digital verification and inclusive access to credit: An empirical investigation. *Available at SSRN 3556554*.

Chauvet, M. (1995). *An econometric characterization of business cycle dynamics with factor structure and regime switching*. PhD thesis, University of Pennsylvania.

Chauvet, M. and Piger, J. (2008). A comparison of the real-time performance of business cycle dating methods. *Journal of Business & Economic Statistics*, 26(1):42–49.

Ching, A. T. (2010). A dynamic oligopoly structural model for the prescription drug market after patent expiration. *International Economic Review*, 51(4):1175–1207.

Cogley, T. and Sargent, T. J. (2005). The conquest of us inflation: Learning and robustness to model uncertainty. *Review of Economic dynamics*, 8(2):528–563.

Crespo, F. and Weber, R. (2005). A methodology for dynamic data mining based on fuzzy clustering. *Fuzzy Sets and Systems*, 150(2):267–284.

Cunningham, S. (2014). Default rates at lending club y prosper: When loans go bad. *Lending Memo*.

Dew, R., Ansari, A., and Li, Y. (2020). Modeling dynamic heterogeneity using gaussian processes. *Journal of Marketing Research*, 57(1):55–77.

Domingos, P. and Pazzani, M. (1997). On the optimality of the simple bayesian classifier under zero-one loss. *Machine learning*, 29(2):103–130.

Doraszelski, U., Lewis, G., and Pakes, A. (2018). Just starting out: Learning and equilibrium in a new market. *American Economic Review*, 108(3):565–615.

Erdem, T., Imai, S., and Keane, M. P. (2003). Brand and quantity choice dynamics under price uncertainty. *Quantitative Marketing and economics*, 1(1):5–64.

Evans, G. W., Honkapohja, S., et al. (2013). Learning as a rational foundation for macroeconomics and finance. *Rethinking expectations: The way forward for macroeconomics*, pages 68–111.

Forman, G. (2006). Tackling concept drift by temporal inductive transfer. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 252–259.

Freedman, S. M. and Jin, G. Z. (2011). Learning by doing with asymmetric information: Evidence from prosper. com. Technical report, National Bureau of Economic Research.

Friedman, N., Geiger, D., and Goldszmidt, M. (1997). Bayesian network classifiers. *Machine learning*, 29(2):131–163.

Fu, R., Huang, Y., and Singh, P. V. (2021). Crowds, lending, machine, and bias. *Information Systems Research*, 32(1):72–92.

Gama, J., Medas, P., Castillo, G., and Rodrigues, P. (2004). Learning with drift detection. In *Brazilian symposium on artificial intelligence*, pages 286–295. Springer.

Hamilton, J. D. and Lin, G. (1996). Stock market volatility and the business cycle. *Journal of applied econometrics*, 11(5):573–593.

Hitsch, G. J. (2006). An empirical model of optimal dynamic product launch and exit under demand uncertainty. *Marketing Science*, 25(1):25–50.

Hoens, T. R., Polikar, R., and Chawla, N. V. (2012). Learning from streaming data with concept drift and imbalance: an overview. *Progress in Artificial Intelligence*, 1(1):89–101.

Humpherys, S. L., Moffitt, K. C., Burns, M. B., Burgoon, J. K., and Felix, W. F. (2011). Identification of fraudulent financial statements using linguistic credibility analysis. *Decision Support Systems*, 50(3):585–594.

Iyer, R., Khwaja, A. I., Luttmer, E. F., and Shue, K. (2016). Screening peers softly: Inferring the quality of small borrowers. *Management Science*, 62(6):1554–1577.

Kawai, K., Onishi, K., and Uetake, K. (2021). Signaling in online credit markets. Technical report, National Bureau of Economic Research.

Kelly, M. G., Hand, D. J., and Adams, N. M. (1999). The impact of changing populations on classifier performance. In *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 367–371.

Kim, C.-J. and Nelson, C. R. (1998). Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime switching. *Review of Economics and Statistics*, 80(2):188–201.

Klinkenberg, R. (2004). Learning drifting concepts: Example selection vs. example weighting. *Intelligent data analysis*, 8(3):281–300.

Koren, Y. (2009a). The bellkor solution to the netflix grand prize. *Netflix prize documentation*, 81(2009):1–10.

Koren, Y. (2009b). Collaborative filtering with temporal dynamics. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 447–456.

Kozlowski, J., Veldkamp, L., and Venkateswaran, V. (2020). The tail that wags the economy: Beliefs and persistent stagnation. *Journal of Political Economy*, 128(8):2839–2879.

Liechty, J. C., Fong, D. K., and DeSarbo, W. S. (2005). Dynamic models incorporating individual heterogeneity: Utility evolution in conjoint analysis. *Marketing Science*, 24(2):285–293.

Lin, M., Prabhala, N. R., and Viswanathan, S. (2013). Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending. *Management Science*, 59(1):17–35.

Lin, M. and Viswanathan, S. (2016). Home bias in online investments: An empirical study of an online crowdfunding market. *Management Science*, 62(5):1393–1414.

Liu, X., Wei, Z., and Xiao, M. (2020). Platform mispricing and lender learning in peer-to-peer lending. *Review of Industrial Organization*, 56(2):281–314.

McCullagh, P. and Nelder, J. A. (2019). *Generalized linear models*. Routledge.

Ni, J. and Xin, Y. (2020). Financing micro-entrepreneurship in online crowdfunding markets: Local preference versus information frictions. *Johns Hopkins Carey Business School Research Paper*, (20-09).

Pechenizkiy, M., Bakker, J., Žliobaitė, I., Ivannikov, A., and Kärkkäinen, T. (2010). Online mass flow prediction in cfb boilers with explicit detection of sudden concept drift. *ACM SIGKDD Explorations Newsletter*, 11(2):109–116.

Sargent, T. J. et al. (1993). Bounded rationality in macroeconomics: The arne ryde memorial lectures. *OUP Catalogue*.

Schlimmer, J. C. and Granger, R. H. (1986). Incremental learning from noisy data. *Machine learning*, 1(3):317–354.

Sornette, D. (2004). Why stock markets crash: critical events in complex financial systems. *Physics Today*, 57(3):78–79.

Sriram, S., Chintagunta, P. K., and Neelamegham, R. (2006). Effects of brand preference, product attributes, and marketing mix variables in technology product markets. *Marketing Science*, 25(5):440–456.

Tsymbal, A. (2004). The problem of concept drift: definitions and related work. *Computer Science Department, Trinity College Dublin*, 106(2):58.

Viaene, S., Derrig, R. A., Baesens, B., and Dedene, G. (2002). A comparison of state-of-the-art classification techniques for expert automobile insurance claim fraud detection. *Journal of Risk and Insurance*, 69(3):373–421.

Wang, H., Fan, W., Yu, P. S., and Han, J. (2003). Mining concept-drifting data streams using ensemble classifiers. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 226–235.

Wei, Z. and Lin, M. (2017). Market mechanisms in online peer-to-peer lending. *Management Science*, 63(12):4236–4257.

Yao, S., Mela, C. F., Chiang, J., and Chen, Y. (2012). Determining consumers' discount rates with field studies. *Journal of Marketing Research*, 49(6):822–841.

Zhang, J. and Liu, P. (2012). Rational herding in microloan markets. *Management science*, 58(5):892–912.

Table 1: Interest Rates and Service Fee Rates

	AA	A	B	C	D	E	HR
Average Interest Rate	7.74%	10.80%	15.88%	19.92%	24.85%	30.43%	31.78%
Average Annual Return	7.03%	8.58%	10.11%	10.99%	12.15%	13.13%	12.17%
Origination Fee Rate	0.5%	3%	3%	4.5%	4.5%	4.5%	4.5%
Posted Loss Rate	1.42%	3.03%	5.56%	7.94%	10.83%	14.41%	17.08%

Notes: The interest rate of each rating is set by Prosper. Annual return is calculated by subtracting real loss rate from interest rate. The real loss rate is calculated from the data since we can observe all the loans repayment outcomes. If a loan defaults, we can observe the principal loss as well. The posted loss rate in this Table is reported by Prosper. For loan applications with the same rating, Prosper posts the same loss rate. The posted loss rate is not necessary to coincide with the real loss rate. This is why the the average annual return and the posted loss rate do not add up to the average interest rate.

Table 2: Loan Application Status by Prosper Ratings

Prosper Rating	Completed	Expired	Withdrawn	Total
AA	1,030 (69.59%)	165 (11.15%)	285 (19.26%)	1,480
A	2,714 (72.55%)	408 (10.91%)	619 (16.55%)	3,741
B	2,965 (72.99%)	351 (8.64%)	746 (18.37%)	4,062
C	2,729(75.49%)	310 (8.58%)	576 (15.93%)	3,615
D	4,675 (71.23%)	573(8.73%)	1,315 (20.04%)	6,563
E	3,402 (76.43%)	263(5.91%)	786 (17.66%)	4,451
HR	4,762 (60.32%)	1,635 (20.71%)	1,498 (18.97%)	7,895
Total	22,277(70.04%)	3,705 (11.65%)	5,825 (18.31%)	31,807

Notes: Percentages are calculated rowwise. For instance, 1,030 completed AA loan applications account for 69.59% of the 1,480 total AA loan applications.

Table 3: List of Variables

Variable Name	Description
	Loan Application Features
Loan Application Number	The unique identifier for a loan application.
Amount Requested*	Amount requested by borrower in a loan application.
Amount Funded	Amount funded by lenders.
Prosper Rating	A series of dummy variables assigned by Prosper to indicate borrower's credit grade (AA, A, B, C, D, E, HR). AA indicates the best credit and HR indicates the worst credit.
Origination Fee Rate*	Origination fees are a percentage of the amount borrowed varying by Prosper Rating. It ranges from 0.5% to 4.5% according to the loan application's rating.
Interest Rate*	The interest rate a borrower needs to pay for this loan.
Loan Application Status	Withdrawn - The loan application was withdrawn by customer request. Expired - The loan application failed to fund in time. Completed - The loan application ran to completion and funded. Cancelled - The loan application was canceled by Prosper.
	Borrower Characteristics
Bankcard Utilization*	The percentage of available revolving credit that is utilized at the time the credit profile was pulled.
Home Owner*	A borrower will be classified as a homowner if they have a mortgage on their credit profile or provide documentation confirming they are a homeowner.
Posted Loss Rate*	The posted principal loss percentage on default. Prosper reports the same posted loss rate for loan applications with the same rating.
Credit Score Range Lower*	The lower value of the range of the borrower's credit score provided by the consumer credit rating agency
Current Credit Lines*	Number of current credit lines at the time the credit profile was pulled.
Credit Lines Last 7 Years*	Number of credit lines in the past seven years at the time the credit profile was pulled.
Current Delinquencies*	Number of accounts delinquent at the time the credit profile was pulled.
Delinquencies Last 7 Years*	Number of delinquencies in the past 7 years at the time the credit profile was pulled.
Monthly Income*	The monthly income the borrower stated at the time the loan application was created.
Income Verifiable*	Prosper will request documents such as recent paystubs, tax returns, or bank statements to verify a borrower's income.
Inquiries Last 6 Months*	Number of inquiries in the past five months at the time the credit profile was pulled.
Total Inquiries*	Total number of inquiries at the time the credit profile was pulled.
Open revolving accounts*	Number of open revolving accounts at the time the loan application was created.
Revolving Credit Balance*	The monetary amount of revolving credit balance at the time the loan application was created.
Prior Prosper Loans*	Number of Prosper loans the borrower has borrowed at the time they created this loan application.

Table 3: List of Variables (Continued)

Prior Prosper Loans Active*	Number of ongoing Prosper loans the borrower has at the time they created this loan application.
Monthly Debt*	The amount of debt the borrower needs to pay each month.
Month*	The month in which the loan application was submitted.
Group*	Equals 1 if the borrower belongs to a group. A Group is a collection of Members who share a common interest or affiliation.
Real Estate Balance*	The mortgage balance.
Market Level Variables	
Mortgage Rate*	30-Year Fixed Rate Mortgage Average in the United States.
TED Spread*	The difference between the interest rate on short-term US government debt and the interest rate on interbank loans.
Adjusted Closing Price*	S&P 500 daily closing price.

Notes: Variables with * enter our lender side, borrower side and risk access model to help us calculate each loan application's funding, withdrawal and default probabilities. We also use these variables in the logistic regressions in section 3.3 to show how to detect concept drift.

Table 4: Variables: Summary Statistics

Variable	Mean	SD	Max	Min
Panel A: Loan Application Features				
Amount Requested ($1000)	6.83	0.095	25	2
Amount Funded ($1000)	4.98	4.58	25	0
Funded Percentage (%)	75.48	39.04	100	0
Interest Rate (%)	23.34	8.14	32.20	5.99
Origination Fee (%)	4.36	0.92	4.95	0.5
Panel B: Borrower Credit Variables				
Bank Card Utilization (%)	51.50	33.12	223	0
Home Owner (0/1)	0.49	0.50	1	0
Posted Loss Rate (%)	10.53	5.29	20.3	0.49
Credit Score Range Lower	697.47	44.55	778	600
Current Credit Lines	9.11	5.39	56	0
Credit Lines Last 7 Years	25.76	13.93	120	2
Current Delinquencies	0.48	1.32	27	0
Delinquencies Last 7 Years	0.48	1.32	27	0
Monthly Income ($1000)	5.67	13.84	1,750	0
Income Verifiable (0/1)	0.86	0.35	1	0
Inquiries Last 6 Months	1.28	1.74	27	0
Total Inquiries	4.33	4.07	73	0
Prior Prosper Loans	0.33	0.69	7	0
Prior Prosper Loans Active	0.15	0.35	2	0
Monthly Debt ($1000)	0.87	1.35	100.28	0
Real Estate Balance ($1000)	107.60	162.81	3,830.08	0
Open revolving accounts	6.19	4.31	47	0
Revolving Credit Balance ($1000)	19.04	37.15	1,066.76	0
Group (0/1)	0.033	0.18	1	0
Panel C: Market Level Variables				
Mortgage Rate (%)	4.38	0.51	5.33	3.58
TED Spread (%)	0.33	0.11	0.57	0.14
Adjusted Closing Price	1,322	78.30	1,466	1,099

Notes: Summary statistics in this table are calculated based on loan applications submitted between Dec 19, 2010 and Dec 31, 2012.

Table 5: Estimation Results of Prosper's Objective Function

	EWM	MWM	RPM	GFM	E-RPM
Log-Likelihood	-51999.34	-53343.86	-51121.85	-52219.56	-49314.87
BIC	104081.62	106770.66	102326.64	104522.06	98712.68
$\hat{\delta}_1$	2.21 (0.378)	1.58 (0.132)	4.41 (0.169)	2.09 (0.401)	4.26 (0.185)
$\hat{\delta}_2$	-11.0 (0.161)	-6.41 (0.120)	-10.44 (0.145)	-10.94 (0.165)	-13.98 (0.154)
$\hat{\alpha}_1$	-2.33 (0.035)	-1.44 (0.033)	-1.70 (0.037)	-2.38 (0.035)	-1.41 (0.037)
$\hat{\alpha}_2$	-1.41 (0.024)	-0.69 (0.021)	-1.27 (0.023)	-1.44 (0.024)	-0.99 (0.023)
$\hat{\alpha}_3$	-1.30 (0.023)	-0.69 (0.021)	-1.18 (0.022)	-1.30 (0.023)	-0.96 (0.023)
$\hat{\alpha}_4$	-1.30 (0.024)	-0.84 (0.022)	-1.27 (0.023)	-1.30 (0.024)	-1.06 (0.023)
$\hat{\alpha}_5$	-0.61 (0.020)	-0.32 (0.018)	-0.57 (0.019)	-0.59 (0.020)	-0.42 (0.019)
$\hat{\alpha}_6$	-0.74 (0.020)	-0.67 (0.020)	-0.75 (0.020)	-0.74 (0.020)	-0.68 (0.020)

Notes: For GFM, the forgetting factor takes value 0.997, 0.99, 0.995 and 0.997 for the borrower side, lender side, default prediction and LGD prediction models, respectively. In MWM, the window size equals to 90, 210, 1080 and 1440 days for the borrower side, lender side, default prediction and LGD prediction models, respectively. The model comparison is done by evaluating their performance on one-period ahead out of sample prediction. For instance, when predicting loan applications' rating assignment in period $t+1$, we use data until period t to calibrate our model and then make predictions. After we have our model's prediction accuracy for period $t+1$, we move one period forward to period $t+2$ and re-estimate our model by taking data from period $t+1$ into account and make predictions for period $t+2$. We repeat this process until the last period of our data. The length of period for different adaptive learning models can be found in sections 5.1-5.5. We normalize $\alpha_7 = 0$. BICs are calculated based on one-period ahead out of sample log-likelihood. Standard errors are in brackets.

Table 6: Comparison of Prediction Results

	EWM	MWM	RPM	GFM	E-RPM
Funding AUC	0.712	0.847	0.796	0.638	0.854
Withdrawal AUC	0.534	0.618	0.587	0.535	0.595
Default AUC	0.597	0.609	0.603	0.615	0.617
Loss Rate Given Default MSE	0.045	0.092	0.110	0.095	0.035

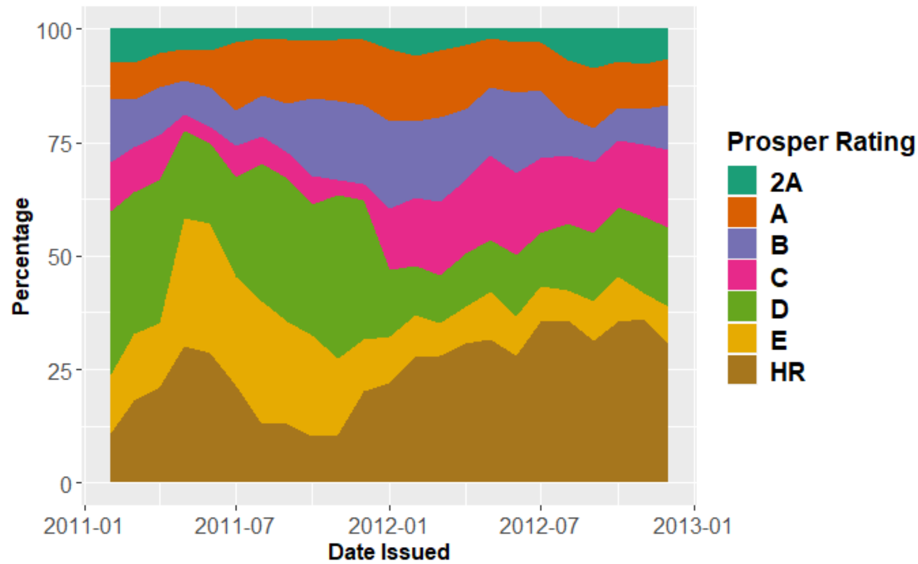
Notes: AUC represents area under ROC. The larger the AUC is, the better prediction performance the corresponding method has.

Table 7: Simulated Revenue in Counterfactual Exercises (E-RPM as Baseline)

	E-RPM	$\delta_1 = 0$	$\delta_2 = 0$	No AL	True DGP
Revenue ($)	7,380,725	7,066,938	7,348,733	7,154,574	7,974,155
ΔRevenue ($)	0	-313,787	-31,992	-226,151	593,430
ΔRevenue (%)	0	-4.25%	-0.40%	-3.06%	8.04%

Notes: This table shows the simulated expected revenue in different scenarios. We use E-RPM as the comparison baseline. $\delta_1 = 0$ represents the scenario where Prosper did not care about short-term revenue. $\delta_2 = 0$ represents the scenario where Prosper did not care about long-term reputation. No AL represents the scenario where Prosper did not adaptively learn at all. True DGP represents the scenario where Prosper knows the true data generating process.

Figure 1: Rating Distribution of Loans Over Time



Notes: Each color represents the percentage of a certain rating category over time.

Figure 2: Conversion Rate and Default Rate

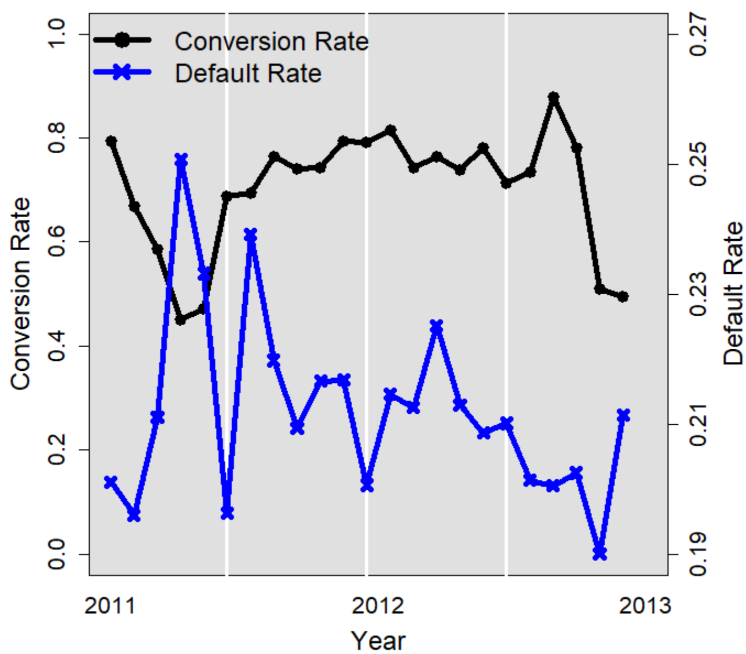
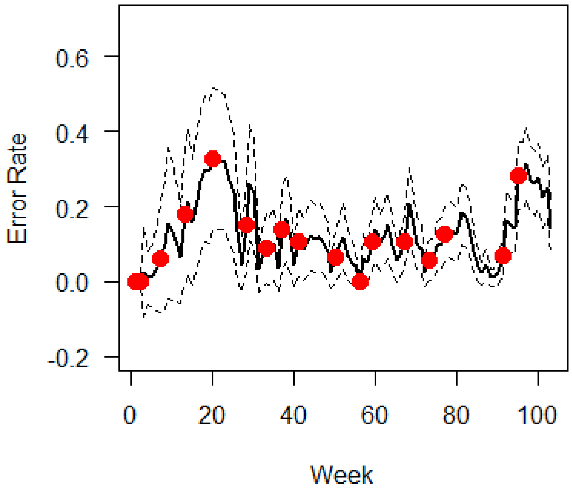
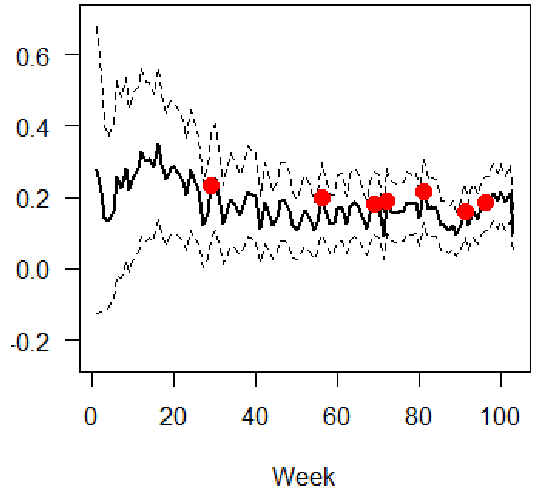


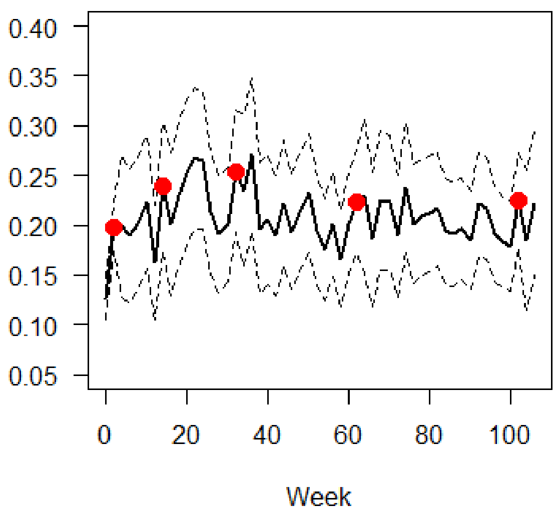
Figure 3: Concept Drift Detection



(a) Prediction Error Rate in Loan Completion



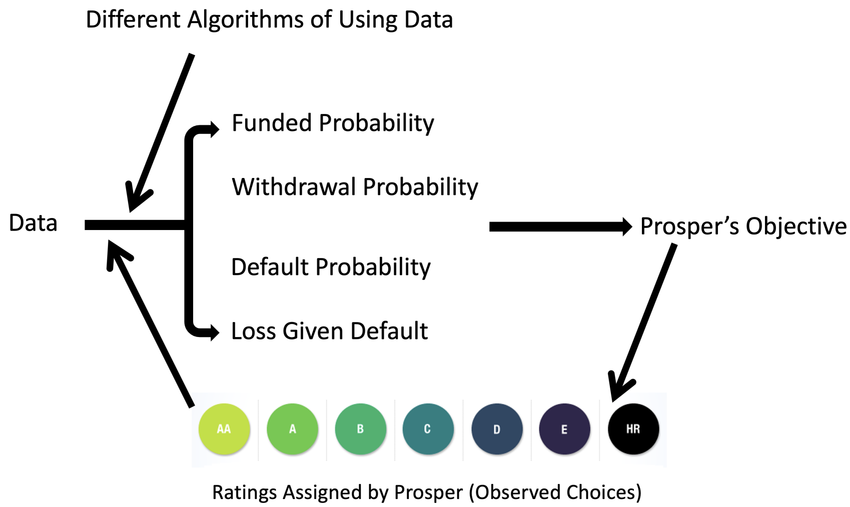
(b) Prediction Error Rate in Withdrawal Decisions



(c) Prediction Error Rate in Default Outcomes

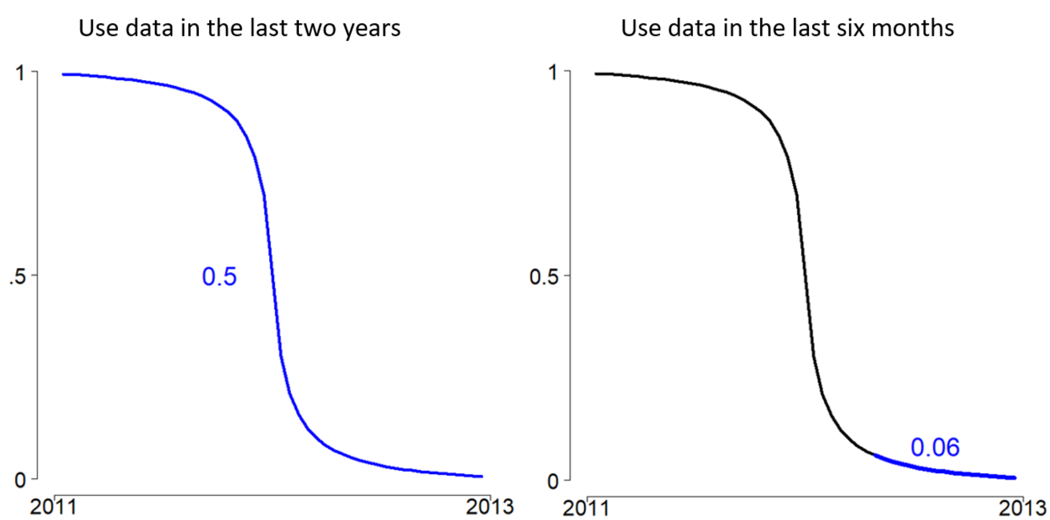
Notes: The red dots indicate when concept drifts are detected by Gama et al. (2004) Test. Concept drifts in lenders' investing behavior, borrowers' withdrawal and default behavior were detected 17, 7 and 5 times, respectively. Dotted lines represent the 95% confidence interval.

Figure 4: Model Framework



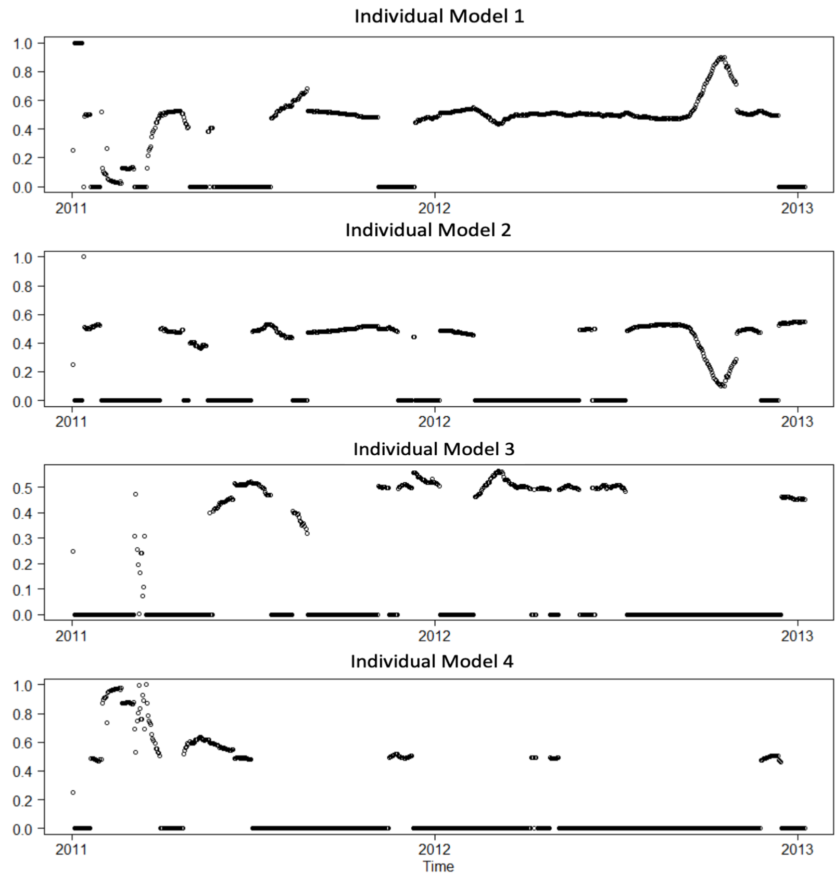
Notes: We use lender side model and borrower side mdoel to compute each loan application's funding probability and withdrawal probability, respectively. Default probability and LGD are computed using Naive Bayes classifier. Prosper then puts these four variables into its objective function and chooses one of the seven ratings to maximize its objective. What we are interested in is to infer the way Prosper uses historical data to compute funding, withdrawal, default probabilities and LGD. Because these calibrated components are a function of the way Prosper weights past data, Prosper's choice in classifying loans should reveal their data weighting mechanism.

Figure 5: Average Default Rate of Restaurant Workers



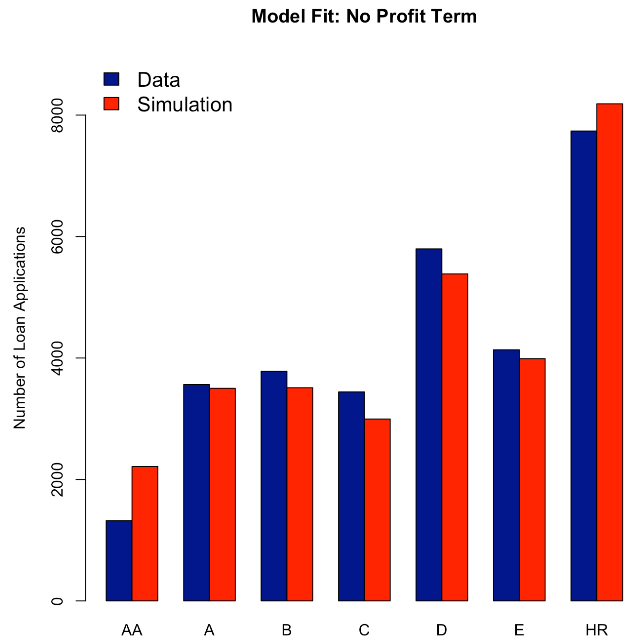
Notes: This is a simplified example. In this example, the average default rate of restaurant workers is 50% from Jan 2020 to Dec 2021 and 6% from Jul 2021 to Dec 2021.

Figure 6: Weights on Individual Funding Model in E-RPM



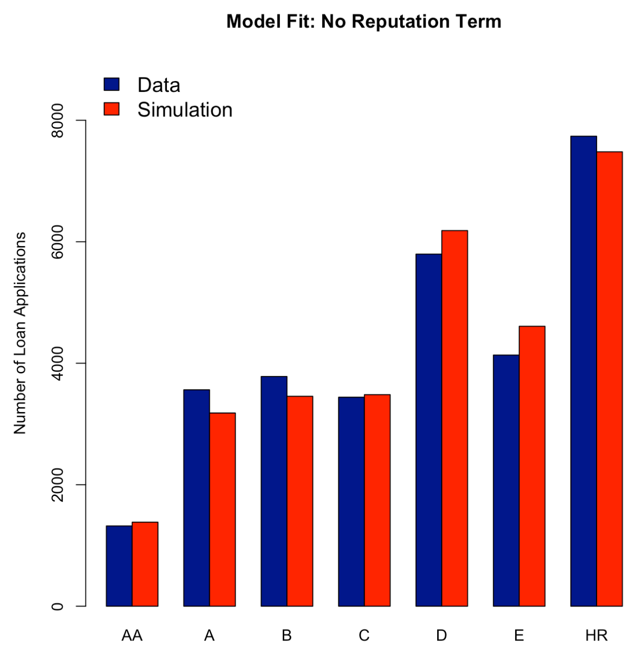
Notes: This figure shows the weights E-RPM puts on each individual model on funding predictions. The weights are updated every day.

Figure 7: Rating Distribution: No Revenue Term



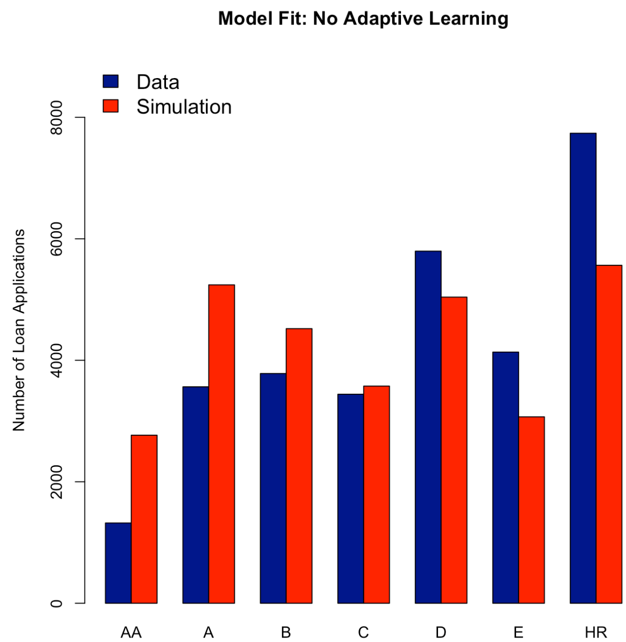
Notes: Red bars represent the expected rating distribution if Prosper does not care about short-term revenue when making rating assignments.

Figure 8: Rating Distribution: No Reputation Term



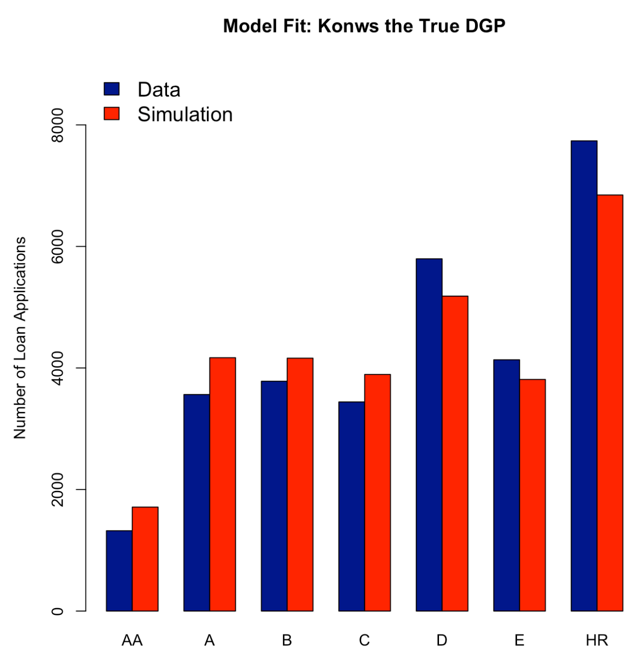
Notes: Red bars represent the expected rating distribution if Prosper does not care about long-term reputation when making rating assignments.

Figure 9: Rating Distribution: No Adaptive Learning



Notes: Red bars represent the expected rating distribution if Prosper did not adaptively update its model but kept using the model calibrated using data from the auction period.

Figure 10: Rating Distribution: Knows the True DGP



Notes: Red bars represent the expected rating distribution if Prosper knows the true DGP and updates its models accordingly.

Appendix

A Concept Drift Detection

We follow Gama et al. (2004) to test for the existence of concept drift. As an example, let us consider how to test the existence of concept drift in lenders' investing behavior. Consider a set of loan applications, in the form of pairs (X_i, y_i), where X_i represents borrower i's characteristics, and y_i takes value 1 if the loan application is funded and 0 otherwise. Suppose that we use a logistic regression model to predict each loan application's funding probability. Let $\hat{y}_i$ denote the predicted funding outcome, with the value 1 if the loan application is funded and 0 otherwise. The event of false prediction, $\hat{y}_i \neq y_i$, is a random variable from Bernoulli trials. For a sequence of n_t observations, the number of false prediction events follows a Binomial distribution. Let p_t denote the percentage of false predictions (error rate) in period t, the corresponding standard deviation is $s_t = \sqrt{p_t(1-p_t)/n_t}$.

If lenders' investing behavior is stationary, the error rate (p_t) of our model should remain stable across different periods. A significant change in the error rate suggests a change in lenders' investing behavior. The drift detection method manages two threshold values, p_{min} and s_{min}, during the training of the learning algorithm. When $t = 1$, we set $p_{min} = p_1$ and $s_{min} = s_1$. As more loan applications arrive, these values are updated when $p_t + s_t$ is lower than $p_{min} + s_{min}$. In particular, if in period t we have $p_t + s_t < p_{min} + s_{min}$, then we update p_{min} and s_{min} as $p_{min} = p_t$ and $s_{min} = s_t$. For a sufficiently large number of observations (>30), the Binomial distribution in each period is closely approximated by a Normal distribution with the same mean and variance. Under the hypothesis that concept drift does not happen, the confidence interval for p_t can be approximated by $p_t \pm \iota \cdot \delta_t$. The parameter ι depends on the desired confidence interval. A commonly used confidence level for warning is 95% with the threshold $p_t + s_t \geq p_{min} + 2 \times s_{min}$, and for concept drift detection is 99% with the threshold $p_t + s_t \geq p_{min} + 3 \times s_{min}$. To be specific, for period t, our concept drift detection system will be in one of the following three states:

(1) $p_t + s_t < p_{min} + 2 \times s_{min}$. The system is stable. That is, we cannot reject the hypothesis that lenders' investing behavior does not change.

(2) $p_t + s_t \geq p_{min} + 3 \times s_{min}$. The error rate has increased significantly. With a very high probability that concept drift is detected.

(3) $p_{min} + 2 \times s_{min} \leq p_t + s_t < p_{min} + 3 \times s_{min}$. This is the warning state, which is in between of the two previous states. Intuitively, there is some evidence that concept drift might happen, but the evidence is not particularly strong yet.

Suppose the warning level is reached in period t_w and concept drift is detected in period t_d. Then we believe the underlying data generation process for our observations has changed from t_d onwards. Gama, et al.

(2004) then propose using observations between t_w and t_d to train our model and make predictions for future observations, until another concept drift is detected.

B Update Naive Bayes Classifier in Each Period

In the actual implementation, we keep updating the prior ω_k and posterior P_k over time. We use $\mathbf{\Omega_t} = \{\omega_{1t}, \omega_{2t}, ..., \omega_{Kt}\}$ represent prior belief at time t and $\boldsymbol{P}_{kt} = (P_{k1t}, P_{k2t}, ..., P_{kmt})$ represent the joint distribution of all m characteristics in class k at time period t.

Let $\mathbf{\Omega_t} = \{\omega_{1t}, \omega_{2t}, ..., \omega_{Kt}\}$ represent prior belief at time t and $\boldsymbol{P}_{kt} = (P_{k1t}, P_{k2t}, ..., P_{kmt})$ represent the joint distribution of all m characteristics in class k at time period t. In particular, $P_{kjt} = (p_{kj1t}, ..., p_{kjn_jt})$ represent characteristic j's distribution in class k at the beginning of period t. The prior belief of class k in period t can be calculated as the empirical frequency of observing class k up to time t:

$$\omega_{kt} = \frac{1}{\sum_{z=1}^{t} I_z} \sum_{z=1}^{t} \sum_{i=1}^{I_z} \mathbf{1}[y_i = k], k = 1, 2, ..., K \tag{10}$$

where I_z represents the number of loan applications in period z. The probability of observing characteristic j's h level in class k in period t, p_{kjht}, can be approximated by the empirical frequency of observing X_{ijh} up to time t,

$$p_{kjht} = \frac{\sum_{z=1}^{t} \sum_{i=1}^{I_z} \mathbf{1}[y_i = k] \cdot X_{ijh}}{\sum_{z=1}^{t} \sum_{i=1}^{I_z} \mathbf{1}[y_i = k]}, \tag{11}$$

where $k = 1, 2, ..., K; j = 1, 2, ..., m; h = 1, 2, ..., n_j$.

C Receiver Operating Characteristic (ROC) Curve

In a classification problem, we need to build a mapping between instances and certain categories. If the output of the classifier happens to be continuous, the classification boundary between classes must be determined by a threshold value. For example, in a binary classification problem, the categorical outcome can be either positive (P) or negative (N). If the predicted outcome is P and the actual value is also P, then it is called a true positive (TP); however, if the actual value is n then it is said to be a false positive (FP). Conversely, a true negative (TN) has occurred when both the prediction outcome and the actual value are N, and false negative (FN) is when the prediction outcome is n while the actual value is p.The true positive rate (TPR) is defined as the number of true positive divided by the number of positive and the false positive rate (FPR) is defined as the number of false positive divided by the number of negative.

To draw an ROC curve, only the TPR and FPR are needed. The TPR defines how many correct positive results occur among all positive samples available during the test. FPR, on the other hand, defines how many incorrect positive results occur among all negative samples available during the test.

An ROC space is defined by FPR and TPR as X and Y axes, respectively, which depicts relative trade-offs between true positive (benefits) and false positive (costs). Intuitively, the more the ROC curve to the upper left, the more prediction power the corresponding classifier has. The best possible prediction method would yield a point at coordinate (0,1) of the ROC space, representing no false negatives and no false positives. The (0,1) point is also called a perfect classification. A random guess would give a point along a diagonal line from the left bottom to the top right corners.

D Smoothed U.S. Recession Probabilities

Smoothed recession probabilities for the United States are obtained from a dynamic-factor markov-switching model applied to four monthly coincident variables: non-farm payroll employment, the index of industrial production, real personal income excluding transfer payments, and real manufacturing and trade sales. This model was originally developed in Chauvet (1998).

In the proposed model, business cycles are empirically characterized by a dynamic factor model with regime switching. The dynamic factor is an unobserved variable that summarizes the common cyclical movements of some coincident macroeconomic variables. This factor is subject to discrete shifts in order to capture the asymmetric nature of business cycle phases - expansions are gradual and display a high mean duration while recessions are shorter and steeper. Hence, the approach used in this paper models the idea of business cycles as the simultaneous movement of economic activity in various sectors by using an unobserved dynamic factor. In addition, the asymmetric nature of expansions and contractions is captured by assuming that the underlying factor switches regimes according to a Markov process.

To estimate the model it is necessary to make inferences about both the unobserved nonlinear factor and the latent Markov state. The estimation procedure consists of a combination of Hamilton's algorithm (Hamilton 1989) and a nonlinear discrete version of the Kalman filter, as proposed by Kim (1994).

The goals in building a dynamic factor model with regime switching are to obtain optimal inferences of business cycle turning points. Empirical results show that the inferred probabilities from this model are strongly correlated with the NBER business cycle dates and all recessions are well characterized for both quarterly and monthly data. In comparison to traditional approaches, the dynamic-factor markov-switching model allows a more rigorous and timely method for real time assessment of the economy, and results can be consistently reproduced.

even though assigning rating C yields higher utility contribution due to the revenue term (i.e., $0.25 = 1.18 - 0.93$), it will lead to a larger decline in utility contribution due to the reputation term (i.e., $-0.29 = -0.32 + 0.03$). Hence, Prosper is more likely to assign rating B instead of rating C.

E Estimation

E.1 Logistics Regression with A Forgetting Factor

This section demonstrates the estimation procedure of the funded and withdrawal probabilities using GFM. Given the assumption that the residual terms follow type I extreme value distribution in equation 1, we basically need to estimate two logistic regression models in each period on the borrower and lender sides. The estimation process could be time consuming because we need to update the weight on each observation and re-estimate the model in each period. To make the estimation more efficient, we employ an estimation scheme proposed by Balakrishnan et al. (2008). This method is based on a quadratic Taylor approximation to the log-likelihood. A forgetting factor can be easily incorporated into this estimation scheme and the model can be recursively estimated, which significantly reduces the computational burden.

With a slight abuse of notation, we present the details of adaptive logistic regression below. Suppose the data set we have is $\{X, Y\} = \{X_i, y_i\}_{i=1}^n$. y_i is the class label which takes value 0 or 1. Under the logistic regression model, the log-likelihood function is:

$$\log L(\beta|X, Y) = \sum_{i=1}^{n} f_i(\beta^T X_i) \tag{12}$$

where

$$f_i(\beta^T X_i) = \begin{cases} \log \Phi(\beta^T X_i), & \text{if } y_i = 1 \\ \log(1 - \Phi(\beta^T X_i)) & \text{o.w.} \end{cases} \tag{13}$$

Assume that our current estimate of β is $\widetilde{\beta}$, Balakrishnan et al. (2008) point out that each f_i may be replaced by the first few terms of its Taylor expansion around the current location $z_i = \beta^T X_i$. Adding up all f_i's, we get

$$\log L(\beta|X, C) \approx \frac{1}{2}\beta^T \Psi(\widetilde{\beta})\beta - \beta^T \theta(\widetilde{\beta}) \tag{14}$$

where

$$\Psi(\widetilde{\beta}) = \sum_{i=1}^{n} a(\widetilde{\beta}^T X_i) X_i X_i^T, \quad \theta(\widetilde{\beta}) = \sum_{i=1}^{n} b(\widetilde{\beta}^T X_i, y_i) X_i \tag{15}$$

and

$$a(\widetilde{\beta}^T X_i) = -\Phi(\widetilde{\beta}^T X_i)(1 - \Phi(\widetilde{\beta}^T X_i))), \quad b(\widetilde{\beta}^T, y_i) = \Phi(\widetilde{\beta}^T X_i) - y_i + \widetilde{\beta}^T X_i a(\widetilde{\beta}^T X_i) \tag{16}$$

We can get a better estimation of β by maximizing the log-likelihood function with respect to β:

$$\widetilde{\beta}^* = \underset{\beta}{\operatorname{argmax}} \left(\frac{1}{2}\beta^T \Psi(\widetilde{\beta})\beta - \beta^T \theta(\widetilde{\beta})\right) \tag{17}$$

Anagnostopoulos, et al. (2009) modify this estimation scheme to make it recursively update its parameter estimates upon the arrival of new data points. Assume our current estimate of β is $\widehat{\beta}_t$ and the new data point arrives is (X_{n+1}, y_{n+1}), our new estimate of β is updated as follows:

$$\widehat{\Psi}_{n+1} = \widehat{\Psi}_n + a(\widehat{\beta}_n^T X_{n+1}) X_{n+1} X_{n+1}^T \tag{18}$$

$$\widehat{\theta}_{n+1} = \widehat{\theta}_n + b(\widehat{\beta}_n^T X_{n+1}, y_{n+1}) X_{n+1} \tag{19}$$

$$\widehat{\beta}_{n+1} = \widehat{\Psi}_{n+1}^{-1} \widehat{\theta}_{n+1} \tag{20}$$

Now, it is straightforward to introduce a forgetting factor into the recursions. The idea is to put less weight on more distant observations. For a forgetting factor $0 < \lambda \leq 1$, equations (11) and (12) can be revised as:

$$\widehat{\Psi}_{n+1} = \lambda \widehat{\Psi}_n + a(\widehat{\beta}_n^T X_{n+1}) X_{n+1} X_{n+1}^T \tag{21}$$

$$\widehat{\theta}_{n+1} = \lambda \widehat{\theta}_n + b(\widehat{\beta}_n^T X_{n+1}, y_{n+1}) X_{n+1} \tag{22}$$

The weights this model puts on historical observations are discounted exponentially at rate λ. When λ equals 1, it is equivalent to a binary logistic regression model.

We estimate the adaptive Naive Bayes model following the procedure described in section 6.1.2. We update the labels of all the loan in each period because in each period some borrowers may fail to repay their monthly installments and make their non defaulted loan become defaulted.

Given the parameter estimates on the borrower, lender sides and risk assessment model, we can compute the expected funding probability, withdrawal probability, default probability and LGD for each loan application. Then we maximize the likelihood function (7) by choose the optimal δ.

E.2 Naive Bayes Classifier with A Forgetting Factor

On the risk assessment side, we incorporate a forgetting factor to introduce temporal adaptivity in the naive Bayes model in the following way: let $\lambda \in [0,1]$ be the user-defined forgetting factor. Let $\widetilde{\boldsymbol{\Omega}}_{\boldsymbol{t}} = \{\widetilde{\omega}_{1t}, \widetilde{\omega}_{2t}, ..., \widetilde{\omega}_{Kt}\}$ and $\widetilde{\boldsymbol{P}}_{kt} = (\widetilde{P}_{k1t}, \widetilde{P}_{k2t}, ..., \widetilde{P}_{kmt})$ denote the corresponding prior probability and feature distribution at time t, where $\widetilde{P}_{kjt} = (\widetilde{p}_{kj1t}, ..., \widetilde{p}_{kjn_jt})$ represent feature j's distribution in class k at the beginning of period t. Let t_i denote the period that loan application i comes in, $v_i = \lambda^{t-t_i}$, $V_t = \sum_{z=1}^{t} \sum_{i=1}^{I_z} v_i$, and I_z denotes all the loan applications in period z. Similar to equations (10) and (11), we have the following expressions:

$$\widetilde{\omega}_{kt} = \frac{1}{V_t} \sum_{z=1}^{t} \sum_{i=1}^{I_z} v_i \cdot \mathbf{1}[y_i = k], \tag{23}$$

$$\widetilde{p}_{kjht} = \frac{\sum_{z=1}^{t} \sum_{i=1}^{I_z} v_i \cdot X_{ijh} \cdot \mathbf{1}[y_i = k]}{\sum_{z=1}^{t} \sum_{i=1}^{I_z} v_i \cdot \mathbf{1}[y_i = k]}. \tag{24}$$

where $k = 1, 2, ..., K; j = 1, 2, ..., m; h = 1, 2, ..., n_j$

Similar to equation (6), we have

$$p(X_i|y_i = k) \propto \prod_{j=1}^{m} \prod_{h=1}^{n_j} \widetilde{p}_{kjht}^{X_{ijh}}, \tag{25}$$

According to Bayes' rule, we have:

$$\begin{aligned} P_i^k = p(y_i = k|X_i) &= \frac{p(y_i = k, X_i)}{p(X_i)} \\ &\propto p(X_i|y_i = k) \cdot p(y_i = k) \\ &= p(X_i|y_i = k) \cdot \widetilde{\omega}_{kt_i} \\ &\propto \prod_{j=1}^{m} \prod_{h=1}^{n_j} \widetilde{p}_{kjht}^{X_{ijh}} \cdot \widetilde{\omega}_{kt_i}, \end{aligned} \tag{26}$$

F Model Weights in E-RPM

In this section, we illustrate how we update each individual model and the model weights in E-RPM. As an example, we will show how to use E-RPM to train each individual model and determine weights to predict each loan application's funding probability. The cases for predicting withdrawal probability, default probability and LGD are similar.

(i) Initial individual models: We use the data from Feb 01, 2007 to Dec 19, 2010 (i.e., the pre-posted price period) as our initial data set. We first divide the data into four sub-samples, Q_0^1, ..., Q_0^4, using the first, second, and third quantiles of the data's recession probability. Here, the subscript 0 represents the initial period. Then we train four sub-sample specific logistic regression models, C_0^1,...,C_0^4. Notice that the specifications of the four individual logistics regression models are the same. But their parameters are different because they are trained using different sub-samples of the data set. We can think of C_0^1,...,C_0^4 as four experts who specialize in different economic environments.

(ii) Initial weights: Initialize the weights we put on each individual model to be $H = (0.25, 0.25, 0.25, 0.25)$.

(iii) Predictions: At time t, we denote the four individual models as C_t^1,...,C_t^4. Those four individual models are trained using data sets Q_t^1, ..., Q_t^4, respectively. Q_t^1, ..., Q_t^4 represent the four sub-samples divided by the first, second, and third quantiles of the data's recession probability in period t. We use C_t^1,...,C_t^4 to make predictions for loan applications arriving in period $t+1$. For loan application i with characteristics X_i, Z_{li} (rating specific characteristics) and macro environment index E_{t+1} at time $t+1$, we use each individual model to predict a corresponding funding probability. Let $F_{jit} = C_t^j(X_i, Z_{li}, E_t)$ represent individual model j's predicted funding probability for loan application i. We take the weighted average of F_{1it}, F_{2it}, F_{3it}, and F_{4it} as the predicted funding probability for loan application i. That is:

$$F_{it} = \sum_{j=1}^{4} h_{jt} F_{jit} = \sum_{j=1}^{4} h_{jt} C_t^j(X_i, Z_{li}, E_t), \tag{27}$$

where h_{jt} is the weight we put on individual model j.

(iv) Update weights: We now explain how we update the weight for each individual model for $t > 1$. Following Wang et al. (2003), the weight for each individual model in time t is a function of the inverse of its prediction mean squared error (MSE) in time $t-1$. Let n_{t-1} be the number of loan applications which borrowers decided to proceed (i.e., they decide not to withdraw after knowing its Prosper rating) in time $t-1$. The MSE for individual model j is given by,

$$B_{jt} = \frac{1}{n_{t-1}} \sum_{i=1}^{n_{t-1}} (y_i - F_{jit})^2, t = 2, 3, ..., T, \tag{28}$$

where y_i takes value 1 if loan application i is indeed funded and 0 otherwise; F_{jit} represents individual model j's predicted funding probability for loan application i. Wang et al. (2003) proposes setting weight for poorly performed model to zero. We experimented this approach, and find that if we only assign positive weights to the top 2 best models in each period (and setting the weights for the other two inferior models to zero), we get better prediction results. The weights for each individual model are determined as follows. Let

$$h'_{jt} = \begin{cases} 0 & \text{if individual model } j \text{ has the largest or second largest MSE in period t.} \\ [B_{jt}]^{-1} & \text{otherwise.} \end{cases}$$

Then we set the weight $h_{jt} = \frac{h'_{jt}}{\sum_j h'_{jt}}$.

(v) Update individual models: Let Q_t denote the whole data set we have at the end-of-period t. We divide the data into four sub-samples, Q_t^1, ..., Q_t^4, using the first, second, and third quantiles of Q_t's recession probability. We train individual model C_t^j using data Q_t^j.

(vi) Repeat steps (iii) to (v) until the last period of the data.

We apply the same procedures to estimate each loan application's withdrawal probability, default probability and LGD.

G Prediction Performance of Different Methods

This section presents the prediction performance comparison between different methods. All the results are summarized in Table 6.

G.1 Equal Weight Model

Figure A1 compares the prediction performances of EWM and E-RPM. E-RPM outperforms EWM in all four prediction tasks. The corresponding AUCs for funding, withdrawal and default predictions of EWM are 0.712, 0.534 and 0.597, respectively, while the corresponding AUCs of E-RPM are 0.854 and 0.617, respectively. Moreover, EWM gives a MSE of LGD prediction at 0.045, while the MSE of the E-RPM model is 0.035.

G.2 Moving Window Model

Figure A2 compares the prediction performances of MWM and E-RPM. E-RPM outperforms MWM in funding, default and LGD predictions. The corresponding AUCs for funding and withdrawal predictions of MWM are 0.847 and 0.609, respectively, and MWM gives a MSE of LGD prediction at 0.092.

G.3 Recession Probability Method

Figure A3 shows the ROC curves for RPM and E-RPM. E-RPM significantly outperforms RPM in predicting a loan application's funding, withdrawal and default outcomes. RPM's AUC for the three prediction tasks are 0.796, 0.587 and 0.603 respectively. The MSE of LGD prediction is 0.037 using the RPM model, which is larger than the MSE of the E-RPM model.

G.4 Gradual Forgetting Method

Figure A4 shows the ROC curves for GFM and E-RPM. E-RPM significantly outperforms GFM in predicting loan application's funding and withdrawal outcomes. They have similar prediction powers as for the default prediction. GFM's AUCs for the three prediction tasks are 0.638, 0.535, and 0.615, respectively, The MSE of LGD prediction is 0.095 using the GFM model.

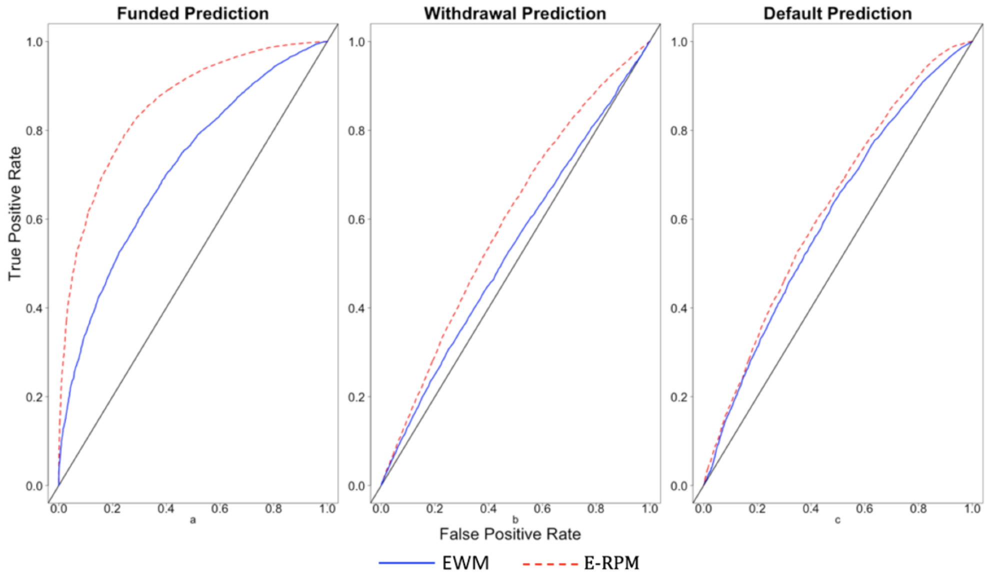
Table A1: Summary Stats of Recession Probability (%)

Mean	Median	SD	Max	Min
0.36	0.32	0.097	0.62	0.26

Notes: Summary statistics of the recession probability from 2011 to 2012.

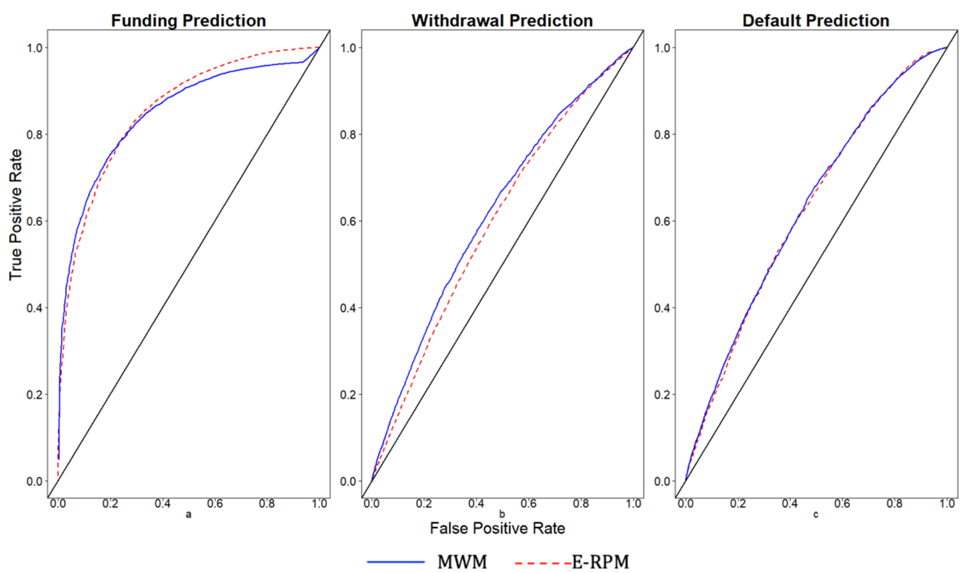
H Supplementary Figures

Figure A1: E-RPM vs. EWM



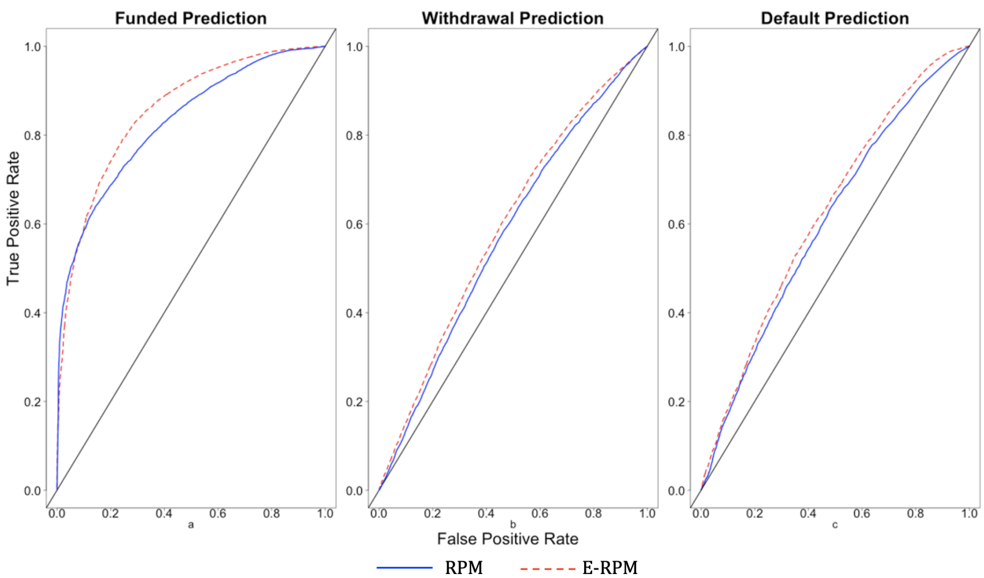
Notes: The larger the area under ROC is, the better prediction performance the model has. E-RPM outperforms EWM in predicting funding, withdrawal and default outcomes.

Figure A2: E-RPM vs. MWM



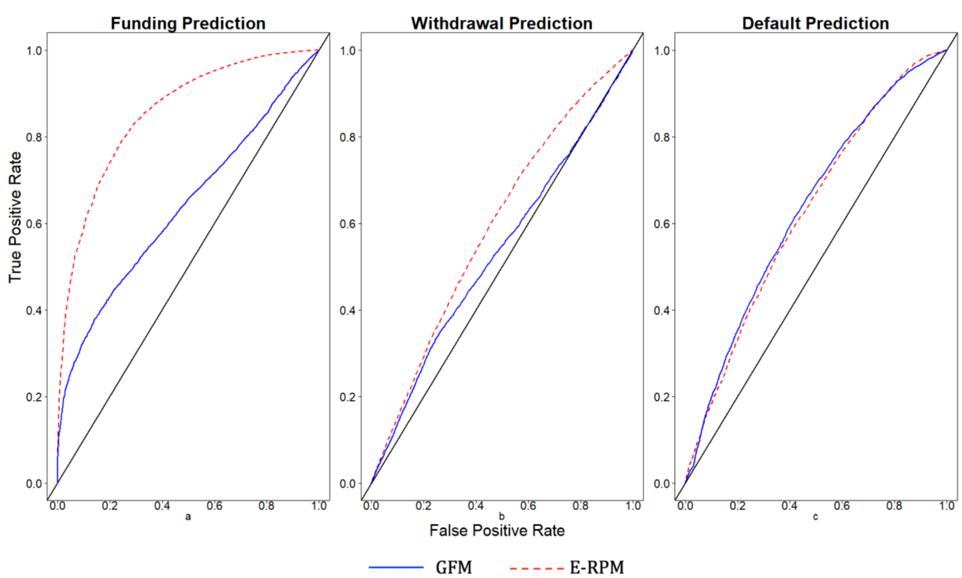
Notes: The larger the area under ROC is, the better prediction performance the model has. E-RPM outperforms MWM in predicting funding outcome. MWM outperforms E-RPM in withdrawal prediction. The two algorithms are very close when predicting default.

Figure A3: E-RPM vs. RPM



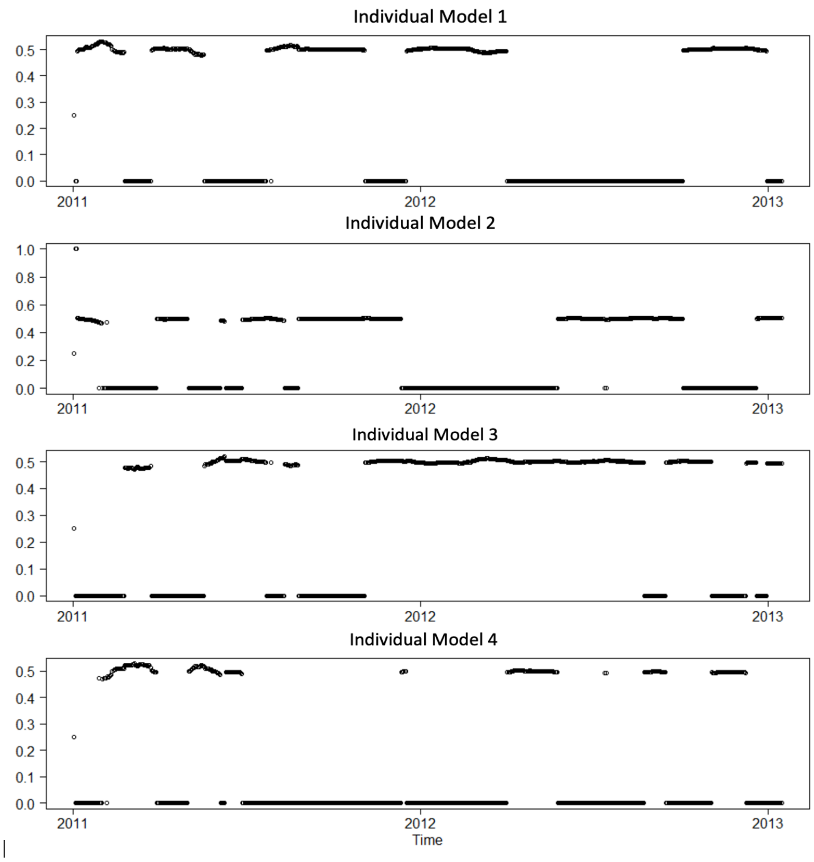
Notes: The larger the area under ROC is, the better prediction performance the model has. E-RPM outperforms RPM in predicting funding, withdrawal and default outcomes.

Figure A4: E-RPM vs. GFM



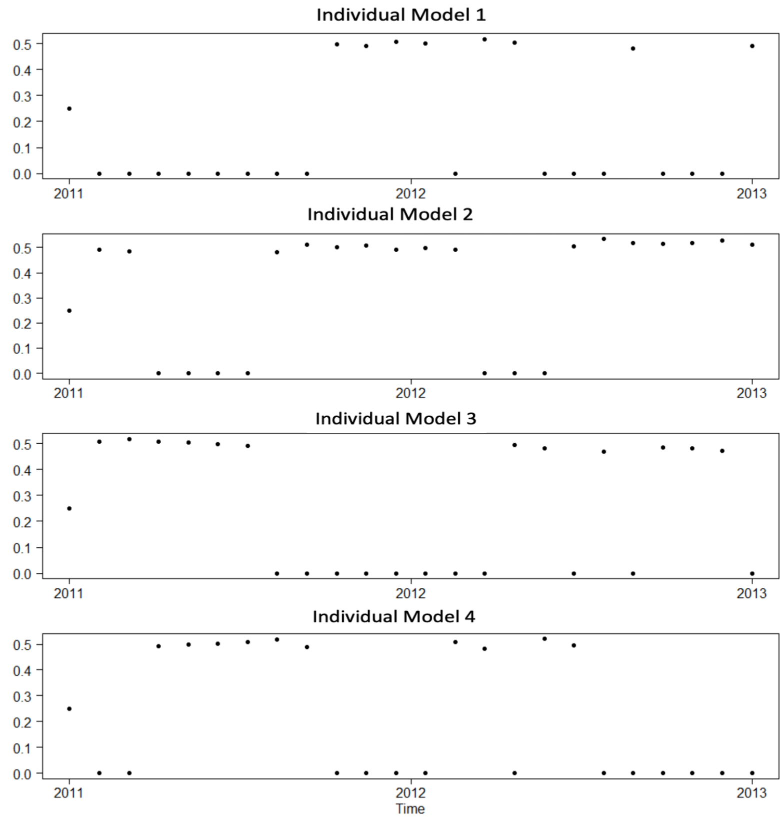
Notes: The larger the area under ROC is, the better prediction performance the model has. E-RPM outperforms GFM in predicting funding, and withdrawal outcomes. The two algorithms are very close when predicting default.

Figure A5: Weights on Individual Withdrawal Model in E-RPM



Notes: This figure shows the weights E-RPM puts on each individual model on withdrawal predictions. The weights are updated every day.

Figure A6: Weights on Individual Default Model in E-RPM



Notes: This figure shows the weights E-RPM puts on each individual model on default predictions. The weights are updated every 30 days.