

Research Article

Dynamic Credit Quality Evaluation with Social Network Data

Stanley Sewe ¹, Philip Ngare,^{1,2} and Patrick Weke^{1,2}

¹Pan African University Institute for Basic Sciences, Technology and Innovation, Kenya

²School of Mathematics, University of Nairobi, Kenya

Correspondence should be addressed to Stanley Sewe; stanleysewe@gmail.com

Received 2 November 2018; Accepted 3 February 2019; Published 1 April 2019

Academic Editor: Junzo Watada

Copyright © 2019 Stanley Sewe et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We investigate the filtering problem where the borrower's time varying credit quality process is estimated using continuous time observation process and her (in this paper we refer to the borrower as female and the lender as male) ego-network data. The hidden credit quality is modeled as a hidden Gaussian mean-reverting process whilst the social network is modeled as a continuous time latent space network model. At discrete times, the network data provides unbiased estimates of the current credit state of the borrower and her ego-network. Combining the continuous time observed behavioral data and network information, we provide filter equations for the hidden credit quality and show how the network information reduces information asymmetry between the borrower and the lender. Further, we consider the case when the network information arrival times are random and solve stochastic optimal control problem for a lender having linear quadratic utility function.

1. Introduction

In this study we consider the problem of stochastic filtering in the presence of network generated information. Consider a continuous time credit quality process X_t which is hidden and is only partially observed through its randomized function Y_t . Additionally, at discrete time points the observer has access to unbiased signals of the process X_t and of nodes directly linked to the X_t node. Thus the standard stochastic filtering is carried out before the arrival of the network information. The unbiased signals from the network are used to improve the estimates of the hidden true credit quality.

We assume that the nodes are individual (potential) borrowers in a dynamic social network with the process X_t being borrower's true credit quality modeled as an Ornstein-Uhlenbeck process. The hidden process X_t is partially observed through its function Y_t in continuous time. Thus Y_t models the typical borrower's behavior and financial information available to the lender (e.g., account turnover, periodic account balance). X_t and Y_t are modeled as scalar processes for simplicity, though the results apply to the vector case too. The borrower knows her true credit quality, and this is also visible to her direct social contacts. A social link is formed upon mutual consent between the parties. The probability of a link forming is influenced by the distance

between the borrowers' credit quality. Thus links are formed by credit type homophily. Homophily ([1]) is the idea that individuals with similar characteristics are more likely to have a network tie than individuals with different characteristics. The network is thus modeled as an undirected continuous time latent space model. The lender is able to observe at discrete time points the borrower's ego-network. Hence at time points $0 = t_0 < t_1 < \dots < t_{N-1} < T$ the lender is able to update his estimate of the borrower's credit quality using the network data.

Kalman-Bucy filtering techniques provide the estimates for the hidden process X_t and the conditional variance based on the continuous time observations Y_t . At time t_k , $k = 0, 1, N \in \mathbb{N}$ the estimates and the conditional variances are updated using the vector \mathbf{Z}_k of network signals. The dimension of the vector is depended on the borrower's degree (number of friends) $\eta_{t,k}$ at time t_k . Thus we combine Kalman-Bucy filtering and Bayesian updating techniques to improve the lender's estimate of the borrower's credit quality. Further, we prove that the network data leads to a lower variance for the estimated credit quality \hat{X}_t . In addition, we consider the case whereby the network information arrives at random times with intensity of a Poisson process. The dynamics of the filtered estimates and conditional variances are the same as in the case of discrete time information arrivals, though we

have the conditional variance being a piecewise deterministic stochastic process. We solve a stochastic optimization problem for a lender using the linear quadratic utility. Due to the random times t_k , the problem is reduced to a linear quadratic Gaussian utility with jumps. Our credit model can be seen as continuous time version of the static credit scoring model of [2]. Also related is the work of [3] who derived filtered estimates and conditional variance of a hidden drift process in the presence of continuous time observation and discrete time scalar expert opinion. Our stochastic optimal control problem is close to the optimal credit limit problem of [4] whereby the lender optimizes the credit limit available to a corporate entity based on the status of the entity's surplus process. Reference [5] solved a singular optimization problem within a random time horizon to obtain the optimal interest rate payable by a corporate entity. Reference [6] provides a good account of stochastic optimization for jump diffusion processes whilst [7] is an excellent treatment of discrete and stochastic optimization.

Application of social network analysis and social network data in consumer credit has become ubiquitous; see [8] for a review. With the availability of social network data, digital credit and mobile phone based lending have emerged as key areas of study in finance. Some previous studies of dynamic social network include the coevolution model of [9] where the node's attributes and network ties influence each other. The nodal attributes and network ties are modeled as continuous time Markov chains. The static latent distance models of [10] and discrete time model of [11] are able to capture key network properties like homophily and transitivity (two nodes connected to a common third node are likely to be connected) and group structures. Reference [12] provides a Bayesian dynamic model of relational structure on a latent variable continuous time social network. Reference [13] gives a good review of the static and dynamic latent distance and latent space network models.

There exist two approaches in modelling credit risk: the structural and reduced form approach, respectively. In the structural approach, default occurs when the observed process related to the value of the firm (or firm cash flow) passes below a deterministic default threshold. The threshold is chosen by the managers of the firm, with the default time being a predictable stopping time. In the reduced form approach, default is modeled as an inaccessible stopping time, with the default threshold modelled as a random variable. Reference [14] provides a good introduction to credit risk modelling. There have been attempts to apply the corporate credit risk models in modelling unsecured consumer credit. The challenge therein lies in the fact that it is hard to measure a borrower's assets. In addition, default in consumer loans is related more to cash flow issues than to the borrower's debts exceeding her assets. Reference [15] modeled the consumer credit risk using the structural approach of credit risk modeling. Treating the consumer's behavioral score as a continuous time jump diffusion process, the authors were able to apply the option pricing theory to model the probability of default. Our model is related to the structural approach of credit risk modeling in that the credit worthiness score is seen as an asset to the borrower. However, in the model the lender manages

credit risk through continuous updating of the credit limit and borrower's credit worthiness. The credit limit is affected by the borrower's credit worthiness, which is related to the probability of default. In turn, the sanctioned credit limit (credit exposure) affects the borrower's credit worthiness.

Some existing studies on consumer credit scoring include [16] whereby the borrower's credit rating is modelled as a discrete time Markov chain process upon incorporating a latent variable driven by economic conditions. Reference [2] used social network links and signals from the borrower's ego-network to estimate her credit quality. The model assumed that the credit quality which is normally distributed is hidden and can only be estimated using the noisy signals from the borrower and her ego-network. Thus the model was based on homophily by preference. Reference [17] proposed the SEN-HMM-CSD (Social Economic Network-Hidden Markov Model-Credit Scores and Default) model whereby economic and social network variables including trust, distrust, and reputation were used to estimate the hidden credit quality of a borrower in a social network. In the SEN-HMM-CSD model, the true credit quality was modeled as a discrete time, discrete state hidden Markov model. Assuming a fully connected social network on a population with no drift, the network based variables were used to generate the credit risk analysis factors, which in turn were used to estimate the hidden Markov parameters for each individual borrower. Our work generalizes the SEN-HMM-CSD model in the continuous time continuous state Hidden Markov model direction through the following: inclusion of neighbouring nodes signals in the estimation of credit worthiness, credit limit variability tied to individual's credit worthiness, and the network ties based on credit type homophily; thus there is no assumption on a fully connected network model.

The objective of stochastic filtering is to find the best estimate of a hidden process X partially seen through the observation process Y . Excellent textbook treatment of stochastic filtering includes [18–20]. Recent studies on stochastic filtering and information modeling include [3] who studied a market model whereby the unobserved drift process (modeled as an Ornstein-Uhlenbeck process) is filtered conditioned on a continuous time stock return observations and discrete time expert opinion. By incorporating expert opinion in their estimation, the model was a continuous time version of the Black-Litterman model [21]. Reference [22] considered a market model whereby the unobserved drift parameter is estimated on an observation filtration initially enlarged with the terminal value of the stock price perturbed by some constant variance noise. Reference [23] considered an enlargement of filtration problem with a random variable combined with partial information, with an application in linear stochastic control.

The paper is organized as follows. In Section 2, we present our credit model, dynamic social network model, and the information setup. Stochastic filtering and Bayesian updating results are presented in Section 3. Within this section, we also present the properties of the conditional variance. Credit limit management optimization problem is solved in Section 4. Brief numerical results are presented in Section 5, whilst Section 6 concludes.

2. The Credit Model

Consider a filtered probability space $(\Omega, \mathcal{A}, \mathbb{A} = (\mathcal{A}_t)_{t \geq 0}, \mathbb{P})$ satisfying the usual conditions of right continuity and completeness. All processes are assumed to be \mathbb{A} adapted. We are interested in the following filtering problem.

Borrower's Behavioral Dynamics. We model the borrower's observed behavioral dynamics as a linear diffusion process defined as

$$dY_t = \alpha X_t dt + \sigma dW_t \quad (1)$$

The parameters $\alpha, \sigma > 0$ are assumed to be constants and W_t is a \mathbb{A} adapted one dimensional Brownian motion. The hidden credit quality process X_t driving the drift is modeled as a mean-reverting Ornstein Uhlenbeck process defined as

$$\begin{aligned} dX_t &= \mu(\delta - X_t)dt + \gamma dB_t \\ X_0 &= x \end{aligned} \quad (2)$$

where $\delta \in \mathbb{R}$, $\mu > 0$, $\gamma > 0$ are constants and B_t is a Brownian motion. We assume that B_t and W_t are independent. x is a \mathbb{A} measurable normally distributed random variable independent of B and W with mean $m_0 \in \mathbb{R}$ and variance $v_0 > 0$. X_t is a Gaussian process; its mean and variance are given by

$$\begin{aligned} &\delta + e^{-\mu t} (m_0 - \delta) \\ &\frac{\gamma^2}{2\mu} + e^{-2\mu t} \left(v_0 - \frac{\gamma^2}{2\mu} \right) \end{aligned} \quad (3)$$

Network Dynamics. We consider a society with a population \mathcal{X} . Each individual in the population has a normally distributed time varying credit quality X_{it} . The credit qualities X_{it} are assumed to be independent across the individuals. At each time $t > 0$, with no other additional borrower information, an individual's true credit quality X_t is assumed to be normally distributed with mean 0 and variance q^{-1} . Individuals i and j interact and a network tie is formed by mutual consent. Affecting equations (4) and (5)-the highlighted section should read: Assuming an undirected network i.e. $\mathcal{Y}_{ij}(t) = \mathcal{Y}_{ji}(t)$, then we let

$$\mathcal{Y}_{ij}(t) \mid \pi_{ij}(t) \sim \text{Bern}(\pi_{ij}(t)) \quad (4)$$

for every $i \neq j$ and $t \geq 0$ with

$$\pi_{ij}(t) = e^{-(1/2)(X_{it} - X_{jt})^2} \quad (5)$$

Thus the network ties $\mathcal{Y}_{ij}(t) \in \{0, 1\}$ are conditionally independent Bernoulli random variables given the corresponding probabilities of tie formation $\pi_{ij}(t) \in (0, 1)$. $\pi_{ij}(t)$ is the survival function of a standard Raleigh distribution. Existence of a network tie between individual i and j is dependent on the distance between their credit types, and shorter distance leads to higher probability of network tie. Thus the network model captures homophily, whereby individuals with similar characteristics are more likely to have social network ties than individuals with different characteristics. We assume there is no cost incurred in network tie formation or destruction.

Let G_t denote the set of friendship ties in the society at time t . The set of borrower i 's friends at time t , known as her ego-network, is defined as $g_t^i = \{ij \mid ij \in G_t\}$. For a particular borrower i , we consider her true credit quality process X_t and behavioral score Y_t .

Lender's information. As a result of the information asymmetry between the lender and the borrowers, the lender is unable to directly observe the true credit quality process X_t for the individual i . He however gets noisy observation of the hidden process via the behavioural process Y_t . Additionally, the lender is able to receive at discrete time points $0 = t_0 < t_1 < \dots < t_{N-1} < T$, $N \in \mathbb{N}$ noisy signals about the current state of X_t and neighbouring nodes within her ego-network. For each $k \in 0, 1, \dots, N$, let the set $\mathbf{Z}_k = \{Z_{jk} \mid ij \in g_t^i\}$ be the set of signals of the borrower's ego-network such that $Z_{jk} = X_{jt_k} + \sqrt{\Lambda_k} \varepsilon_{jk}$. The variable $\varepsilon_{jk} \sim N(0, 1)$ is i.i.d across individuals and $\mathbb{E}(\varepsilon_{js} \varepsilon_{lu}) = 0$ for $j \neq l$ $s \neq u$. Thus at time point t_k the lender observes borrower i 's ego-network and the vector \mathbf{Z}_k comprising of the noisy observation of her current credit state $Z_k = X_{t_k} + \varepsilon_{ik}$ and the credit states of her immediate social contacts $\{Z_{kj}, ij \in g_{t_k}^i\}$. The variance Λ_k is a measure of the reliability of the network information received by the lender.

The information available to the lender can thus be represented by the following filtrations:

$$\begin{aligned} \mathbb{F}^Y &= (\mathcal{F}_t^Y) \\ &t \in [0, T] \text{ with } \mathcal{F}_t^Y \text{ generated by } \{Y_s, s \leq t\} \\ \mathbb{F}^Z &= (\mathcal{F}_t^Z) \\ &t \in [0, T] \text{ with } \mathcal{F}_t^Z \text{ generated by } \{Y_s, s \leq t, Z_k, t_k \leq t\} \\ \mathbb{F}^O &= (\mathcal{F}_t^O) \\ &t \in [0, T] \text{ with } \mathcal{F}_t^O \text{ generated by } \{Z_k, t_k \leq t\} \end{aligned} \quad (6)$$

\mathbb{F}^Y corresponds to the continuous time behavioral score only and \mathbb{F}^O consists of the network information received at discrete times whilst \mathbb{F}^Z is the combination of behavioral score and the network information. We assume that the σ -algebras \mathcal{F}_t^Y and \mathcal{F}_t^Z are augmented by the \mathbb{P} null sets.

Assumption 1. It is obvious that $\mathcal{F}_t^Y \subset \mathcal{F}_t^Z$ for all $t > 0$. \mathcal{F}_t^Y is generated by the innovation process $dM_t = dW_t + \alpha \sigma^{-1}(X_t - \widehat{X}_t)dt$. We assume that $\mathbb{E}(\mathbf{Z}_k, M_t - M_{t,k}) = 0$ for all $t \geq t_k$. Thus \mathcal{F}_t^Y is immersed in \mathcal{F}_t^Z ; hence every \mathbb{F}^Y square integrable martingale is a \mathbb{F}^Z square integrable martingale.

3. Stochastic Filtering

Stochastic filtering entails obtaining estimate of the hidden process X_t conditioned on the observed process. Let \widehat{X}_t^H be the projection of the process X_t onto the observed filtration \mathbb{F}^H , $H \in \{O, Y, Z\}$, i.e., $\widehat{X}_t^H = \mathbb{E}(X_t \mid \mathcal{F}_t^H)$. \widehat{X}_t^H is the optimal estimator of X_t in the mean square sense. In this section,

we derive the recursive equations for the filtered estimates conditioned on the different information settings.

3.1. Behavioural Observations. In the case whereby the lender observes the borrower's behavioral information only, i.e., where the lender's observation filtration is \mathcal{F}_t^Y , the state and observation processes constitute a linear system of equations. Besides, the bivariate process (X_t, Y_t) is Gaussian. Thus the usual Kalman-Bucy filtering technique, see, e.g., [19, 20], can be used to obtain the conditional mean $\widehat{X}_t^Y = \mathbb{E}(X_t | \mathcal{F}_t^Y)$ and the conditional variance $\lambda_t^Y = \mathbb{E}[(X_t - \widehat{X}_t^Y)^2 | \mathcal{F}_t^Y]$.

The dynamics of \widehat{X}_t^Y is given by the following SDE:

$$\begin{aligned} d\widehat{X}_t^Y &= (\mu(\delta - \widehat{X}_t^Y) - \alpha^2 \sigma^{-2} \lambda_t^Y \widehat{X}_t^Y) dt + \alpha \sigma^{-2} \lambda_t^Y dY_t, \\ \widehat{X}_0^Y &= m_0 \end{aligned} \quad (7)$$

whilst the dynamics of λ_t^Y is given by the deterministic ODE

$$\begin{aligned} \frac{d\lambda_t^Y}{dt} &= -2\mu\lambda_t^Y + \gamma^2 - \alpha^2 \sigma^{-2} (\lambda_t^Y)^2, \\ \lambda_0^Y &= v_0 \end{aligned} \quad (8)$$

Equation (8) is the wellknown Riccati equation. With initial value $\lambda_0^Y = v_0$, the unique solution for the equation can be given as

$$\lambda_t^Y = \frac{-\mu\sigma^2}{\alpha^2} + C_0 \frac{C_1 + C_2 e^{-2(\alpha^2/\sigma^2)C_0 t}}{C_1 - C_2 e^{-2(\alpha^2/\sigma^2)C_0 t}} \quad (9)$$

with $C_0 = (\sigma/\alpha)\sqrt{(\mu^2\sigma^2/\alpha^2) + \gamma^2}$, $C_1 = v_0 + C_0 + \mu\sigma^2/\alpha^2$, and $C_2 = v_0 - C_0 + \mu\sigma^2/\alpha^2$ (see, e.g., [24]).

3.2. Behavioral Observations and Network Information. This is the case whereby the lender uses both the observed behavioral score and the network information to obtain an estimate of the borrower's credit quality. The following lemma shows that the expected number of friends for individual i at any time $t \geq 0$ can be treated as a constant.

Lemma 2. Let η_t be the borrower's degree (number of friends) at time t . At each time $t \geq 0$, conditioned on the borrower's true credit quality X_t , the expected degree $\mathbb{E}(\eta_t | X_t)$ is a constant denoted as η .

Proof. Conditioned on the borrower's true credit type X_t , the probability of having a network tie with any other individual is

$$\int_{-\infty}^{\infty} e^{-(X_t-s)^2/2} \sqrt{\frac{q}{2\pi}} e^{-q(s^2/2)} ds = \sqrt{\frac{q}{q+1}} e^{-(q/2(q+1))X_t^2} \quad (10)$$

Thus the conditional expected number of friends $\mathbb{E}(\eta_t | X_t)$ is given by

$$\mathbb{E}(\eta_t | X_t) = \mathcal{Z} \sqrt{\frac{q}{q+1}} e^{-(q/2(q+1))X_t^2} \quad (11)$$

As $q \rightarrow 0$, $\mathcal{Z} \rightarrow \infty$ we make the simplifying assumption that $\mathcal{Z} \sqrt{q/(q+1)}$ is a constant denoted by η . Thus in the limit $\mathbb{E}(\eta_t | X_t) = \eta$. \square

The simplifying assumption is justified by the consideration that in a typical human social network; types are diffuse and the population size is large. Besides, in a small social network with less diffuse individual types, the benefits of social network scoring would be limited.

Proposition 3. For any $k \in 0, 1, \dots, N$ and $t > 0$, let $p_k = \Lambda_k^{-1}$ be the precision of the network information at time t_k . Further define the parameter

$$\theta_k^Z = \frac{(p_k + 1)}{\lambda_{t_k-}^Z p_k (p_k + 1 + \eta_k) + (p_k + 1)} \quad (12)$$

Then it holds that

(i) For any $t \in [t_k, t_{k+1})$, \widehat{X}_t^Z is Gaussian and its dynamics satisfies the following:

$$d\widehat{X}_t^Z = (\mu(\delta - \widehat{X}_t^Z) - \alpha^2 \sigma^{-2} \lambda_t^Z \widehat{X}_t^Z) dt + \alpha \sigma^{-2} \lambda_t^Z dY_t \quad (13)$$

whilst the conditional variance is given by

$$\lambda_t^Z = \frac{-\mu\sigma^2}{\alpha^2} + C_0 \frac{C_{1k} + C_{2k} e^{-2(\alpha^2/\sigma^2)C_0 t}}{C_{1k} - C_{2k} e^{-2(\alpha^2/\sigma^2)C_0 t}} \quad (14)$$

with initial values $\widehat{X}_{t_k}^Z, \lambda_{t_k}^Z$. C_0 is same as in (9) whilst $C_{1k} = \lambda_{t_k}^Z + C_0 + \mu\sigma^2/\alpha^2$ and $C_{2k} = \lambda_{t_k}^Z - C_0 + \mu\sigma^2/\alpha^2$

(ii) At information date t_k , $\widehat{X}_{t,k}^Z$ is Gaussian with the conditional mean and variance $\widehat{X}_{t,k}^Z, \lambda_{t,k}^Z$ updated from their respective values at t_{k-} (that is, before the arrival of the network information) to be mean

$$\widehat{X}_{t,k}^Z = \theta_k^Z \lambda_{t_{k-}}^Z \left(\frac{\widehat{X}_{t,k-}^Z}{\lambda_{t_{k-}}^Z} + p_k Z_{ik} + \frac{p_k}{p_k + 1} \sum_{j \in \mathcal{Z}} Z_{jk} \right) \quad (15)$$

and variance

$$\theta_k^Z \lambda_{t_{k-}}^Z = \frac{\lambda_{t_{k-}}^Z (p_k + 1)}{\lambda_{t_{k-}}^Z p_k (p_k + 1 + \eta_k) + (p_k + 1)} \quad (16)$$

Proof.

Part (i). Between two information dates, $t \in [t_k, t_{k+1})$, there is no arrival of new network information. The only observed process is the continuous time returns Y_t with the lender's information set being $\mathcal{F}_t^Z = \mathcal{F}_{t,k}^Z \vee \sigma\{Y_s, t_k < s \leq t\}$. Thus, we have the standard Kalman-Bucy filtering case, with the initial values for conditional mean and conditional variance being $\widehat{X}_{t_k}^Z$ and $\lambda_{t_k}^Z$, respectively. The result thus follows from (8) and (9).

Part (ii). At the information date, t_k , the estimates for the mean and variance are updated from their values at time t_{k-} using the Bayesian method. The borrower's ego network at time t_k comprises all her direct network ties. We assume that

the lender is able to observe the borrower's complete ego-network at time t_k . The posterior probability of the borrower's credit type $\mathbb{P}(X_{t_k} | \mathbf{Z}_k)$ is obtained by

$$\begin{aligned} \mathbb{P}(X_{t_k} | \mathbf{Z}_k) &\propto \mathbb{P}(X_{t_k}, \mathbf{Z}_k) = \int_{-\infty}^{\infty} \mathbb{P}(\mathbf{Z}_k | X_{t_k}, X_{j,t_k}) \\ &\cdot \mathbb{P}(X_{t_k}) \mathbb{P}(X_{j,t_k} | X_{t_k}) dX_{j,t_k} \\ &= \int_{-\infty}^{\infty} \mathbb{P}(\mathbf{Z}_k | X_{t_k}, X_{j,t_k}) \mathbb{P}(X_{t_k}) \\ &\cdot \mathbb{P}(X_{j,t_k}) dX_{j,t_k} \end{aligned} \quad (17)$$

The last equality is as a result of the assumption of independence for the X_{j,t_k} . We have $\mathbb{P}(X_{j,t_k}) = \prod_{ij \in g_k^1} e^{-(q/2)X_{j,t_k}^2}$ being the assumed density of any individual X_j for $j \neq i$. The lender does not update his knowledge of individual j 's signal. The remaining part of the integrand is

$$\begin{aligned} &\mathbb{P}(\mathbf{Z}_k | X_{t_k}, X_{j,t_k}) \mathbb{P}(X_{t_k}) \\ &= e^{-((X_{t_k} - X_{t_k}^Z)^2 / 2\lambda_{t_k}^Z)} \times e^{-(p_k/2)(Z_{ik} - X_{t_k})^2} \\ &\times \underbrace{\prod_{j \in Z_k} e^{-(p_k/2)(Z_{jk} - X_{j,t_k})^2} \times \prod_{ij \in g_k^1} e^{-(1/2)(X_{t_k} - X_{j,t_k})^2}}_{(a)} \\ &\times \underbrace{\prod_{j \in g_k^0, j \notin Z_k} \left(1 - \frac{\mathbb{E}(\eta_{t_k} | X_{t_k})}{\mathcal{Z}}\right)}_{(b)} \end{aligned} \quad (18)$$

where

- (i) The first term denotes the product of the prior density (before the arrival of network information) and the likelihood function for the observation $Z_{it,k}$.
- (ii) (a) denotes the probability that at time t_k borrower i is friends with the individuals within her ego-network (in g_k^1) whose signals are in \mathbf{Z}_k and that these friends have the signals as collected in \mathbf{Z}_k
- (iii) (b) denotes the probability that at time t_k borrower i is not friends with anyone outside g_k^1 .

As $q \rightarrow 0$, $\mathcal{Z} \rightarrow \infty$, by the monotone convergence theorem and applying Lemma 2 above then

$$\prod_{j \in g_k^0, j \notin Z_k} \left(1 - \frac{\mathbb{E}(\eta_{t_k} | X_{t_k})}{\mathcal{Z}}\right) \rightarrow e^{-\eta} \quad (19)$$

which has no X_{t_k} term. Hence we have

$$\begin{aligned} \mathbb{P}(X_{t_k} | \mathbf{Z}_k, g_k) &\propto \int_{-\infty}^{\infty} e^{-(X_{t_k} - X_{t_k}^Z)^2 / 2\lambda_{t_k}^Z} \\ &\times e^{-(p_k/2)(Z_{ik} - X_{t_k})^2} \end{aligned}$$

$$\begin{aligned} &\times \prod_{j \in Z_k} e^{-(p_k/2)(Z_{jk} - X_{j,t_k})^2} \\ &\times \prod_{ij \in g_k^1} e^{-(1/2)(X_{t_k} - X_{j,t_k})^2} dX_{j,t_k} \end{aligned} \quad (20)$$

whereby the integrand is a product of Gaussian densities. Upon integrating out X_{j,t_k} and matching the terms of X_{t_k} and $X_{t_k}^2$ we obtain the posterior density which is Gaussian with the given expectation and variance. \square

For the case whereby the lender observes only the network information, i.e., $H = O$, we have that between information dates $t \in [t_k, t_{k+1})$, there is no arrival of new information. We thus get the following corollary.

Corollary 4. When the lender's information set is \mathcal{F}_t^O we have the following:

- (i) For $t \in [t_k, t_{k+1})$, the estimates for the mean and variance are given by

$$\widehat{X}_t^O = \delta + e^{-\mu(t-t_k)} (\widehat{X}_{t_k}^O - \delta) \quad (21)$$

$$\lambda_t^O = e^{-2\mu(t-t_k)} \lambda_{t_k}^O + \frac{\gamma^2}{2\mu} (1 - e^{-2\mu(t-t_k)}) \quad (22)$$

- (ii) At information date t_k , it holds that $X_{t_k}^O$ is Gaussian with mean and variance

$$\widehat{X}_{t_k}^O = \theta_k^0 \lambda_{t_k-}^0 \left(\frac{\widehat{X}_{t_k-}^O}{\lambda_{t_k-}^0} + p_k Z_{ik} + \frac{p_k}{p_k + 1} \sum_{j \in Z} Z_{jk} \right) \quad (23)$$

$$\theta_k^0 \lambda_{t_k-}^0 = \frac{\lambda_{t_k-}^0 (p_k + 1)}{\lambda_{t_k-}^0 p_k (p_k + 1 + \eta_k) + (p_k + 1)} \quad (24)$$

respectively.

Proof. Between information dates, it follows that $\mathcal{F}_t^O = \mathcal{F}_{t_k}^O$ for $t \in [t_k, t_{k+1})$. Thus $\widehat{X}_t^O = \mathbb{E}(X_t | \mathcal{F}_{t_k}^O)$ and $\lambda_t^O = \mathbb{E}[(X_t - \widehat{X}_t^O)^2 | \mathcal{F}_{t_k}^O]$. From the state/observation model, we get that X_t has the form

$$X_t = \delta + e^{-\mu(t-t_k)} \left(X_{t_k} + \gamma \int_{t_k}^t e^{\mu s} dB_s \right) \quad (25)$$

Therefore

$$\begin{aligned} \widehat{X}_t^O &= \delta + e^{-\mu(t-t_k)} \mathbb{E}(X_{t_k} - \delta | \mathcal{F}_{t_k}^O) \\ &= \delta + e^{-\mu(t-t_k)} (\widehat{X}_{t_k}^O - \delta) \end{aligned} \quad (26)$$

The conditional variance is given by

$$\begin{aligned} \mathbb{E}((X_t - \widehat{X}_t^O)^2 | \mathcal{F}_{t_k}^O) &= \mathbb{E}\left(\left(e^{-\mu(t-t_k)} (X_{t_k} - \widehat{X}_{t_k}^O) \right. \right. \\ &\quad \left. \left. + e^{-\mu(t-t_k)} \gamma \int_{t_k}^t e^{\mu(s-t_k)} dB_s \right)^2 | \mathcal{F}_{t_k}^O\right) \end{aligned}$$

$$\begin{aligned}
&= e^{-2\mu(t-t_k)} \mathbb{E} \left((X_{t_k} - \widehat{X}_{t_k}^O)^2 \mid \mathcal{F}_{t_k}^O \right) \\
&+ e^{-2\mu(t-t_k)} \gamma^2 \int_{t_k}^t e^{2\mu(s-t_k)} ds = e^{-2\mu(t-t_k)} \lambda_{t_k}^O + \frac{\gamma^2}{2\mu} (1 \\
&- e^{-2\mu(t-t_k)})
\end{aligned} \tag{27}$$

Here we have employed the martingale property of the Brownian motion process and used the Ito isometry to obtain the variance. At information date t_k , the estimates are updated by the information received from the random vector \mathbf{Z}_k . With a Gaussian prior for X_{t_k} , we use the Bayesian update similar to Proposition 3 part (ii) to get the posterior expectation and variance. \square

Remark 5.

- (i) For the case where the individual has no network ties (isolated node), i.e., when $\eta_{tk} = 0$, then θ_k^0 reduces to $\Lambda_k/(\Lambda_k + \lambda_{t_k}^0)$. Considering the behavior of the estimates with limiting values of Λ_k , when $\Lambda_k = 0$ we have that $\theta_k^0 = 0$ and $\widehat{X}_t^O = X_{t_k}$ whilst when $\Lambda_k = \infty$ we have $\theta_k^0 = 1$ and $\widehat{X}_t^O = X_{t_k}$. For $\Lambda_k = \infty$ the implication is that no additional information on the hidden process X_t can be obtained by observing the network data.
- (ii) Further, the estimates \widehat{X}_t^O and λ_t^O can be obtained from \widehat{X}_t^Z and λ_t^Z as limiting cases as $\sigma \rightarrow \infty$. Thus between information dates, the dynamics for the conditional mean and conditional variance are $d\widehat{X}_t^O = (\mu(\delta - \widehat{X}_t^O))dt$ and $d\lambda_t^O = (-2\mu\lambda_t^O + \gamma^2)dt$, respectively. These are deterministic O.D.E equations which can be solved to yield the expressions in Corollary 4. The implication is that on account of $\sigma \rightarrow \infty$ no additional information relating to the hidden process X_t can be obtained by observing the continuous time process Y_t .
- (iii) We may consider a case whereby the information arrival times $t_k, k = 0, 1, \dots, N-1$ are random times with intensity φ akin to the jump times of a Poisson process. The random times are independent of the processes X, Y, Z and the Brownian motions W, B ; thus the timing of jumps does not carry any information. We deal with this case in Section 4 on optimal credit limit management.

In the following proposition, we show how the additional network information \mathbf{Z}_k improves the lender's estimate of the borrower's true credit quality. This is captured by the lower variance associated with the estimates from the information set \mathcal{F}_t^Z compared to the other two alternative information settings. Thus a lender incorporating signals from the borrower's ego-network in his analysis is bound to have a better estimate of the borrower's true credit history.

Proposition 6 (properties of the conditional variance). *For $H \in \{Y, O\}$, $t \in [0, T]$ and $k = 1, 2, \dots, N$*

$$\lambda_t^Z \leq \lambda_t^H \tag{28}$$

Proof. For $t \in [t_k, t_{k+1})$, the dynamics of the conditional variance is given by the function

$$\frac{d\lambda_t^H}{dt} = f^H(\lambda_t^H) \tag{29}$$

for $H \in \{Y, O, Z\}$. From Proposition 3 (i), we have

$$f^Z(r) = f^Y(r) = -2\mu r + \gamma^2 - \alpha^2 \sigma^{-2} r^2 \tag{30}$$

whilst following Remark 5

$$\begin{aligned}
f^O(r) &= -2\mu r + \gamma^2 \text{ which is } f^O(r) \\
&= f^Z(r) + \alpha^2 \sigma^{-2} r^2
\end{aligned} \tag{31}$$

Comparing λ_0^Z and λ_0^Y , we note that, at time $t = 0$, $\lambda_0^Z = \theta_0^0 v_0 \leq v_0 = \lambda_0^Y$ since $\theta_0 \in [0, 1]$ for all values of η_0 . Since $f^Z(r) = f^Y(r)$ and a unique solution for the ODE $f^Z(r)$ exists, we note that the inequality will persist for all $t \in [t_0, t_1)$. At time $t = t_1$, we again have the inequality $\lambda_{t_1}^Z \leq \lambda_{t_1}^Y$. Extending the argument for all $t \in [0, T]$ yields that $\lambda_t^Z \leq \lambda_t^Y$.

To prove that $\lambda_0^Z \leq \lambda_0^O$, we note that, at time $t = 0$, $\lambda_0^Z = \lambda_0^O$. For $t \in [0, t_1)$, we have $f^Z(r) \leq f^O(r)$. This means that the two deterministic functions starting at the same initial value will have the relation $\lambda_t^Z \leq \lambda_t^O$. At time $t = t_1$, since the map $x \rightarrow x(p+1)/(xp_k(p_k+1+\eta_k)+(p_k+1))$ is nondecreasing, we shall have $\lambda_{t_1}^Z \leq \lambda_{t_1}^O$. Iterating this argument for all $t \in [0, T]$ we conclude that $\lambda_0^Z \leq \lambda_0^O$. \square

The implication of this result is that the network information improves the accuracy of the lender's belief about the borrower's true credit quality. Periodically observing the borrower's ego-network at discrete times without even observing the entire network leads to a better estimate of the hidden true credit quality. This is as a result of the network ties being based on borrower type homophily. This is quite in line with industry's practice in digital and mobile phone based lending whereby an individual's social network data can be used to improve the estimate of her credit quality.

4. Optimal Credit Limit Management

Following Remark 5 (iii), in this section, we consider the case whereby the network information arrival times t_k are modeled as Poisson jump times, with jump intensity φ . No information arrives at time t_0 . As such the sequence $(t_k, \mathbf{Z}_k)_{k \in I}$ is a marked point process, though the dimension of the vector \mathbf{Z}_k is governed by the borrower's degree η_{t_k} at time t_k . In this case, the dynamics of the filtered process \widehat{X}_t^Z and λ_t^Z will be same as in Proposition 3, save for the fact that λ_t^Z will be a piecewise deterministic process with stochastic jump times. Denote the jumps of $\widehat{X}_{t,k}^Z$ at time t_k as

$$R_{t,k} = \widehat{X}_{t,k}^Z - \widehat{X}_{t,k-}^Z \tag{32}$$

From the properties of the filtered process $\widehat{X}_{t,k}^Z$, $R_{t,k}$ has a Gaussian distribution. Let $f(\widehat{X}_{t,k}^Z, r_{t,k})$ denote the density of $R_{t,k}$ with first and second moments $\mu_{t,k}^1 = \mathbb{E}(R_{t,k})$ and $\mu_{t,k}^2 = \mathbb{E}(R_{t,k}^2)$, respectively. We note that the jump sizes are not i.i.d like in the compound Poisson case, since the density $f(\widehat{X}_{t,k}^Z, r_{t,k})$ is dependent on the state of the system at time t_k .

Consider a financial market model with terminal time T whereby a lender has access to the continuous time information \mathbb{F}^Y and the network information generated by the marked point process $(t_k, \mathbf{Z}_k)_{k \in I}$. The lender desires to obtain an estimate of the borrower true credit quality X_t based on the information set \mathbb{F}^Z . Besides, the lender desires to optimize his credit decision viz the credit limit l_t sanctioned to the borrower from time to time. Thus we are faced with a stochastic optimal control problem. Due to mathematical tractability, we assume a nondefaulting borrower who is credit constrained and thus will accept and fully utilize the credit limit as and when availed. Thus, the lender avails a credit line facility on revolving fund basis, and the borrower is expected to continuously utilize and repay any outstanding amounts. The lender seeks to adjust the sanctioned credit limit in line with the borrower's time varying credit quality. Different from the consumer credit card limits, the borrower is expected to fully repay any outstanding amounts before any new borrowing. Such a credit arrangement is applicable in unsecured mobile and digital banking products especially within the emerging market economies; see, e.g., [25, 26].

Define the controlled credit quality process as

$$dX_t = (\mu(\delta - X_t) + (\bar{\alpha} - p)l_t)dt + \gamma dB_t \quad (33)$$

When the borrower has no outstanding credit balance, $l_t = 0$. As soon as the borrower obtains credit within the sanctioned limit, the finances are channeled towards short-term working capital/consumption with return of $\bar{\alpha}$ (different from the drift before borrowing) which we net the price per unit of funds p . We assume that borrowing impacts the drift of the borrower's credit quality without affecting her volatility. Let $\phi = (\bar{\alpha} - p) > 0$ and

$$J(t, x) = \mathbb{E}_{t,x} \left(\int_t^T e^{-\omega s} \left(l_s(p - \kappa) - \frac{1}{2} \beta l_s^2 \right) ds \right) \quad (34)$$

be the lender's objective function, where ω is the lender's discounting factor, p is the price per unit of funds, κ is the lender's cost per unit of funds, and β is the marginal lending cost. ω, κ, p , and β are all positive. The lender's optimization problem is

$$\sup_{l \in U} J(t, x) \quad (35)$$

where U , the admissible set of lending strategies, is defined below. Since the controlled state process X is not observable, this is a situation of stochastic optimization with partially observed process. From Proposition 3 and Remark 5 and applying Ito's formula for jump diffusion process, we condition on the information set \mathbb{F}^Z to obtain the filtered controlled state process as

$$d\widehat{X}_t^Z = (\mu(\delta - \widehat{X}_t^Z) + \phi l)dt + R_t dN_s + \frac{\alpha}{\sigma} \lambda_t^Z dM_t^Z \quad (36)$$

where M_t^Z is the \mathcal{F}_t^Z innovation process and N is the Poisson process with jump intensity φ . Let $\varphi dt f(\widehat{X}_{t,k}^Z, r_{t,k}) dR$ be the \mathbb{F}^Z compensator for the jump process. The complete information optimization problem is given as

$$\begin{aligned} \sup_{l \in U} \widehat{J}(t, \widehat{x}) &= -\inf_{l \in U} -\widehat{J}(t, \widehat{x}) \\ &= -\inf_{l \in U} \mathbb{E}_{t, \widehat{x}} \left(\int_t^T e^{-\omega s} \left(\frac{1}{2} \beta l_s^2 - l_s(p - \kappa) \right) ds \right) \end{aligned} \quad (37)$$

Thus the lender seeks to obtain the value function $V(t, \widehat{x}) = \sup_{l \in U} \widehat{J}(t, \widehat{x})$ and the optimal loan limit l^* such that $V(t, \widehat{x}) = V_{l^*}(t, \widehat{x}) = \sup_{l \in U} \widehat{J}(t, \widehat{x})$. (37) with dynamics (36) is a linear quadratic Gaussian regulator with jump diffusion problem. We solve the optimization problem using dynamic programming approach.

Definition 7. The admissible set U is defined as

$$\begin{aligned} U &= \left\{ l = (l_t)_{t \in [0, T]}, l \text{ is } F^Z \text{ adapted, } l_t \in [0, \bar{L}], \widehat{X}_t^Z \right. \\ &> 0, \mathbb{E} \left(\int_0^T l_s \right) < \infty, \mathbb{E} \left(\int_0^T \int_{-\infty}^{\infty} |h(X_u^l + R) \right. \\ &\quad \left. - h(X_u^l) \right| f(X_u^l, r) dR du \Big) < \infty, \forall h(x) \\ &\left. \in C^2 \right\} \end{aligned} \quad (38)$$

where C^2 is the family of functions $h(x)$ such that h and its derivatives h_x, h_{xx} are continuous on \mathbb{R} . The upper bound on the loan amount \bar{L} is established exogenously. The upper bound may be in respect to the credit product specification or the lender's risk management criteria.

4.1. The Dynamic Programming Method. We know that if the value function $V(\widehat{x}) \in C^2$, then it satisfies the Hamilton-Jacobi-Bellman (HJB) equation

$$\begin{aligned} \inf_{l \in H} \left\{ -\omega V + \frac{1}{2} \beta l^2 - l(p - \kappa) + (\mu(\delta - x) + \phi l) V_x \right. \\ \left. + \frac{1}{2} \frac{\alpha^2 \lambda^2}{\sigma^2} V_{xx} + \varphi (\mathbb{E}(V(x + R) - V(x))) \right\} = 0, \end{aligned} \quad (39)$$

$$\forall (t, x) \in [0, T) \times \mathbb{R}$$

We proceed to solve the optimization problem in the following proposition.

Proposition 8. The function

$$v(x) = -\left(Jx^2 + Kx + L \right) \quad (40)$$

solves the HJB equation with the constants J, K , and L given as

$$\begin{aligned} J &= \frac{\beta(\omega + 2\mu)}{2\phi^2} \\ K &= \frac{-(\omega + 2\mu)(\beta\mu\delta + \beta\phi\mu_t^1 + \phi(p - \kappa))}{\mu\phi^2} \\ \text{and } L &= \frac{(\omega + 2\mu)}{2\omega\beta\phi^2\mu^2} \left(\omega(\beta\mu\delta + \beta\phi\mu_t^1 + \phi(p - \kappa))^2 \right. \\ &\quad \left. + 2\beta^2\mu^2 \left(\frac{\alpha^2\lambda^2}{\sigma^2} + \phi\mu_t^2 \right) + \frac{\phi^2\mu^2(p - \kappa)^2}{(\omega + 2\mu)} \right) \end{aligned} \quad (41)$$

Proof. By solving the minimization problem in the HJB equation, we obtain the candidate optimal loan amount as

$$l = \frac{1}{\beta} ((p - \kappa) - \phi\nu_x) \quad (42)$$

Plugging this into the HJB equation yields the following second-order ODE equation:

$$\begin{aligned} -\omega\nu + \mu(\delta - x)\nu_x + \frac{1}{2}\frac{\alpha^2\lambda^2}{\sigma^2}\nu_{xx} \\ + \phi(\mathbb{E}(\nu(t, x + r) - \nu(t, x)) \\ - \frac{1}{2\beta}((p - \kappa) - \phi\nu_x)^2 = 0 \end{aligned} \quad (43)$$

Given that the value function is a twice continuously differentiable function of x , we make the following quadratic ansatz:

$$\nu(x) = -(Jx^2 + Kx + L) \quad (44)$$

whereby J, K , and L are constants to be determined. The derivatives of ν are given by $\nu_x = -2Jx - K$ and $\nu_{xx} = -2J$. Further, we have

$$\begin{aligned} \phi(\mathbb{E}(\nu(t, x + r) - \nu(t, x)) \\ = \phi(-\mathbb{E}(J(x + R)^2 + K(x + R) + L) + Jx^2 + Kx \\ + L) \\ = \phi(-\mathbb{E}(Jx^2 + JR^2 + 2JRx + K(x + R) + L) \\ + Jx^2 + Kx + L) = -\phi(J\mu_t^2 + 2Jx\mu_t^1 + K\mu_t^1) \end{aligned} \quad (45)$$

Inserting the derivatives of ν together with (45) and rearranging yields the following:

$$\begin{aligned} x^2 \left(\omega J + 2J\mu - \frac{2J^2\phi^2}{\beta} \right) + x \left(\omega K - 2J\mu\delta + K\mu \right. \\ \left. - 2J\phi\mu_t^1 - \frac{2J\phi^2K}{\beta} - \frac{2\phi J(p - \kappa)}{\beta} \right) + \left(\omega L - \mu\delta K \right. \\ \left. - \frac{\alpha^2\lambda^2J}{\sigma^2} - J\phi\mu_t^2 - K\phi\mu_t^1 - \frac{(p - \kappa)^2}{2\beta} - \frac{\phi^2K^2}{2\beta} \right. \\ \left. - \frac{\phi K(p - \kappa)}{\beta} \right) = 0 \end{aligned} \quad (46)$$

which holds for all (t, x) . The terms in the large parentheses must equal zero. As such solving for the constants J, K , and L we conclude the proof. The case $J = 0$ is trivial since then, with the optimal control $l(t, x) = (p - \kappa)/\beta$ being a constant. \square

Proposition 9. (i) The value function is given by

$$V(x) = -(Jx^2 + Kx + L) \quad (47)$$

where

$$\begin{aligned} J &= \frac{\beta(\omega + 2\mu)}{2\phi^2} \\ K &= \frac{-(\omega + 2\mu)(\beta\mu\delta + \beta\phi\mu_t^1 + \phi(p - \kappa))}{\mu\phi^2} \end{aligned} \quad (48)$$

$$\begin{aligned} \text{and } L &= \frac{-(\omega + 2\mu)}{2\omega\beta\phi^2\mu^2} \left(\omega(\beta\mu\delta + \beta\phi\mu_t^1 + \phi(p - \kappa))^2 \right. \\ &\quad \left. + 2\beta^2\mu^2 \left(\frac{\alpha^2\lambda^2}{\sigma^2} + \phi\mu_t^2 \right) + \frac{\phi^2\mu^2(p - \kappa)^2}{(\omega + 2\mu)} \right) \end{aligned}$$

(ii) The optimal credit limit process is given by

$$\begin{aligned} l_t^* &= \left(\frac{\omega + 2\mu}{\phi} \right) \widehat{X}_t^Z - \frac{1}{\mu\beta\phi} (\phi(p - \kappa)(\omega + \mu) \\ &\quad + \beta(\omega + 2\mu)(\mu\delta + \phi\mu_t^1)) \end{aligned} \quad (49)$$

Proof.

(i) Given that the solution $\nu(x)$ of the HJB equation is a function in C^2 , it follows by verification theorem for jump diffusion processes that the value function $V(x) = \nu(x)$.

(ii) In addition, the candidate optimal loan limit process l_t^* obtained from optimization problem with the HJB is indeed the optimal loan limit process. Reference [6] provides an excellent review of the verification theorem for jump diffusion processes \square

The optimal loan limit is a straight line with y-intercept depended on the expectation of jump size R_t . Since the

density of R_t is depended on the state of the system at time t , the intercept will change in line with the time varying values of $X_{t,k}^Z$. The positive gradient of the line is independent of the jump sizes.

It may be optimal for the lender to decline to avail any credit line to the borrower, i.e., to have $l_t^* = 0$. This occurs when

$$\widehat{X}_t^Z = \frac{\phi(p - \kappa)(\omega + \mu) + \beta(\omega + 2\mu)(\mu\delta + \phi\mu_t^1)}{\mu\beta(\omega + 2\mu)} \quad (50)$$

In this case, for \widehat{X}_t^Z to be positive we must have $\mu_t^1 > (-1/\phi)(\mu\delta + \phi(p - \kappa)(\omega + \mu)/(\omega + 2\mu))$.

Corollary 10. *For the case when there is no network data, i.e., when $\mathbb{F} = \mathbb{F}^Y$, the value function is given by*

$$V_0(x) = -(Jx^2 + Kx + L) \quad (51)$$

where

$$\begin{aligned} J &= \frac{\beta(\omega + 2\mu)}{2\phi^2}, \\ K &= \frac{-(\omega + 2\mu)(\beta\mu\delta + \phi(p - \kappa))}{\mu\phi^2} \\ \text{and } L &= \frac{(\omega + 2\mu)}{2\omega\beta\phi^2\mu^2} \left(\omega(\beta\mu\delta + \phi(p - \kappa))^2 \right. \\ &\quad \left. + 2\beta^2\mu^2 \left(\frac{\alpha^2\lambda^2}{\sigma^2} \right) + \frac{\phi^2\mu^2(p - \kappa)^2}{(\omega + 2\mu)} \right) \end{aligned} \quad (52)$$

and the optimal credit limit process is

$$\begin{aligned} l_{0,t}^* &= \left(\frac{\omega + 2\mu}{\phi} \right) \widehat{X}_t^Y \\ &\quad - \frac{1}{\mu\beta\phi} (\phi(p - \kappa)(\omega + \mu) + \mu\delta\beta(\omega + 2\mu)) \end{aligned} \quad (53)$$

Proof. To obtain the optimal value function when the observation filtration is \mathbb{F}^Y , simply replace the jump intensity parameter $\varphi = 0$ in Proposition 9 and hence the proof. \square

When $\varphi = 0$, the lender's optimization problem reduces to the normal linear Gaussian quadratic regulator and the objective function $J(t, x) = \widehat{J}(t, \widehat{x})$; see, e.g., [7].

Remark 11. The impact of the network information on the lender's optimal decision process is seen by noting that whenever $\mu_t^1 < 0$ for all t then $l_t^* \geq l_{0,t}^*$. This, coupled with the variance results in Proposition 6, captures the gains made by incorporating the borrower's ego-network information in the lender's credit risk strategy.

TABLE 1: Model parameter values.

Mean reversion speed δ	0.05	Drift α	2
Return's volatility σ	0.3	Volatility γ	1
Mean reversion level μ	2.0	Network variance Λ	0.04

5. Numerical Results

In this section we illustrate our findings in the preceding sections. We model the hidden credit quality process as an Ornstein Ulehnbeck process and the observed returns process as a linear diffusion process. We assume that the network information \mathbf{Z}_k arrives at equidistant time points t_k , $k = 0, 1, \dots, N$. We simulate the processes using the parameter values as found in Table 1.

In order to capture how the network information causes jumps on the filtered estimates and variances, Figure 1 shows the drift X_t , filters \widehat{X}_t^H for $H \in \{Y, O, Z\}$ and variances λ_t^H for the case when the borrower's degree η_t is fixed with values $\{0, 2\}$. The borrower's own network information Z_{itk} is indicated with red cross. We choose the number of network information times $N = 6$. In panels one and two, we note that on information dates the filters \widehat{X}_t^O and \widehat{X}_t^Z jumps towards Z_{itk} whilst between information dates the filters are driven back to the stationary mean of the drift δ . In panels 3 and 4, we note that λ_t^Y moves towards its asymptotic value $C_0 - \mu\sigma^2/\alpha^2$ as $t \rightarrow \infty$. λ_t^O and λ_t^Z rise between information dates but jump down on information dates.

To capture well the effect of number of direct friends on the conditional variances λ_t^O and λ_t^Z Figure 2 plots a comparison between the variances of λ_t^O and λ_t^Z with and without friends. The left panel plots a comparison of the variances for the case when the number of friends is zero and five. For the right panel, we model the random number of friends at information dates t_k as a Poisson random variable with parameter $\rho = 5$. The conditional variances for the case when there exists friends' data in \mathbf{Z}_k is lower as compared to the case with zero friends. The right panel captures the perturbation of conditional variance as the number of friends varies.

6. Conclusion

In this paper, we have presented the detailed results for the estimates and properties of conditional variance for a hidden Ornstein-Ulehnbeck equation driving the drift process of a borrower's observed credit score. The network information obtained from the borrower's evolving ego-network has been found to improve the quality of the lender's estimate. These results offer a theoretical basis for the emerging area of unsecured digital and mobile phone based lending prevalent in emerging market economies. We have also solved a stochastic optimization problem whereby the lender is able to obtain the optimal credit limit for a borrower, the optimal limit being a linear function of the borrower's credit quality estimate. The optimization problem is reduced to a linear quadratic Gaussian problem with jump

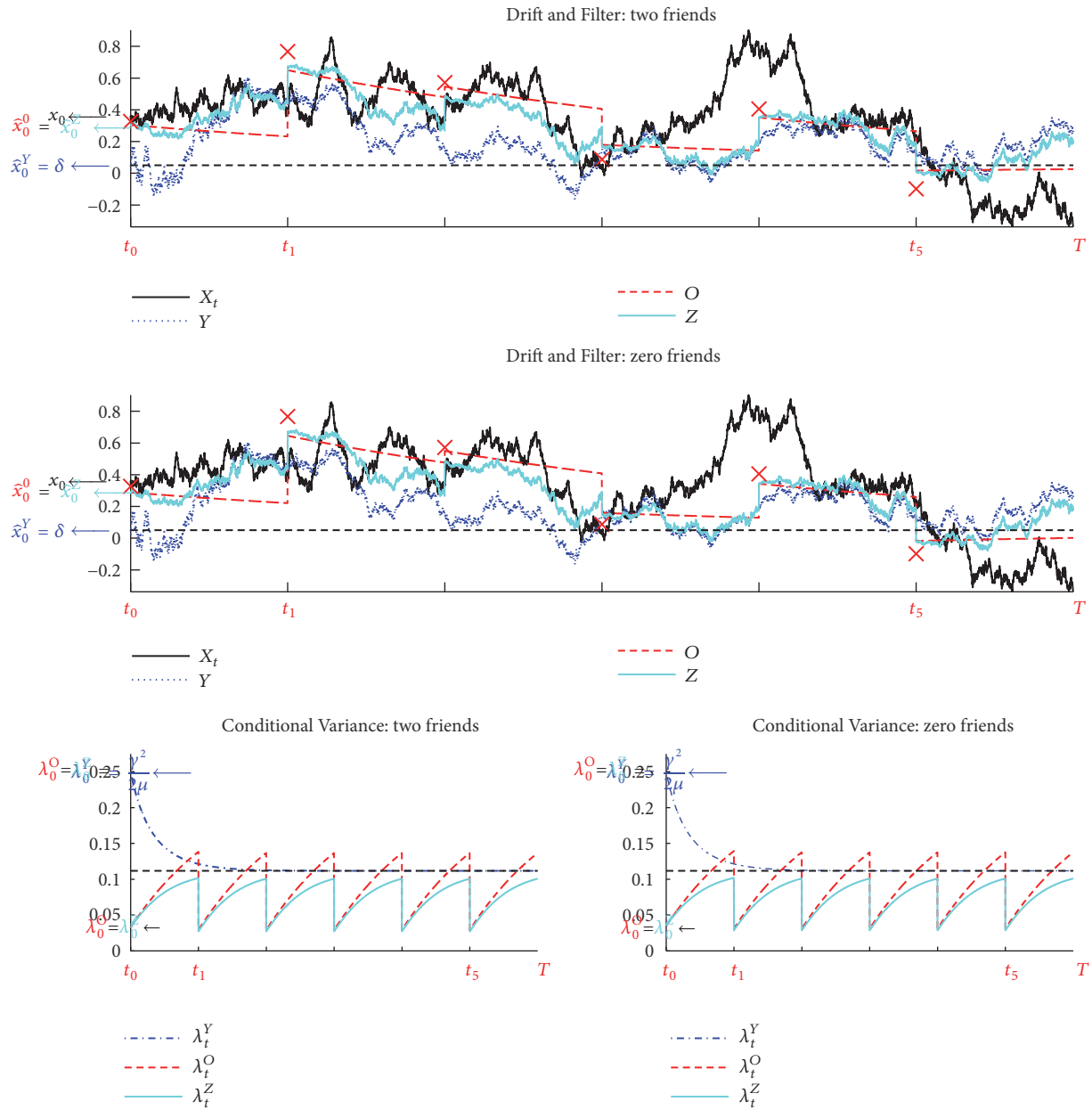


FIGURE 1: Panel 1: The drift X_t and filtered estimates X_t^H for the case when the borrower with degree $\eta = 2$. Panel 2: The drift X_t and filtered estimates X_t^H for the case when the borrower with degree $\eta = 0$. Panel 3 and 4: The conditional variances λ_t^H when the degrees $\eta_t = 2$ and $\eta_t = 0$.

diffusion. The jumps caused by the network information arriving at random times have been found to impact on the credit limit amount availed to the borrower from time to time.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

The authors wish to thank African Union and Pan African University Institute for Basic Sciences Technology and

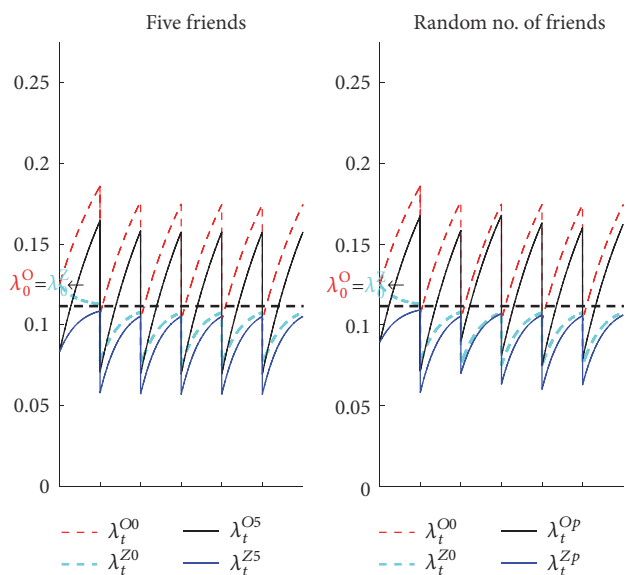


FIGURE 2: Left: Plot of variances λ_t^H when no. of friends $\eta_t = 0$ and $\eta_t = 5$ Right: Plot of variances λ_t^H when no. of friends $\eta_t = 0$ and $\eta_t = \rho$.

Innovation, Kenya, for their financial support for this research.

References

- [1] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: homophily in social networks," *Annual Review of Sociology*, vol. 27, pp. 415–444, 2001.
- [2] Y. Wei, P. Yildirim, C. Van Den Bulte, and C. Dellarocas, "Credit scoring with social network data," *Marketing Science*, vol. 35, no. 2, pp. 234–258, 2016.
- [3] A. Gabih, H. Kondakji, J. Sass, and R. Wunderlich, "Expert opinions and logarithmic utility maximization in a market with Gaussian drift," *Communications on Stochastic Analysis*, vol. 8, no. 1, pp. 27–47, 2014.
- [4] M. Leippold, P. Vanini, and S. Ebnöther, "Optimal credit limit management under different information regimes," *Journal of Banking & Finance*, vol. 30, no. 2, pp. 463–487, 2006.
- [5] R. Cerqueti, "Financing policies via stochastic control: a dynamic programming approach," *Journal of Global Optimization*, vol. 53, no. 3, pp. 539–561, 2012.
- [6] B. K. Øksendal and A. Sulem, *Applied Stochastic Control of Jump Diffusions*, Springer, 2005.
- [7] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer Science & Business Media, New York, NY, USA, 1975.
- [8] E. Francis, J. Blumenstock, and J. Robinson, *Digital credit: A snapshot of the current landscape and open research questions*, CEGA White Paper, 2017.
- [9] T. Snijders, C. Steglich, and M. Schweinberger, *Modeling the coevolution of networks and behavior*, pp. 41–71, 2017.
- [10] P. D. Hoff, A. E. Raftery, and M. S. Handcock, "Latent space approaches to social network analysis," *Journal of the American Statistical Association*, vol. 97, no. 460, pp. 1090–1098, 2002.
- [11] D. K. Sewell and Y. Chen, "Latent space models for dynamic networks," *Journal of the American Statistical Association*, vol. 110, no. 512, pp. 1646–1657, 2015.
- [12] D. Durante and D. B. Dunson, "Nonparametric Bayes dynamic modelling of relational data," *Biometrika*, vol. 101, no. 4, pp. 883–898, 2014.
- [13] B. Kim, K. H. Lee, L. Xue, and X. Niu, "A review of dynamic network models with latent variables, 2017," <https://arxiv.org/abs/1711.10421>.
- [14] T. R. Bielecki, M. Jeanblanc, and M. Rutkowski, "Introduction to mathematics of credit risk modeling," in *Stochastic Models in Mathematical Finance*, CIMPA–UNESCO–MOROCCO School, pp. 1–78, Marrakech, Morocco, 2007.
- [15] F. W. M. De Andrade and L. C. Thomas, "Structural models in consumer credit," *European Journal of Operational Research*, vol. 183, no. 3, pp. 1569–1581, 2007.
- [16] M. Malik and L. C. Thomas, "Transition matrix models of consumer credit ratings," *International Journal of Forecasting*, vol. 28, no. 1, pp. 261–272, 2012.
- [17] D. B. Ntwiga, *Social network analysis for credit risk modeling [Ph.D. thesis]*, School of Mathematics, University of Nairobi, Kenya, July 2016.
- [18] R. J. Elliott, L. Aggoun, and J. B. Moore, *Hidden Markov Models: Estimation and Control*, Springer Science & Business Media, NY, USA, 2008.
- [19] S. N. Cohen and R. J. Elliott, *Stochastic Calculus and Applications*, Springer, 2015.
- [20] A. Bain and D. Crisan, *Fundamentals of stochastic filtering*, vol. 60 of *Stochastic Modelling and Applied Probability*, Springer, New York, 2009.
- [21] F. Black and R. Litterman, "Global portfolio optimization," *Financial Analysts Journal*, vol. 48, no. 5, pp. 28–43, 1992.
- [22] A. Danilova, M. Monoyios, and A. Ng, "Optimal investment with inside information and parameter uncertainty," *Mathematics and Financial Economics*, vol. 3, no. 1, pp. 13–38, 2010.
- [23] J.-D. Benarous, L. Mazliak, and R. Rouge, "Filtering and control with information increasing," *Methodology and Computing in Applied Probability*, vol. 2, no. 2, pp. 123–135, 2000.
- [24] P. Lakner, "Optimal trading strategy for an investor: the case of partial information," *Stochastic Processes and Their Applications*, vol. 76, no. 1, pp. 77–97, 1998.
- [25] B. Ngugi, "CBA plans M-Shwari loan rate cut to fend off rivals," *Daily Nation*, 2017.
- [26] K. Macharia, "Safaricom revels in M-Pesa as borrowers take Sh14m per hour," *The Standard*, 2016.