

Does Adaptive Learning Neutralize Interbank Market Liquidity Hoarding under a Distressed Market Condition?

Cheuk Yin Jeffrey Mo, Steve Y. Yang*, Xingjia Zhang

School of Business, Stevens Institute of Technology, Hoboken, NJ, USA
Email: *steve.yang@stevens.edu

How to cite this paper: Mo, C. Y. J., Yang, S. Y., & Zhang, X. J. (2024). Does Adaptive Learning Neutralize Interbank Market Liquidity Hoarding under a Distressed Market Condition? *Theoretical Economics Letters*, 14, 1267-1293.

<https://doi.org/10.4236/tel.2024.143064>

Received: February 19, 2024

Accepted: June 25, 2024

Published: June 28, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This paper investigates how counterparty risk and precautionary liquidity concerns can lead to dry-up of liquidity in the interbank market with adaptive bank agents sharing common objectives of collecting as much interest payment while minimizing the insolvency risk. The agents are calibrated based on the US bank balance sheet data and are allowed to make autonomous decisions under different simulated market conditions. Through simulation studies, we observe an endogenously formed multi-layer lending network exhibiting the well-known core-periphery structure. Moreover, we find the adaptive learning banks would endogenously form liquidity hoarding phenomenon under an exogenous shock to a subset of the banks in the system. We also find that in an adaptive learning environment, fire sales would lead to a decrease in interbank liquidity and increase the probability of contagion through interbank lending networks.

Keywords

Reinforcement Learning, Agent-Based Modeling, Contagion Risk, Fire Sales, Network Topology, U.S.

1. Introduction

In the mid of the great financial crisis (GFC), problems arose in the interbank lending markets where the U.S. Libor spreads rose significantly, reflecting that banks became much more unwilling to lend money to other banks. The Federal Reserve conducted a variety of measures to improve the smooth functioning of the money markets, and yet the liquidity crisis prevailed in the end resulting in the 2007-09 financial crisis. Since then the emergence of systemic event and sys-

tem instability has drawn increasing attention among the finance researchers (Stiglitz, 2010; Bisiyas et al., 2012; Glasserman & Young, 2015; Roukny et al., 2018) and regulators (Yellen, 2013; Confidence, 2012). Numerous theoretical and empirical studies across different disciplinary areas have investigated financial system stability problems (Bisiyas et al., 2012; Glasserman & Young, 2015). There are also recent studies focusing on liquidity spillover effect in the interbank system in view of incentive of interbank lenders' behaviors (Duarte & Eisenbach, 2021; Li & Lai, 2022). However, many financial models, i.e. search-based and network-based (Gofman, 2017), do not fully reflect the complexity of the financial system and the diverse nature of the individual decisions of different economic agents. In this paper, we aim to examine the effect of individual bank decisions on the interbank collective behaviors.

We argue the existing methods for financial counterparty risk modeling and consequent contagion modeling suffer three major shortcomings: 1) most of the models tackle the problem with highly stylized structures; 2) the assumption of rational agents and equilibrium points in bank lending and borrowing is broadly applied; 3) organizations' decisions and behaviors are largely missing in the dynamic modeling process. The reality is that financial institutions are more performance-driven, and their performance may be individually optimal, but collectively suboptimal (Acharya, 2009). Furthermore, these organizations are autonomous decision-making agents with various constraints, and they learn and adapt to market changes in order to achieve their performance objectives, which are very dynamic in nature. This characteristic is largely not present in a pure network optimization setting or an equilibrium-based approach, and therefore when these networks are used for stress testing of real banking systems, the results likely deviate significantly from the reality.

We propose to build a heterogeneous agent model where agents (i.e. banks) are learning with reinforcement learning style of exploration and exploitation to maximize expected long-term rewards. This agent decision process is modeled as a Markov decision process (MDP). Through interactions of multiple agents, i.e. lending and borrowing from each other, the interbank market is cleared using the established fixed point clearing process (Eisenberg & Noe, 2001). Such a modeling framework eliminates the equilibrium-based and rational expectations approach in the current literature focusing on individual agents' decisions and the emergence of network structure. Moreover, it allows researchers to calibrate the model using empirical observations, and the calibrated models would allow policymakers and regulators to associate bank decision patterns with bank failures and contagion events. This methodology bridges significant gaps in the existing literature and addresses the shortcomings, i.e. highly stylized approaches, rational expectation equilibrium assumption, and lack of empirical organizational decision perspective.

The adaptive learning models assume that agents are faced with repeated decisions using certain learning rules to maximize their individual long-term rewards

under various constraints. These decision rules can be learned from actual economic data we collect. This paper investigates the hypothesis that interbank lending decisions and behaviors are determined by both the economic environment and the individual organization preferences in banking system. By studying adaptive behaviors of learning agents, we aim to answer the following questions: How do adaptive learning banks endogenously form liquidity hoarding? How liquidity hoarding is influenced by fire sales? How does contagion risk evolve with adaptive learning agents?

In this study, we propose a learning agent model with policy gradient style of reinforcement learning method. Through this model, we aim to incorporate the adaptive agents' reaction to macroeconomic environment changes and information diffusion in the interbank market (Gilbert & Terna, 2000; Macy & Willer, 2002). Reinforcement learning methods are often used in developing learning agents in computer artificial intelligence research community. In Figure 1, the method allows agents to select an action (A_t) for each state (S_t) that tends to maximize the expected future value (R_t) of the reinforcement signal (Sutton et al., 2000).

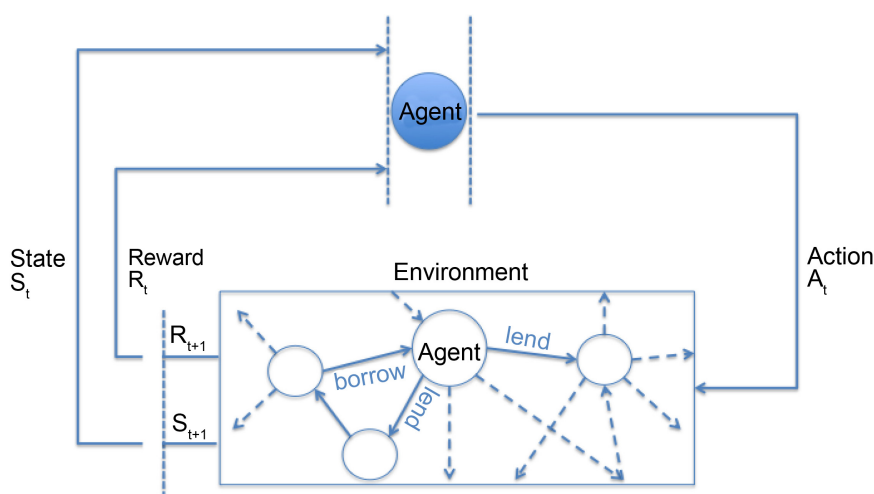


Figure 1. Reinforcement learning agent in an interbank environment.

More importantly, the model can be calibrated with the observed state variables, such as bank balance sheet information, various asset prices, etc., and then a large-scale Monte Carlo simulation is constructed to examine contagion effect and answer various policy-related questions. As such, we calibrate the model with 6600 bank balance sheet data collected by Federal Financial Institutions Examination Council (FFIEC) as well as asset prices from Bloomberg Financial Data from 2001 to 2014. After calibrating the model with the pre-crisis data, we perform two major experiments. First, we induce the mortgage crisis into the system and study liquidity hoarding behaviors of participating banks. Secondly, we introduce fire sales (Cifuentes et al., 2005) and further examine the liquidity hoarding contagion risk with and without fire sales.

The first contribution the paper makes to the literature is that we show when banks learn that others are tightening their balance sheets, they adapt their policies accordingly, which results in overall liquidity hoarding under a distress market condition. Many have argued the existence of liquidity hoarding behaviors (Cifuentes et al., 2005; Keister & McAndrews, 2009; Acharya & Skeie, 2011), but we use a computational learning model to show that hoarding is a result of banks' learning and adaptation behaviors. We further examine the model from network perspective, and compare it with the observed bank failures in the 2007-09 financial crisis, and confirm that the adaptive banks tend to hoard liquidity in order to maximize their total utility given the information they observe in the interbank market.

The second contribution of the paper is that fire sales would drive banks to increase the interbank lending in order to balance their market risk under a distressed market. In other words, adaptive banks hoard less liquidity when fire sales are considered. We show when faced with a trade-off between market risk caused by fire sales and counterparty risk, banks choose to lend more in the interbank market. We show that liquidity hoarding and fire sales jointly change the topology of interbank markets, and in the end, they tend to reduce the clustering, shortest path, and average degree of interbank networks, which would reduce the probability of contagion through interbank markets.

The rest of the paper is organized as follows. In Section 2, we review the related literature and describe data used in Section 3. In Section 4, we present the general methodology of the learning agent model. In Section 5, we validate model by comparing interbank lending network topology with the empirical findings. In Section 6, we examine the liquidity hoarding effect by replicating the 2008-09 mortgage crisis. We then introduce fire sales into the model and examine their effect on liquidity hoarding and contagion risk respectively. Finally, we conclude the paper by summarizing the major findings and suggest future applications.

2. Related Literature

2.1. Liquidity Risk and Contagion Risk

Liquidity played a very important role in the 2008-09 financial crisis. Many theoretical frameworks have been proposed to study the cause of liquidity provision and its contribution to the crisis. Allen and Carletti (2008) argued that following features are related to liquidity and have greatly exacerbated liquidity hoarding in 2008-09 financial crisis: 1) prices of AAA-rated tranches of securitized products went below fundamental values; 2) the dry-up of liquidity in interbank markets is attributed to a mixture of liquidity hoarding by banks. In normal times, high quality asset backed securities are close substitutes for collateral in the money markets. However, in crisis times, they are not. Heider et al. (2015) provided a model to explain how the risk of banks' long-term assets can lead to liquidity hoarding in interbank market. They showed that in a particular level

and distribution of risk among banks, interbank market may not reach an equilibrium in which liquidity is reallocated smoothly. Afonso and Shin (2009) studied liquidity and systemic risk in high-value payment systems. Using lattice-theoretic methods to solve for the unique fixed point of an equilibrium mapping, they found that banks' attempting to conserve liquidity can cause disruption of payments. Gai and Kapadia (2010) showed how liquidity hoarding can cascade through a banking network and cause severe consequences. He argued that such systemic breakdowns of the interbank market can be explained by concerns about future liquidity needs.

Interbank contagion risk is related to liquidity risk. Cifuentes et al. (2005) and Boyson et al. (2010) proposed to use quantile and logit regression models to study contagion among worst returns in the hedge fund industry and found that large adverse shocks to asset and hedge fund liquidity strongly increase the probability of contagion. Ballester et al. (2006) analyzed how economic agents' liquidity holding decisions in a network game generate systemic risk and by structurally estimating network externalities. Eross et al. (2016) liquidity shock spreading within the interbank market is associated with benchmark interest movements which can be used a tool for policy decisions on benchmark interest rate. Brandi et al. (2018) used a pandemic spreading modeling technique and show that the interbank network has improved through the subsequent micro-prudential and liquidity hoarding policies adopted by banks, as a result it increased the network resilience to systemic risk, yet with the undesired side effect of drying out liquidity from the market. Herskovic et al. (2020) embed a similar spatial autoregressive structure in firms' growth rates to study the comovement of firm volatilities. Denbee et al. (2021) provided an empirical framework to attribute systemic risk to individual banks and by characterizing the wedge between decentralized outcome and the planner's solution. They show that banks' liquidity equilibrium holdings can be strategic complements or substitutes which may amplify or dampen shocks to individual banks.

2.2. Multilayer Network of Interbank Markets

Multilayer network has been recently applied in interbank network studies. Buldyrev et al. (2010) found that interdependent networks become significantly more vulnerable to contagion compared to their non-interacting sub-networks. Bargigli et al. (2015) analyzed different layers of Italian banks over an interesting time lapse. They found that layers have different topological properties and persistence over time. Using the total interbank network or focusing on a specific layer as representative of the other layers provides a poor representation of interlinkages in the interbank market and could lead to biased estimation of systemic risk. Montagna and Kok (2016) developed an agent-based multi-layered interbank network models including short-term interbank, long-term interbank, and common exposures in banks' securities portfolios. They found that the impact of different layers are non-negligible and non-linear in the propagation of

shocks to individual banks. That is the accumulative contagion effects from multiple layers can be substantially larger than the sum of the individual layers. Zhang et al. (2018) studied the impact of interbank risk exposures that arise due to fire sales in the interbank lending market. They found the fire sales in portfolio overlap network changed interbank lending network topology. Interbank lending networks with fire sales tend to form a more connected network. Aldasoro et al. (2017) built an interbank network model and demonstrated risk contagion through liquidity hoarding, interbank interlinkages and fire sale externalities.

2.3. Reinforcement Learning in Computational Economics

The “reinforcement” represents “actions or strategies that have yielded relatively high (low) payoffs in the past are more (less) likely to be played in the future” (Amman et al., 1996). It has a long history associated with behavioral psychologists since 1920s. And the models of reinforcement learning first appeared in Bush and Mosteller (1955). Models of reinforcement learning for economics didn’t appear until 1990s. Arthur was among the first economists who studied agent behaviors utilizing reinforcement-type learning (Arthur, 1991). They demonstrated the ability to design artificial learning agents and calibrate their “rationality” to replicate human behaviors in a N-bandit problem. Erev and Roth (1998) extended Arthur’s study to multi-player games. They examine a unique equilibrium in mixed strategies. Sarin et al. (2001) incorporated agents maximizing subjective assessments through period $t - 1$, and the method demonstrated better performance than Arthur’s method. Duffy and Feltovich (1999) claimed that agents can not only learn information from past experience, but also learn from other agents. Lux (2015) built a dynamic multi-agent learning system where banks with the fundamental reinforcement learning concept to select their counterparties and showed that banks will experience regular deposit shocks and they have to act accordingly in order to remain solvent.

Reinforcement learning (RL) has also been developed in machine learning. It is regarded as a learning to maximize accumulative reward. The learner must discover which actions yield the most reward by trying, instead of being told which actions to take (Sutton et al., 2000). People in computer science field developed model-free RL algorithms to solve situations without period payoff or state transition function, like Q-learning (Watkins & Dayan, 1992) and policy gradient (Sutton et al., 2000). It’s not a new area to apply RL in economics (Hopkins, 2007; Bossan et al., 2015; Berardi & Galimberti, 2017).

RL algorithms have been widely used in agent-based modeling maybe because they are straightforward and simple. Axtell et al. (1996) used a variant of reinforcement learning in their sugarscaped model to generate rich “artificial histories”, scenarios that display stylized facts of interest, such as cultural differentiation driven by resource availability, migration, trade, and combat. Nicolaisen et al. (2001) used RL to model electricity market in which buyers and sellers use a

modified Roth-Erev individual reinforcement learning algorithm to determine their price and quantity offers in each auction round. Pemantle and Skyrms (2004) used RL to study social network formation based on three-person interactions with uniform positive reinforcement. Liu et al. (2018) modeled individual bank decisions using the temporal difference reinforcement learning algorithm based on banks' lending preferences and environment. Hu (2021) applied Bayesian neural networks to RL to reduce uncertainty in molecule design in which agents get to choose one action from a pool of sampled actions at each step and investigate its benefits. They demonstrated that the Bayesian approach could offer a balance between optimality and diversity in learning new policies.

3. Data

Financial reports are one of the key information sources that disclose banks' financial fundamentals and business conditions. The Federal Reserve, FDIC, and OCC require all U.S. regulated banks to submit quarterly reports known as *Federal Financial Institutions Examination Council Reports of Condition and Income*. Those banks include national banks, state member banks and insured state nonmember banks. Similar to regulators that rely on balance sheets to monitor banks' liquidity status and banking system structures, we use balance sheets data from March, 2001 to December, 2014, covering around 10,000 banks (see Table 1). The reason we choose this period of data is because the focus of the study is to examine the U.S. bank behavior changes before and after the great financial crisis (GFC) (2008-09). The training and validation data cover 7 years before the crisis and 7 years after the crisis. By introducing the housing price shock during the GFC, we will be able to observe individual bank behavior changes as well as the system wide interbank lending liquidity responses. We use outstanding volume from Securities Industry and Financial Markets Association (SIFMA) and GDP from the Bureau of Economic Analysis (BEA) of the United States Department of Commerce to measure market depth for the same period, and the housing price index from Federal Housing Finance Agency from Jun 2007 to Dec 2014. The housing price index from the Federal Housing Finance

Table 1. Description of the bank balance sheet.

Asset: A	Liability: L
Overnight lending: ON^d	Overnight borrowing: ON^b
Short-term lending: ON^d	Short-term borrowing: ON^b
Long-term lending: ON^d	Long-term borrowing: ON^b
Cash and balance due: C	Other liabilities: OL
Other assets: OA	Equity: E

Notes: This description of a bank's balance sheet focuses on major bank lending and borrowing channels, i.e. overnight, short-term, and long-term markets. The rest of the balance sheet is condensed into cash or other asset or liabilities. The notations introduced here for aspect of the balance sheet will be used throughout this paper.

Agency (FHFA) is a valuable measure to assess the impact of the 2008-09 Financial Crisis on the housing market. The crisis was significantly influenced by the collapse of the housing bubble, making the FHFA index a critical indicator of the price shock. In this study, we use the housing price index from 2007 to 2014 as the exogenous shock to the banking system for the validation/experimentation of the proposed learning agent model.

4. Methodology

The primary objective of the learning agent modeling approach is to take the relationships and behaviors of individual banks into account in their lending and borrowing decisions and examine the emergent risk implications to the entire interbank market. In this study, we follow the framework proposed by Liu et al. (2017) (consequently we refer it as Liu's model) that includes two types of banks (small banks and large banks) and three types of debts (overnight, short-term and long-term debts). When banks are represented as nodes linked through the different types of debts, the interbank market naturally forms a multi-layered network which is more realistic than the single network approach.

In Liu's model, when a bank receives a borrowing request, it must decide two things: 1) whether to provide new loans to the requesting borrowers, and 2) how much to lend. A bank chooses to lend by going through each request until its lending target is fully satisfied or there are no more requests to fill. It is based on a relationship scoring system to maintain a list of preferred counterparties.

Liu et al. (2018) extended Liu's model by incorporating temporal difference reinforcement learning algorithms in determining counterparties by establishing a relationship lending scheme. We further extend (Liu et al., 2018) model by allowing bank agents to adaptively learn in determining the percentage of surplus balance to lend out. In the proposed model, we keep the relationship lending scheme in the first step and apply reinforcement learning method to determine how much of the surplus balance to lend out for each bank based on the balance sheet condition of other banks. This extension allows bank agents to learn to assess the entire interbank market condition represented by all balance sheets. We use stylized balance sheet over the period from 2001 to 2014 collected from FFIEC, as shown in previous section.

4.1. Markov Decision Processes and RL Algorithms

In a reinforcement learning problem, an agent observes a state s_t from time t . It interacts with the environment by taking an action a_t in state s_t . After the action is taken, the agent transitions to a new state s_{t+1} based on the current state and the chosen action. In every time step, the agent gets an immediate reward. Formally, RL can be described as an MDP (Arulkumaran et al., 2017), which consists of: a set of states \mathcal{S} , a set of actions \mathcal{A} , transition dynamics $\mathcal{T}(s_{t+1} | s_t, a_t)$, immediate reward function $\mathcal{R}(s_t, a_t, s_{t+1})$, a discount factor $\gamma \in [0, 1]$, where lower values place more emphasis on immediate rewards.

Generally, policy π maps states to a probability distribution $\pi: \mathcal{S} \rightarrow P(\mathcal{A} = a | \mathcal{S})$. The period from the first to the final time step is referred to as an episode. In an episodic environment, the state is reset after each episode of length T , then the gained immediate reward (r) from each step is accumulated as result $R = \sum_{t=0}^{T-1} \gamma^t r_{t+1}$. The aim of RL is to find out an optimal policy π^* , which maximizes the expected accumulative reward.

$$\pi^* = \arg \max E[R | \pi] \quad (1)$$

An MDP follows Markov property, which means that only the current state affects the next state. This means that any decisions made at s_t can be based solely on s_{t-1} , rather than $\{s_0, s_1, \dots, s_{t-1}\}$.

4.1.1. Monte Carlo Policy Gradient

The policy gradient algorithm, a model-free method, is utilized in the proposed model. There are many advantages of using policy gradient. First, policy-based methods have better convergence properties. Compared with value-based methods, in which the choice of action may change dramatically for an arbitrarily small change in the estimated action values, policy gradient provides a smooth update of policy at each step. It guarantees to converge on a local maximum or global maximum. Second, policy gradients can handle cases with high dimensional action space or continuous actions. The Deep Q-learning predicts future by assigning a score (maximum expected future reward) for each possible action, at each time step, given the current state. It would be impossible to generate a Q-value for each possible action. On the other hand, policy gradient only adjusts the parameters directly.

4.1.2. Policy Gradient Updates

Policy gradient updates action policy to maximize agent's objective function $J(\theta)$. In our model, banks aim to maximize their total profits from lending and borrowing subtracting the losses from the write-downs of their lendings and asset holdings. For example, in the overnight market, banks would calculate their interbank profits using the overnight interest rate which we assume at 2% (it is fairly stable). Under fire sales, banks would have to write down asset losses due to excessive price impact. Policy is estimated as a function with policy parameter θ :

$$\pi(a | s, \theta) = P(\mathcal{A}_t = a | \mathcal{S}_t = s, \theta_t = \theta). \quad (2)$$

In a continuous action space, it's natural to assume a Gaussian policy (Chou et al., 2017) for stochastic policy gradient. So, we make the following assumptions:

$$\begin{aligned} \pi(a | s, \theta) &= N(\mu(s), \sigma^2) \\ \mu(s) &= \phi(s)^\top \theta \end{aligned}$$

where variance σ^2 is fixed, $\phi(s)$ is a vector of state features. Then it's score function is:

$$\nabla_{\theta} \log \pi_{\theta}(s, a) = \frac{(a - \mu(s))\phi(s)}{\sigma^2} \tag{3}$$

In an episodic environment, we can optimize accumulative rewards for the start state. The objective function is:

$$J(\theta) = E \left[\sum_{t=1}^T \gamma^{t-1} r_t \mid s_0, \pi \right] = \sum_s d^{\pi}(s) \sum_a \pi_{\theta}(s, a) R_{s,a} \tag{4}$$

where $d^{\pi}(s) = \lim_{t \rightarrow \infty} P(S_t = s \mid s_0, \pi)$. Use likelihood ratios to compute the policy gradient, we get:

$$\begin{aligned} \nabla_{\theta} J(\theta) &= \sum_s d(s) \sum_a \pi_{\theta}(s, a) \nabla_{\theta} \log \pi_{\theta}(s, a) R_{s,a} \\ &= E_{\pi_{\theta}} [\nabla_{\theta} \log \pi_{\theta}(s, a) r] \end{aligned} \tag{5}$$

Policy gradient learns the policy parameter θ based on the gradient of some performance measure $J(\theta)$. Their updates approximate gradient ascent in J :

$$\begin{aligned} \theta_{t+1} &= \theta_t + \alpha \nabla_{\theta} J(\theta_t) \\ &= \theta_t + \alpha \gamma^t R_{s,a} \nabla_{\theta} \log \pi_{\theta}(s, a) \\ &= \theta_t + \alpha \gamma^t R_{s,a} \frac{(a - \mu(s))\phi(s)}{\sigma^2} \end{aligned} \tag{6}$$

4.2. Agent-Based Interbank Lending Market

As shown in **Figure 2**, in each period, bank agents take the following steps to fulfill their lending and borrowing needs: 1) determine how much to borrow or lend for each period; 2) decide counterparties and the lending/borrowing amount; 3) settle debt payments; 4) clear the market and determine bank defaults; and 5) update balance sheets.

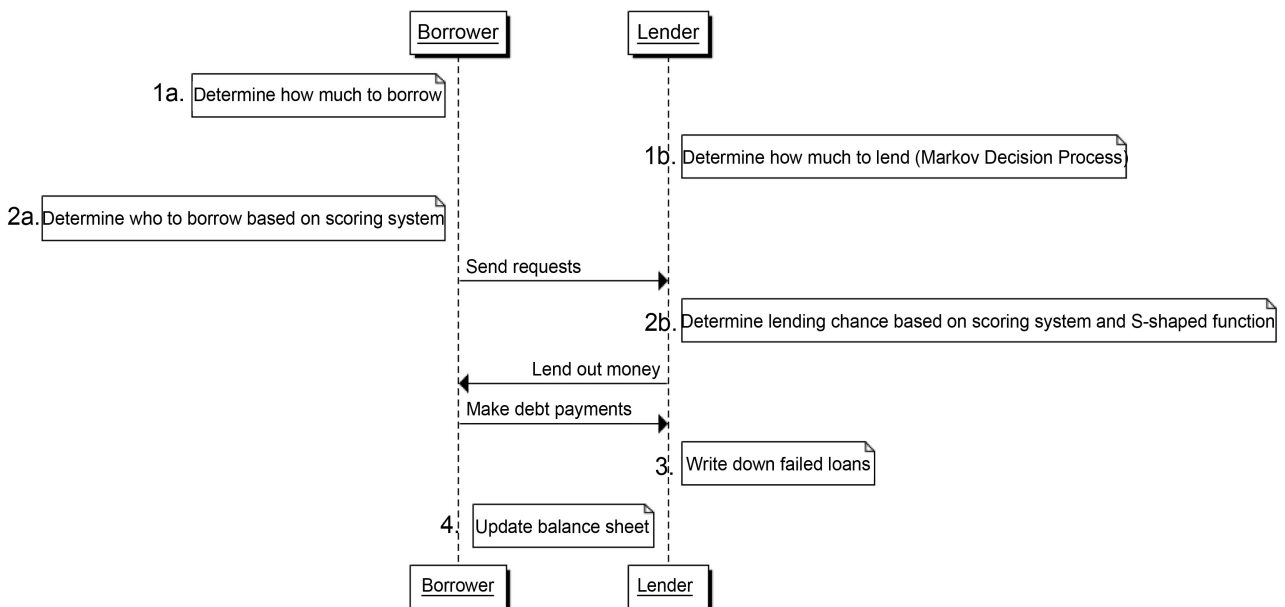


Figure 2. Interbank activities timeline.

At the beginning of each period, banks decide the total lending and borrowing need and send out requests to a list of potential counterparties. At the end of each period, banks pay off debts and write down losses. In every iteration, banks pay back debts with different rates. Overnight debts are paid fully, short-term debt payments and long-term debt payments are drawn from $U(99\%, 100\%)$ and $U(25\%, 100\%)$ respectively. These pay-off plans are based on historical observations. Eisenberg-Noe clearing vector system is applied to clear payments for banks (Eisenberg & Noe, 2001). In this process, banks may default and no longer bear obligation to pay back full debts (see Equations (7) and (9)). In this case, the lenders suffer loss. Secondly banks actively search for lenders in the interbank system. Each bank sends out borrowing requests to potential lenders, and lenders response with approval or rejection. The process continues until demand is fulfilled or requests have been sent to all available banks.

When a bank receives a borrowing request, it must decide whether to provide new loans to requesting borrowers and how much to lend—two primary factors that affect bank lending preferences. Following Liu et al. (2018), we apply a scoring system (described in Equation (7)), which includes relationship score and size score, for each bank to decide on its potential borrowers. Accordingly, a bank chooses to lend by going through each request until its lending target is satisfied or there are no more requests to fill.

$$S_{j,i}^{\text{total}}(t) = \omega S_{j,i}^{\text{relation}}(t) + (1 - \omega) S_{j,i}^{\text{size}}(t), \quad (7)$$

where $S_{j,i}^{\text{total}}$ is the score that lender j assigns to borrower i . $S_{j,i}^{\text{total}}$ is the weighted average of the relationship score and size score of bank i . We set $\omega = 0.5$ in our model.

Temporal difference TD(0) is applied in getting relationship score. When new debts is settled between two banks, their relationship score will be updated based on debt amount using Equation (8).

$$S_{i,j}^{\text{relation}}(t) = S_{i,j}^{\text{relation}}(t) + (1 - \gamma) \left[D_{t+1} + \gamma \frac{D_{t+1}}{1 - \gamma} - S_{i,j}^{\text{relation}}(t) \right] = \gamma S_{i,j}^{\text{relation}}(t) + D_{t+1} \quad (8)$$

where γ is discount rate and learning rate of TD is set as $1 - \gamma$.

A lending bank uses an S-shaped function, $p(S_{j,i}^{\text{total}})$, to decides probability of settling:

$$p(S_{j,i}^{\text{total}}(t)) = \frac{1}{1 + \alpha \times \exp(\beta \times S_{j,i}^{\text{total}}(t))} \quad (9)$$

where $p(S_{j,i}^{\text{total}})$ is the probability that i lends to j , and α and β are two parameters that control the intercept and slope, respectively. In this function, α is a positive real number. The larger the number, the lower probability of lending to a bank scoring 0. To represent different preferences of large banks and small banks, values are chosen from the uniform distribution $U(0.3, 0.5)$ for large banks and from the uniform distribution $U(0.9, 1.1)$ for small banks. This approach allows more lending from large banks to small banks. β is a negative

real number, and the larger it is the slower the probability moves from 0 to 1. In other words, a larger β means a tighter lending policy such that fewer borrowers get loans. Default values are chosen for banks from the uniform distribution $U(-1.1, -0.9)$. To build a dynamic system, we update assets of banks based on an empirically fitted distribution Beta (17.36, -0.1, 0.3).

4.3. Adaptive Learning Behaviors

In the section, we introduce the adaptive learning in banks' decisions in which they learn from experience to determine the percentage of reserve balance to lend out based on balance sheet information of other banks. We map the decision making process as a MDP problem and apply reinforcement learning algorithm to guide bank agents in determining the lending amounts.

4.3.1. RL Elements Definition in Interbank Lending Market

As bank agents interact with each other in interbank lending activities, potential lenders (with a surplus in reserve balance) must decide the amount of reserve balance to lend out based on the overall risk presented by the balance sheet conditions of other banks. Therefore, we define our RL elements as follows:

In each iteration, banks assess the balance sheet conditions of all other banks within the interbank lending market. Specifically in our problem, two types of balance sheet information are considered by bank agents: 1) lending and borrowing amounts from overnight, short-term, and long-term lending markets, 2) balance sheet ratios such as cash/equity and liabilities/equity.

We introduce the liquidity hoarding behavior in agent-based interbank lending market by allowing bank agents to determine the percentage of reserve balance to lend out. Empirical studies of the financial crisis have documented that the reserve balance significantly increased after the crisis as banks started to hoard liquidity due to uncertain conditions of the financial market (Afonso & Shin, 2009; Afonso et al., 2014). Under the RL setup, the action ranges from 0 to 1, denoting the percentage of reserve balance to lend out at each iteration. For example, if the amount of reserve balance is a million, then the final lending amount would be a million multiply by the percentage determined by the policy gradient method.

After each iteration, lenders receive full or partial repayments from their borrowers. Lenders will suffer losses when one of the borrowers defaults and reward will consequently decrease. On the other hand, if borrowers repay in full, reward will be positive since lenders got interests in addition to the lent amount. The choice of reward signal of bank agents is therefore naturally the difference between the lent amount and amount returned to the lenders. This setup of reward signal aligned with empirical studies (Cocco et al., 2009) where the authors showed that non-performing loans have a significant impact on the willingness of lenders fulfilling borrower request. Allen and Gale (2000) also claim that banks tend to form lending network to decrease their liquidity risk. Therefore, banks prefer to provide liquidity to the interbank lending market if lenders can

get their repayments in full.

4.3.2. Monte Carlo Policy Gradient

Monte Carlo policy gradient (MCPG) relies on an estimated return by Monte-Carlo methods to update the policy parameter θ . It works because the expectation of the sample gradient is equal to the actual gradient. In this study, we adopt a fixed rolling window approach for bank agents to update their policy. Banks reevaluate their policy over 30 episodes of learning using a fixed 4-year period in each quarter. The advantage of adopting this method is twofold: 1) In normal times, banks that need to borrow from the interbank market often settle their liquidity needs from large banks. As a result, smaller banks may not have enough experience evaluating incoming borrowing request during crisis when large banks begin to exhibit liquidity hoarding behavior. This approach allows smaller banks to evaluate the overall market condition more accurately; 2) By adopting a rolling window approach with MCPG, bank agents can react to the most recent event swiftly. This method enables bank agents to dynamically update and learn to assess the overall counterparty risks according to the latest balance sheets condition. Consequently, bank agents in the model evolve and change behaviors on how they react through learning from the past and display liquidity hoarding behavior after the crisis. Therefore we are able to measure $R_{s,a}$ from real sample episodes and use that to update our policy gradient. The process is as following:

- At the beginning of the process, banks have no knowledge about the banking system and behave randomly. We initialize the policy parameter θ at random.
- Simulate dynamic system based on banking behaviors and generate one episode on policy $\pi_\theta: s, a_1, r_2, s_2, a_2, \dots, s_T$. Each episode consists of the most recent 4-year period.
- Banks get feedback from the banking system for their behaviors and update their policy with respect to environment dynamics.
- At the end of each step, banks update their policy by updating policy parameters θ .
- With one episode complete, the banking system is reset to its initial condition, start a new episode and simulation continues. Banks have memory of the past episode, and they will use their knowledge gained from past episode in current episode. This step will repeat until banks reach a stable policy, as shown in Section 5.1.
- Banks then use the learned policy and decide the lending amount for the next quarter.

5. Model Validation

5.1. Convergence of Monte Carlo Policy Gradient

As a stochastic gradient method, MCPG has good theoretical convergence prop-

erties. By construction, the expected update over an episode is in the same direction as the performance gradient. We begin by showing the improvement and convergence of the accumulative reward using MCPG. In the model, lenders receive reward signal after their counterparties repay their loans, either in full or partially due to bank defaults. We collect all the reward signals in each step and run each episode for 4 years period prior to the financial crisis. At first, banks have no knowledge of the environment and lend out their surplus deposit randomly. Banks continue to update their policy as they receive repayments from counterparties during an episode. At the end of each episode, the banking system is restored to the initial condition and banks continuous to update their policies. As the number of episodes increases, banks gain more knowledge in assessing the overall risk condition presented in the balance sheet information of other banks and make better informed decision on the lending amount. We find that accumulative reward set in our reinforcement learning problem gradually increases and stabilizes within 30 episodes of training. **Figure 3** shows the average performance of 25 runs of simulation.

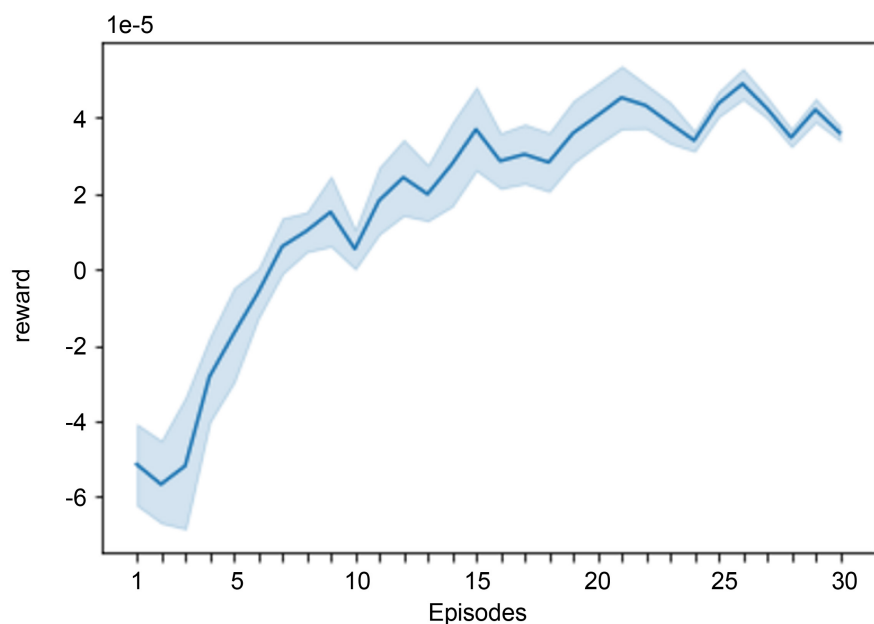


Figure 3. Accumulative reward for Monte Carlo policy gradient. Notes: The picture represents the accumulative reward for each iteration. From the plot of the accumulative reward, we observe that bank agents' learning reaches a plateau around 20 episodes and then stabilizes after 30 episodes. We also observe that 95% confidence interval of the accumulative reward becomes smaller after 25 episodes. It is also another indication of the convergence of the policy gradient training. Overall, we use 30 episodes as the cut-off for all agents training.

5.2. Interbank Network Topology Validation

We begin the model validation by comparing the interbank network topology between our proposed model and the U.S. federal fund market. Bech and Atalay

summarized the interbank markets topology using bank transaction data from Fedwire in 2006. They presented different measures of the interbank lending market using degree, clustering, and shortest path. Network topology of our simulated interbank lending markets is computed, and we find that the network structure closely resembles empirical findings. To validate the use of policy gradient reinforcement learning agents, we experiment with different time horizon that agents can use to maximize their cumulative rewards. With stopping time ranging from 2 years to 4 years, we find that the network topology remains relatively robust. **Table 2** presents network topology comparison with empirical interbank network.

Table 2. Network topology.

	Num. of Nodes	Clustering Coeff.	Shortest Path	Degree
The ABM Model	6600	0.32	3.20	14.80
U.S. Fed Funds Market	6600	0.53	2.55	15

6. The Subprime Mortgage Crisis

Empirical findings (Afonso & Shin, 2009; Afonso et al., 2014) found that market breakdown triggered liquidity hoarding during the subprime mortgage. In this section, we simulate the 2008-09 subprime mortgage crisis by triggering exogenous shock (Liu et al., 2017), that corresponds to a rapid drop of the U.S. housing price index, to the simulated model. As described in previous sections, the bank agents can adaptively learn to determine the lending amount based on their assessment of the overall counterparty risks presented in the balance sheet of other banks. This experiment examines how agents endogenously form liquidity hoarding behavior in the presence of increasing counterparty risk.

6.1. Simulation of Subprime Mortgage Crisis

We apply the exogenous shock to the model that correspond to the U.S. housing price index and compare the number of failed banks in the model with the empirical data. In the fast-changing environment during the financial crisis, bank agents need to decide lending amount with the converged action policy learned in normal times. **Figure 4** compares the number of bank failures in two simulated scenarios: 1) simulated model with adaptive learning agents, and 2) simulated model with non-adaptive learning agents with the empirical data.

We find that the number of failed banks from the model and the empirical data follow a similar trajectory. However, we observe two key differences between our simulated model and the empirical data: 1) Immediately after the crisis in 2009, the number of failed banks from our model is less than the number

of failed banks empirically; 2) There exist more failed banks from our simulated model than the empirical data 2 years after the crisis. We note that the rapid increase in the number of failed banks during 2010 to 2011 outpaces the increase of number of failed banks empirically. This can be attributed to the banks' ability to quickly react to the increasing counterparty risks. Recall that bank agents in the simulated model are only trained by MCPG to determine the lending amount during normal times. They need to adapt to the deteriorating environment by continuously exploring and receiving negative feedbacks through default loans. In reality, banks react and start hoarding liquidity following the news of major bankruptcy news occurred during the crisis from 2007-2009. While the liquidity hoarding behavior triggered more bank failures during 2009 to 2011, the banking system empirically had less overall bank failures in 2014. By comparing the two simulated scenarios between adaptive and non-adaptive banks, we demonstrate the ability of reinforcement learning agents to adapt from the crisis and reduce counterparty risks.

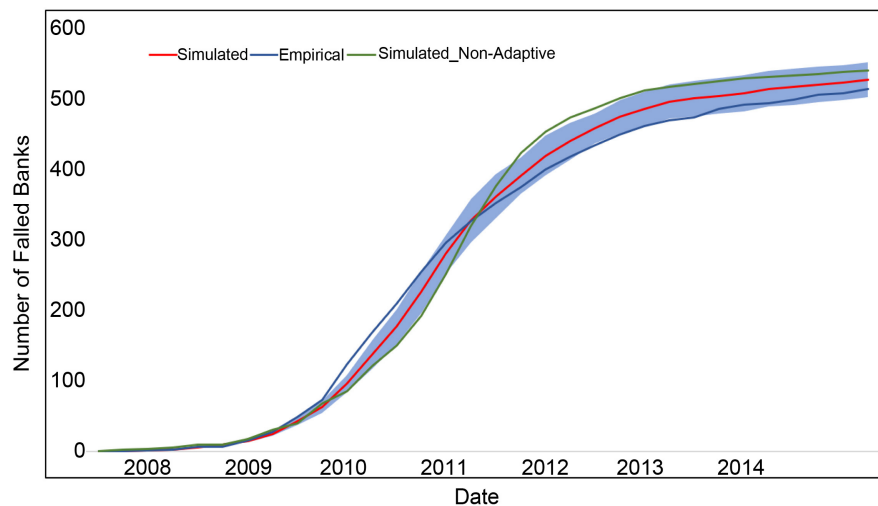


Figure 4. Failed banks in the 2007-09 financial crisis. Notes: This figure simulates banking system during the 2007-09 financial crisis. The red line represents the cumulative number of bank defaults from 2007 Q2 to 2014 Q3. The green line represents the number of bank defaults in the model. The experiment runs for 20 iterations. The shaded area of the green line shows the standard error of the mean. Source: Federal Financial Institutions Examination Council Reports of Condition and Income, authors' model.

6.2. Liquidity Hoarding

Interbank markets allow liquidity to be transferred from banks with a surplus to banks with a deficit. Under normal circumstances the interbank markets work rather well (Allen et al., 2009). However, the functioning of interbank markets has been impaired since August 2007. As the financial crisis deepened, liquidity in the interbank market was further dried up as banks preferred to hoard liquidity in anticipation of deteriorating market condition (Heider et al., 2015). We observe that agents, according to the learned lending policy, exhibit liquidity hoard-

ing behavior after the shock is triggered. In **Figure 5**, we observe that the lending percentage of the simulated banks resembles the decline in lending percentage of the empirical lending percentage during the financial crisis from 2007 to 2011. This demonstrates that balance sheets provide counterparty risk information and adaptive bank can learn from it.

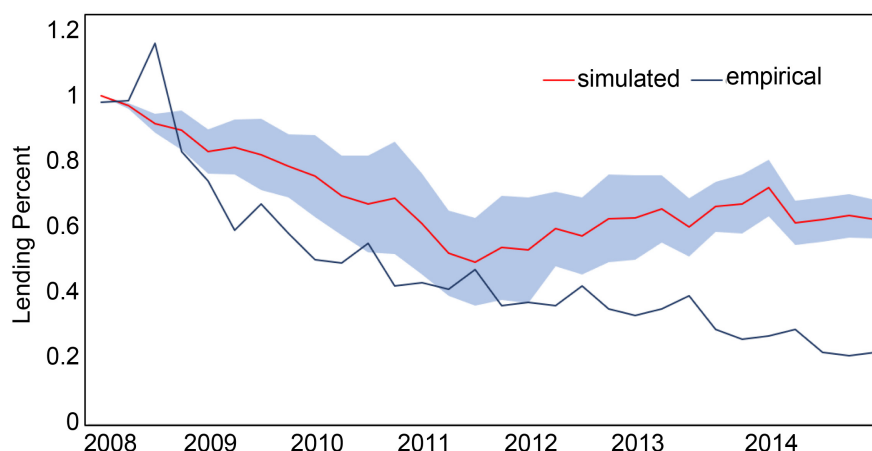


Figure 5. Liquidity hoarding. Notes: Empirical relative lending percentage is the ratio of lending amount over lending amount at the beginning of financial crisis. This figure displays the relative lending percentage for 6600 banks. The result is averaged for 20 runs, grey area shows standard error of the mean.

At the end of the financial crisis after 2011, the empirical lending percentage continued to decline while the simulated banks begin to lend more. This phenomena can be explained by the regulatory reforms to the financial sector. The Basel III framework was first issued in 2010 and it is aimed to provide a foundation for a more resilient banking system. Some notable measures from the framework included: 1) significantly increases the required amount of capital held by banks, 2) emphasizes the need for a capital buffer, and 3) specifies a lower limit for the leverage ratio to discourage excess leverage (*Bank for International Settlements, 2011a, 2011b*). In addition, we can also observe that the major write down losses from the housing price index started to rebound after 2011 as shown in **Figure A1** in **Appendix**. Consequently, the banking environment became stable and bank continued to hold more capital after 2011.

7. Effect of Fire Sales

Fire sales contributed to the deterioration of the financial system by exacerbating the distressed balance sheets of banks during the subprime crisis. Such effect has attracted more attention from researchers and a number of theoretical models have been proposed to study such effect (*Cifuentes et al., 2005; Coval & Stafford, 2007*). In this section, from empirical perspective, we examine the role of liquidity hoarding and fire sales in determining contagion and aggregate losses in the banking system using ABM model.

7.1. Simulation of Fire Sales and Systemic Risk

Fire sales happen when banks lack liquidity. The asset liquidation will temporarily lower market values and impact bank asset sizes, which will further amplify losses. Similarly, we incorporate fire sales in the model and study how fire sales impact bank behaviors and interbank lending market. We follow the methods in Zhang et al. (2018) and fire sales occur when a bank faces liquidity risk or violates capital constraint. The minimum capital ratio is defined as equity to risk weighted assets. We adopt the most recent capital requirement, which is 10.5% from 2019 under Basel III. The selling of securities in the market can affect the price of securities, causing other banks holding same securities suffer financial shocks Cifuentes et al. (2005), see Equation (10):

$$p_u = p_u^0 * \exp\left(-\sum_i^N \frac{sell_u^i}{\sum_i^N \alpha_u s_u^i}\right) \tag{10}$$

where α_u indicates the market depth of the security, $sell_u^i$ represents the amount of security u sold by bank i , and s_u^i represents the total amount of security u in the banking system.

The Eisenberg Noe clearing system is modified by incorporating fire sales (Zhang et al., 2018) (see Equation (10)). Let the relative obligation matrix of bank i to bank j be:

$$\Pi_{ij} = (P_{ij}^{ON} + P_{ij}^{ST} + P_{ij}^{LT}) / P_i \tag{11}$$

where P_i represents total loan payment bank i gets, the clearing vector P^* is:

$$P^* = \min[l; \Pi^T * P^* + C + Z] \tag{12}$$

where vector l represents the total obligations of the banks, vector C represents cash holdings, vector Z represents the mark-to-market value of liquid and illiquid securities.

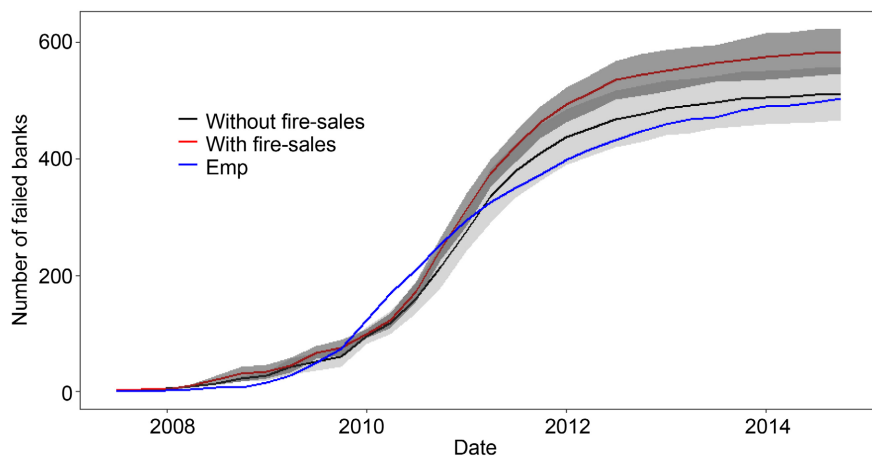


Figure 6. Bank failures comparison. Notes: This figure shows accumulative bank defaults for a scaled down version of the US banking system (with 6600 banks) from Jun 2007 to Dec 2012. The red line represents the number of defaults for model with fire sales incorporated. The blue line represents the number of defaults for model without fire sales.

In **Figure 6**, we compare the bank defaults for two models with and without fire sales. We find that the model with fire sales have more bank defaults, demonstrating that fire sales can contribute to systemic risk. This finding is consistent with the result in *Zhang et al. (2018)*.

7.2. Impact of Fire Sales on Liquidity Hoarding

We consider the effect of fire sales due to the adaptive learning behaviors. When fire sales happen, market price drop will be reflected in the individual banks' reward functions. Then we allow banks relearn the optimal policy. **Figure 7** and table in **Appendix** show how the relative lending percentage evolves over time after the exogenous shock. With fire sales, we observe that banks hold more liquidity when a shock comes. Our result is consistent with *Diamond and Rajan (2009)*, which shows that the fire sales of risky banks can drive healthy banks to hoard liquid funds.

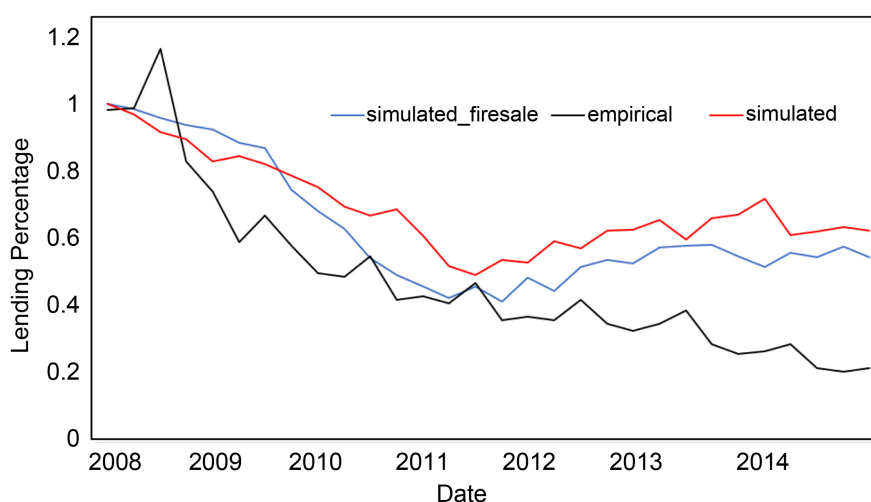


Figure 7. Relative lending percentage. Notes: This figure compares the empirical relative lending percentage with the averaged relative lending percentage in the simulated interbank lending markets with and without fire sale. The simulated result is averaged for 20 runs.

7.3. Impact of Fire Sales on Contagion Risk

To quantify contagion risk, we use the contagion index (CI) in *Glasserman and Young (2015)* to measure the potential impact of bank failures on the rest of the banking system. It can be regarded as dollar amount of a bank's borrowings in interbank markets.

$$CI_i = IL_i * \left(1 - \frac{IA_i}{L_i}\right) \quad (13)$$

where IL and IA respectively represent interbank liability and interbank asset from overnight, short-term, and long-term markets. **Figure 8** shows the evolution of the average contagion index for three different scenarios. By observing

the average contagion index of non-adaptive agents (grey line) and adaptive agents (blue line) without fire sales incorporated, it shows that contagion index with adaptive agents is lower due to the liquidity hoarding behaviors. It can be interpreted by the network structure changes due to reduced interbank connections. When the interbank market gets less connected, the probability of a failed node causing others to fail is also reduced as reflected in the decreasing contagion index. Despite the observation from **Figure 7** where fire sales would cause adaptive bank agents to hoard more liquidity, **Figure 8** shows that fire sales lead to a higher contagion index comparing the averaged contagion index with and without fire sales (red and blue line respectively) using adaptive agents. This indicates that fire sales lead to a higher probability of contagion through interbank lending networks.

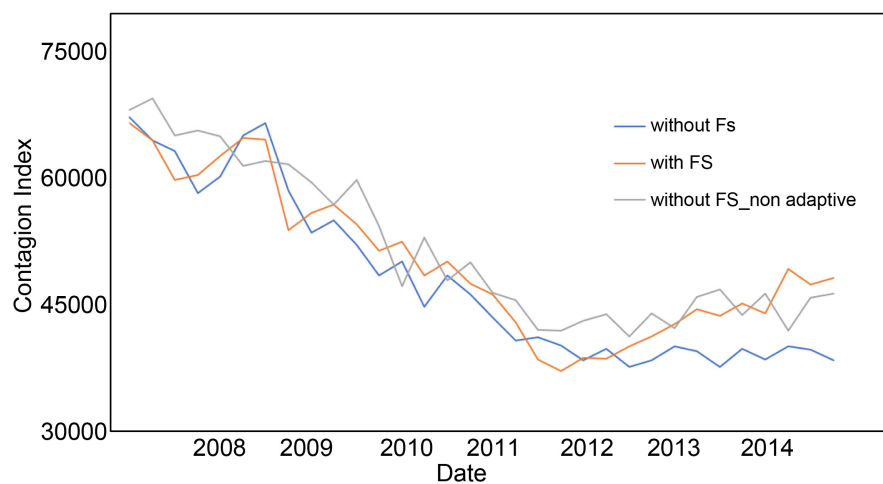


Figure 8. Contagion index comparison. Notes: This figure displays contagion index in interbank lending markets. The grey line shows the averaged contagion index using non-adaptive bank agents without incorporating fire sales. The other two lines compare the contagion index before and after incorporating fire sales with adaptive agents. The red line represents the averaged contagion index when fire sales are incorporated. The blue line represents averaged contagion index without fire sales incorporated. The result is averaged for 20 runs.

8. Conclusion

This paper presents a heterogeneous agent model with adaptive learning agents to model liquidity hoarding in the U.S. interbank lending markets. We allow banks to have information about balance sheets in the banking system. We show that individual counterparty lending decisions can affect overall funding liquidity in the interbank market. We evaluate the model from network perspective and compare the simulation results with the empirical bank failures in the 2008-09 financial crisis. The model illustrates the effects of the banks' adaptive learning behaviors through analyzing interbank liquidity and contagion risk.

Overall, we find that adaptive learning banks tend to hoard liquidity to maximize their total utility under a distressed market condition. Such liquidity-hoarding

behaviors create naturally a contagion-dampening effect from the network perspective to reduce the impact of the failed banks. In a multi-layered network setting, the learning agent model generates a more sparse network structure compared with a similar non-adaptive agent model (Liu et al., 2017), and the model traces the empirical bank failure pattern better than the comparison model.

Another major finding of the paper is that fire sales will worsen the liquidity hoarding problem under a distressed market condition. Through examining the average percentage of lending and the contagion index (Glasserman & Young, 2015), we find that fire sales would lead to a decrease in interbank liquidity and increase the probability of contagion through interbank lending networks.

In summary, we propose a learning agent model that has several advantages over the other existing models, such as it is not equilibrium-based, agents are adaptive in nature, and it can be calibrated with real bank and market data. More importantly, the agents' behaviors are calibrated based on the real bank and market data, and thus a macro-prudential-based stress test can be easily performed with many perceivable scenarios. Such modeling paradigm can also be used to evaluate the effectiveness of new regulations.

Acknowledgements

The authors would like to thank Blake LeBaron, Nathan Palmer, German Creamer, Jason Barr, and other attendees at the 45th Eastern Economic Association Annual Meeting in 2019 for their valuable comments. Additionally, the authors would like to thank Victor Luo, Anand Goel, Pape Ndyie, Hamed Ghoddusi, and others for their valuable feedback at Stevens' Finance and Financial Engineering seminar.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Acharya, V. V. (2009). A Theory of Systemic Risk and Design of Prudential Bank Regulation. *Journal of Financial Stability*, 5, 224-255. <https://doi.org/10.1016/j.jfs.2009.02.001>
- Acharya, V. V., & Skeie, D. (2011). A Model of Liquidity Hoarding and Term Premia in Inter-Bank Markets. *Journal of Monetary Economics*, 58, 436-447. <https://doi.org/10.1016/j.jmoneco.2011.05.006>
- Afonso, G., & Shin, H. S. (2009). *Systemic Risk and Liquidity in Payment Systems*. FRB of New York Staff Report.
- Afonso, G., Kovner, A., & Schoar, A. (2014). *Trading Partners in the Interbank Lending Market*. FRB of New York Staff Report.
- Aldasoro, I., Gatti, D. D., & Faia, E. (2017). Bank Networks: Contagion, Systemic Risk and Prudential Policy. *Journal of Economic Behavior & Organization*, 142, 164-188. <https://doi.org/10.1016/j.jebo.2017.05.022>
- Allen, F., & Carletti, E. (2008). *The Role of Liquidity in Financial Crises*. In *Jackson Hole*

Economic Policy Symposium (No. 48).

- Allen, F., & Gale, D. (2000). Financial Contagion. *The Journal of Political Economy*, 108, 1-33. <https://doi.org/10.1086/262109>
- Allen, F., Carletti, E., & Gale, D. (2009). Interbank Market Liquidity and Central Bank Intervention. *Journal of Monetary Economics*, 56, 639-652. <https://doi.org/10.1016/j.jmoneco.2009.04.003>
- Amman, H. M., Tesfatsion, L., Kendrick, D. A., Judd, K. L., & Rust, J. (1996). *Handbook of Computational Economics, Volume 2*. Elsevier.
- Arthur, W. B. (1991). Designing Economic Agents That Act Like Human Agents: A Behavioral Approach to Bounded Rationality. *The American Economic Review*, 81, 353-359.
- Arulkumaran, K., Deisenroth, M. P., Brundage, M., & Bharath, A. A. (2017). Deep Reinforcement Learning: A Brief Survey. *IEEE Signal Processing Magazine*, 34, 26-38. <https://doi.org/10.1109/MSP.2017.2743240>
- Axtell, R., Axelrod, R., Epstein, J. M., & Cohen, M. D. (1996). Aligning Simulation Models: A Case Study and Results. *Computational & Mathematical Organization Theory*, 1, 123-141. <https://doi.org/10.1007/BF01299065>
- Ballester, C., Calvó-Armengol, A., & Zenou, Y. (2006). Who's Who in Networks. Wanted: The key Player. *Econometrica*, 74, 1403-1417. <https://doi.org/10.1111/j.1468-0262.2006.00709.x>
- Bank for International Settlements (2011a). *Basel III: A Global Regulatory Framework for More Resilient Banks and Banking System*. <https://www.bis.org/publ/bcbs189.htm>
- Bank for International Settlements (2011b). *Global Systemically Important Banks: Assessment Methodology and the Additional Loss Absorbency Requirement*. <https://www.bis.org/publ/bcbs201.htm>
- Bargigli, L., Di Iasio, G., Infante, L., Lillo, F., & Pierobon, F. (2015). The Multiplex Structure of Interbank Networks. *Quantitative Finance*, 15, 673-691. <https://doi.org/10.1080/14697688.2014.968356>
- Berardi, M., & Galimberti, J. K. (2017). Empirical Calibration of Adaptive Learning. *Journal of Economic Behavior & Organization*, 144, 219-237. <https://doi.org/10.1016/j.jebo.2017.10.004>
- Bisias, D., Flood, M., Lo, A. W., & Valavanis, S. (2012). A Survey of Systemic Risk Analytics. *Annual Review of Financial Economics*, 4, 255-296. <https://doi.org/10.1146/annurev-financial-110311-101754>
- Bossan, B., Jann, O., & Hammerstein, P. (2015). The Evolution of Social Learning and Its Economic Consequences. *Journal of Economic Behavior & Organization*, 112, 266-288. <https://doi.org/10.1016/j.jebo.2015.01.010>
- Boyson, N. M., Stahel, C. W., & Stulz, R. M. (2010). Hedge Fund Contagion and Liquidity Shocks. *The Journal of Finance*, 65, 1789-1816. <https://doi.org/10.1111/j.1540-6261.2010.01594.x>
- Brandi, G., Di Clemente, R., & Cimini, G. (2018). Epidemics of Liquidity Shortages in Interbank Markets. *Physica A: Statistical Mechanics and its Applications*, 507, 255-267. <https://doi.org/10.1016/j.physa.2018.05.104>
- Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E., & Havlin, S. (2010). Catastrophic Cascade of Failures in Interdependent Networks. *Nature*, 464, 1025-1028. <https://doi.org/10.1038/nature08932>
- Bush, R. R., & Mosteller, F. (1955). *Stochastic Models for Learning*. John Wiley & Sons. <https://doi.org/10.1037/14496-000>

- Chou, P. W., Maturana, D., & Scherer, S. (2017). Improving Stochastic Policy Gradients in Continuous Control with Deep Reinforcement Learning Using the β Distribution. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70* (pp. 834-843). PMLR.
- Cifuentes, R., Ferrucci, G., & Shin, H. S. (2005). Liquidity Risk and Contagion. *Journal of the European Economic Association*, 3, 556-566.
<https://doi.org/10.1162/jeea.2005.3.2-3.556>
- Cocco, J., Gomes, F., & Martins, N. (2009). Lending Relationships in the Interbank Market. *Journal of Financial Intermediation*, 18, 24-48.
<https://doi.org/10.1016/j.jfi.2008.06.003>
- Confidence, R. (2012). *Global Financial Stability Report: Restoring Confidence and Progressing on Reforms*. <https://www.imf.org/external/pubs/ft/gfsr/2012/02/>
- Coval, J., & Stafford, E. (2007). Asset Fire Sales (and Purchases) in Equity Markets. *Journal of Financial Economics*, 86, 479-512. <https://doi.org/10.1016/j.jfineco.2006.09.007>
- Denbee, E., Julliard, C., Li, Y., & Yuan, K. (2021). Network Risk and Key Players: A Structural Analysis of Interbank Liquidity. *Journal of Financial Economics*, 141, 831-859.
<https://doi.org/10.1016/j.jfineco.2021.05.010>
- Diamond, D. W., & Rajan, R. G. (2009). The Credit Crisis: Conjectures about Causes and Remedies. *American Economic Review*, 99, 606-610.
<https://doi.org/10.1257/aer.99.2.606>
- Duarte, F., & Eisenbach, T. M. (2021). Fire-Sale Spillovers and Systemic Risk. *The Journal of Finance*, 76, 1251-1294. <https://doi.org/10.1111/jofi.13010>
- Duffy, J., & Feltovich, N. (1999). Does Observation of Others Affect Learning in Strategic Environments? An Experimental Study. *International Journal of Game Theory*, 28, 131-152. <https://doi.org/10.1007/s001820050102>
- Eisenberg, L., & Noe, T. (2001). Systemic Risk in Financial Systems. *Management Science*, 47, 236-249. <https://doi.org/10.1287/mnsc.47.2.236.9835>
- Erev, I., & Roth, A. E. (1998). Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria. *American Economic Review*, 88, 848-881.
- Eross, A., Urquhart, A., & Wolfe, S. (2016). Liquidity Risk Contagion in the Interbank Market. *Journal of International Financial Markets, Institutions and Money*, 45, 142-155.
<https://doi.org/10.1016/j.intfin.2016.07.005>
- Gai, P., & Kapadia, S. (2010). Liquidity Hoarding, Network Externalities, and Interbank Market Collapse. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 466, 439. <https://doi.org/10.1098/rspa.2009.0410>
- Gilbert, N., & Terna, P. (2000). How to Build and Use Agent-Based Models in Social Science. *Mind & Society*, 1, 57-72. <https://doi.org/10.1007/BF02512229>
- Glasserman, P., & Young, H. P. (2015). How Likely Is Contagion in Financial Networks? *Journal of Banking & Finance*, 50, 383-399.
<https://doi.org/10.1016/j.jbankfin.2014.02.006>
- Gofman, M. (2017). Efficiency and Stability of a Financial Architecture with Too-Interconnected-To-Fail Institutions. *Journal of Financial Economics*, 124, 113-146.
<https://doi.org/10.1016/j.jfineco.2016.12.009>
- Heider, F., Hoerova, M., & Holthausen, C. (2015). Liquidity Hoarding and Interbank Market Spreads: The Role of Counterparty Risk. *Journal of Financial Economics*, 118, 336-354.
<https://doi.org/10.1016/j.jfineco.2015.07.002>

- Herskovic, B., Kelly, B., Lustig, H., & Van Nieuwerburgh, S. (2020). Firm Volatility in Granular Networks. *Journal of Political Economy*, *128*, 4097-4162.
<https://doi.org/10.1086/710345>
- Hopkins, E. (2007). Adaptive Learning Models of Consumer Behavior. *Journal of Economic Behavior & Organization*, *64*, 348-368. <https://doi.org/10.1016/j.jebo.2006.02.010>
- Hu, W. (2021). Reinforcement Learning of Molecule Optimization with Bayesian Neural Networks. *Computational Molecular Bioscience*, *11*, 69-83.
<https://doi.org/10.4236/cmb.2021.114005>
- Keister, T., & McAndrews, J. (2009). Why Are Banks Holding So Many Excess Reserves? *Current Issues in Economics and Finance*, *15*, 1-10.
- Li, Z., & Lai, R. N. (2022). Interbank Borrowing and Bank Liquidity Risk. *Journal of Financial Research*, *45*, 53-91.
- Liu, A., Mo, C., Paddrik, M., & Yang, S. (2018). An Agent-Based Approach to Interbank Market Lending Decisions and Risk Implications. *Information*, *9*, Article 132.
<https://doi.org/10.3390/info9060132>
- Liu, A., Paddrik, M., Yang, S. Y., & Zhang, X. (2017). Interbank Contagion: An Agent-Based Model Approach to Endogenously Formed Networks. *Journal of Banking & Finance*, *112*, Article ID: 105191.
- Lux, T. (2015). Emergence of a Core-Periphery Structure in a Simple Dynamic Model of the Interbank Market. *Journal of Economic Dynamics and Control*, *52*, A11-A23.
<https://doi.org/10.1016/j.jedc.2014.09.038>
- Macy, M. W., & Willer, R. (2002). From Factors to Factors: Computational Sociology and Agent-Based Modeling. *Annual Review of Sociology*, *28*, 143-166.
<https://doi.org/10.1146/annurev.soc.28.110601.141117>
- Montagna, M., & Kok, C. (2016). *Multi-Layered Interbank Model for Assessing Systemic Risk*. Kiel Working Paper No. 1873.
- Nicolaisen, J., Petrov, V., & Tesfatsion, L. (2001). Market Power and Efficiency in a Computational Electricity Market with Discriminatory Double-Auction Pricing. *IEEE transactions on Evolutionary Computation*, *5*, 504-523.
<https://doi.org/10.1109/4235.956714>
- Pemantle, R., & Skyrms, B. (2004). Time to Absorption in Discounted Reinforcement Models. *Stochastic Processes and Their Applications*, *109*, 1-12.
<https://doi.org/10.1016/j.spa.2003.08.003>
- Roukny, T., Battiston, S., & Stiglitz, J. E. (2018). Interconnectedness as a Source of Uncertainty in Systemic Risk. *Journal of Financial Stability*, *35*, 93-106.
<https://doi.org/10.1016/j.jfs.2016.12.003>
- Sarin, R., Vahid, F., et al. (2001). Predicting How People Play Games: A Simple Dynamic Model of Choice. *Games and Economic Behavior*, *34*, 104-122.
<https://doi.org/10.1006/game.1999.0783>
- Stiglitz, J. E. (2010). Risk and Global Economic Architecture: Why Full Financial Integration May Be Undesirable. *American Economic Review*, *100*, 388-392.
<https://doi.org/10.1257/aer.100.2.388>
- Sutton, R. S., McAllester, D. A., Singh, S. P., & Mansour, Y. (2000). Policy Gradient Methods for Reinforcement Learning with Function Approximation. In M. I. Jordan, Y. Lecun, & S. A. Solla (Eds.), *Advances in Neural Information Processing Systems* (pp. 1057-1063). MIT Press.
- Watkins, C. J., & Dayan, P. (1992). Q-Learning. *Machine Learning*, *8*, 279-292.
<https://doi.org/10.1023/A:1022676722315>

Yellen, J. (2013). *Interconnectedness and Systemic Risk: Lessons from the Financial Crisis and Policy Implications: A Speech at the American Economic Association/American Finance Association Joint Luncheon, San Diego, California, January 4, 2013*. Technical Report, Board of Governors of the Federal Reserve System (US).

Zhang, X., Mo, C. Y., & Yang, S. Y. (2018). Bank Contagion Risk through Fire-Sales: A Heterogeneous Agent Model. In *2018 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 2321-2328). IEEE. <https://doi.org/10.1109/SSCI.2018.8628902>

Appendix

Housing Price Index

We collect housing price index from Federal Housing Finance Agency from Jun. 2007 to Dec. 2014 and get changes of index for each quarter. We divide total other assets OA of each bank into real estate OA_R and non-real estate assets OA_{NR} based on empirical ratio of real estate loans to other assets. At each step, banks write down losses based on changes of HPI , ΔHPI , following Equation (14):

$$OA(t) = \Delta HPI(t) \cdot OA_R(t-1) + OA_{NR} \tag{14}$$

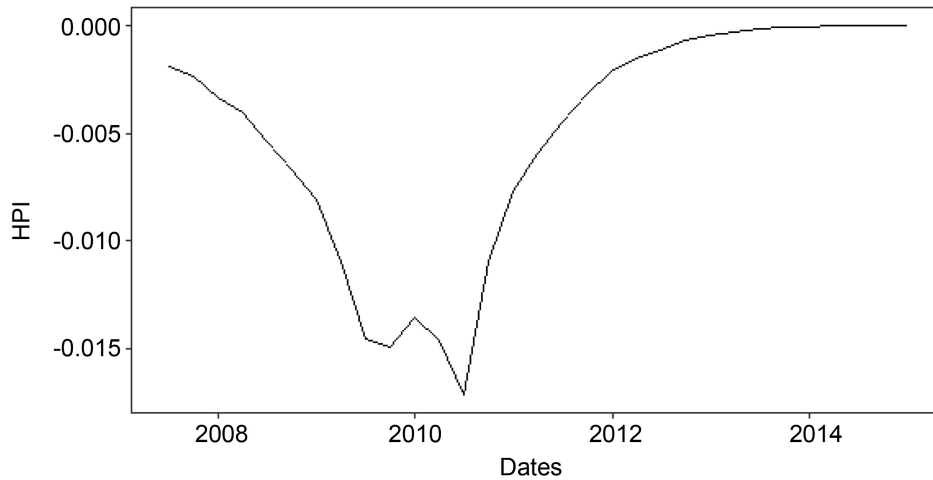


Figure A1. Housing price index from federal housing finance agency from Jun 2007 to Dec 2014.

Relative Lending Percentage

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
6/30/2007	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9/30/2007	0.951	0.989	0.956	0.978	0.970	0.960	0.972	0.971	0.952	0.950	0.987	0.959	0.983	0.990	0.982	0.993	0.993	0.953	0.988	0.950
12/31/2007	0.954	0.976	0.969	0.875	0.878	0.963	0.946	0.986	0.872	0.877	0.916	0.921	0.887	0.969	0.854	0.870	0.842	0.922	0.972	0.955
3/31/2008	0.980	0.834	0.859	1.003	0.848	0.972	0.880	0.898	0.981	0.990	1.007	0.949	0.792	0.858	0.785	0.881	0.833	0.950	0.904	0.809
6/30/2008	0.905	0.892	0.878	0.736	0.906	0.947	0.751	0.943	0.947	0.867	0.820	0.922	0.707	0.740	0.697	0.731	0.891	0.883	0.787	0.857
9/30/2008	0.677	0.816	0.950	0.794	0.951	0.780	0.844	0.971	0.971	0.846	0.713	0.872	0.804	0.723	0.762	0.976	0.871	0.728	0.905	0.767
12/31/2008	0.866	0.707	0.944	0.822	0.755	0.725	0.863	0.713	0.685	0.681	0.701	0.827	0.995	0.902	0.968	0.943	0.935	0.684	1.007	0.712
3/31/2009	0.647	0.611	0.786	0.820	0.939	0.813	0.902	0.734	0.738	0.915	0.874	0.755	0.808	0.909	0.702	0.666	0.795	0.876	0.719	0.847
6/30/2009	0.866	0.660	0.701	0.842	0.626	0.887	0.640	0.851	0.869	0.820	0.497	0.821	0.742	0.517	0.762	0.693	0.869	0.869	0.716	0.871
9/30/2009	0.526	0.621	0.713	0.718	0.757	0.597	0.785	0.704	0.690	0.787	0.714	0.700	0.616	0.798	0.687	0.838	0.786	0.872	0.463	0.495
12/31/2009	0.494	0.833	0.501	0.881	0.442	0.679	0.712	0.794	0.678	0.619	0.725	0.918	0.496	0.672	0.904	0.669	0.575	0.653	0.817	0.499
3/31/2010	0.659	0.923	0.767	0.800	0.562	0.826	0.923	0.473	0.473	0.719	0.434	0.908	0.726	0.861	0.474	0.410	0.786	0.722	0.751	0.934
6/30/2010	0.515	0.628	0.852	0.737	0.610	0.402	0.562	0.799	0.557	0.428	0.417	0.431	0.800	0.652	0.736	0.503	0.438	0.673	0.798	0.741
9/30/2010	0.610	0.787	0.503	0.590	0.396	0.333	0.400	0.441	0.537	0.506	0.566	0.438	0.636	0.529	0.599	0.326	0.611	0.779	0.651	0.529
12/31/2010	0.389	0.420	0.529	0.712	0.770	0.338	0.460	0.350	0.382	0.634	0.467	0.461	0.426	0.536	0.695	0.400	0.393	0.497	0.395	0.759

Continued

3/31/2011	0.793	0.552	0.348	0.470	0.708	0.436	0.808	0.678	0.618	0.441	0.452	0.687	0.373	0.781	0.376	0.696	0.417	0.643	0.555	0.367
6/30/2011	0.359	0.349	0.387	0.443	0.696	0.367	0.365	0.528	0.647	0.721	0.780	0.465	0.647	0.459	0.466	0.555	0.462	0.719	0.853	0.395
9/30/2011	0.678	0.797	0.811	0.682	0.696	0.637	0.624	0.431	0.536	0.366	0.344	0.529	0.545	0.524	0.767	0.748	0.530	0.456	0.420	0.633
12/31/2011	0.432	0.696	0.740	0.806	0.645	0.456	0.393	0.742	0.671	0.587	0.485	0.459	0.462	0.652	0.395	0.360	0.639	0.767	0.495	0.408
3/31/2012	0.638	0.654	0.436	0.579	0.540	0.378	0.863	0.904	0.472	0.618	0.851	0.772	0.452	0.580	0.410	0.822	0.634	0.688	0.430	0.888
6/30/2012	0.620	0.640	0.634	0.441	0.377	0.519	0.862	0.629	0.466	0.666	0.756	0.528	0.809	0.508	0.449	0.690	0.632	0.532	0.794	0.401
9/30/2012	0.791	0.567	0.822	0.830	0.468	0.426	0.786	0.875	0.467	0.620	0.828	0.774	0.658	0.566	0.498	0.430	0.535	0.853	0.608	0.571
12/31/2012	0.418	0.533	0.755	0.701	0.587	0.839	0.593	0.547	0.433	0.789	0.385	0.401	0.577	0.631	0.490	0.634	0.656	0.640	0.544	0.518
3/31/2013	0.719	0.598	0.748	0.715	0.609	0.628	0.638	0.817	0.875	0.534	0.562	0.782	0.615	0.759	0.569	0.720	0.606	0.848	0.474	0.630
6/30/2013	0.449	0.679	0.520	0.799	0.877	0.781	0.834	0.836	0.703	0.875	0.803	0.641	0.454	0.565	0.460	0.734	0.599	0.451	0.743	0.677
9/30/2013	0.825	0.670	0.539	0.481	0.867	0.514	0.686	0.484	0.893	0.913	0.859	0.520	0.655	0.712	0.612	0.834	0.836	0.910	0.825	0.795
12/31/2013	0.613	0.648	0.668	0.572	0.730	0.539	0.451	0.788	0.440	0.657	0.799	0.516	0.530	0.665	0.470	0.545	0.654	0.816	0.462	0.811
3/31/2014	0.458	0.508	0.491	0.547	0.800	0.737	0.606	0.545	0.487	0.702	0.830	0.481	0.609	0.567	0.447	0.805	0.654	0.540	0.575	0.767
6/30/2014	0.527	0.471	0.516	0.728	0.790	0.740	0.651	0.847	0.537	0.670	0.560	0.629	0.699	0.778	0.483	0.739	0.590	0.449	0.843	0.430
9/30/2014	0.687	0.772	0.439	0.720	0.591	0.800	0.807	0.677	0.421	0.497	0.679	0.521	0.674	0.723	0.427	0.614	0.703	0.506	0.762	0.690
