AN AGENT-BASED MODEL OF REGIONAL FOOD SUPPLY CHAIN DISINTERMEDIATION

Teri J. Craven
Department of Industrial and Manufacturing Systems Engineering
Iowa State University
3004 Black Engineering
Ames, IA, USA
tjcraven@iastate.edu

Caroline C. Krejci
Department of Industrial and Manufacturing Systems Engineering
Iowa State University
3004 Black Engineering
Ames, IA, USA
ckrejci@iastate.edu

ABSTRACT

Regional food hubs provide logistics services for small and mid-sized food producers, giving them the ability to reach larger markets and customers than they can reach on their own. However, once food hub managers have helped to establish connections between producers and new customers, they often find themselves cut out of the regional food supply chain when the farmers decide to sell their products directly to the customers, thereby avoiding the food hub's service fees. Widespread disintermediation can eventually lead to food hub failure, which can disrupt the entire regional food system. This paper describes an agent-based model that incorporates reinforcement learning to study disintermediation behavior in a regional food supply network in Iowa. The model is designed to serve as a decision support tool for food hub managers, allowing them to simulate the effects of various supply chain management strategies on producer decision making and long-term system success.

Keywords: agent-based modeling, reinforcement learning, regional food hubs, supply chain management.

1 INTRODUCTION

Consumers are increasingly seeking fresh and healthy food that has been sustainably produced in the regions in which they live (Jones, Comfort, and Hillier 2004). However, many consumers also value convenience and efficiency and prefer to purchase food from retailers and restaurants, rather than farmers' markets. Regional food hubs offer a potential solution to this problem. A regional food hub acts as an intermediary in a regional food supply chain by aggregating, marketing, and distributing food from small and mid-sized producers to customers that are located in the same region, facilitating a connection between them (Barham et al. 2012). This allows the producers to focus on food production, rather than logistics and marketing, and enables them to reach larger markets and customers (e.g., grocery stores) than they could reach on their own.

However, once food hub managers have helped to establish connections between producers and new customers, they often find themselves cut out of the regional food supply chain when the farmers decide to sell their products directly to the customers, thereby avoiding the food hub's service fees (McCann and Crum 2015). This phenomenon, known as disintermediation, is the removal of an intermediary in a supply chain in return for lower costs for the customers and suppliers (Mills and Camek 2004). While this can have short-term financial benefits for the producers, widespread disintermediation can eventually lead to food hub failure, which can disrupt the entire regional food system. To avoid this, food hub managers must develop and implement policies that will support long-term and mutually beneficial relationships with their producers and customers.

SpringSim-ADS 2017, April 23-26, Virginia Beach, VA, USA

One particular regional food hub in western Iowa provides transportation and marketing services for small and mid-sized regional food producers across Iowa and eastern Nebraska. These producers list their products on the hub's website, where customers order the products that they want to purchase through the hub. At the end of each weekly order cycle, food hub personnel pick up products from the producers' farms and deliver them to the customers. This food hub primarily sells to wholesale customers (e.g., restaurants and grocery stores) that are concentrated in the metropolitan areas that encompass Omaha and Des Moines. The food hub strives to help small and mid-sized producers reach new markets, as well as helping customers meet their demand for regionally-produced food. The food hub's services allow some producers to sell products to customers that are distant from their farms.

Despite its mission to be a supportive supply chain partner to small and mid-sized producers, the food hub manager has noticed that some of its producers have begun selling products directly to customers, thereby cutting out the hub. These same producers have continued to use the food hub's services to sell products to other customers. The food hub manager believes the producers are initiating transactions around the hub to the food hub customers. The manager is extremely concerned that if more producers choose to sell products around the hub, the hub will no longer be financially viable. If the food hub fails, all of the hub's producers will be forced to either deliver their products via their own transportation, make separate transportation arrangements, or lose customers. The manager has attempted to curb this behavior by asking new producers to sign an agreement in which they promise not to sell around the hub. However, she is unsure it of its effectiveness and does not know how to prevent this behavior in the long term.

The purpose this paper is to demonstrate how an agent-based model (ABM) can be used as a tool for food hub managers to evaluate the effectiveness of various management strategies in preventing producers from selling around the hub. Experimentation with an ABM can enable a food hub manager to understand the impact of a management policy without assuming any of the risks of actual implementation. In addition to studying the effects of management strategies, in this paper, modeling the case study supply chain using ABM also allows the modeler to examine whether the food hub is necessary for small and mid-sized producers to be successful. If the food hub is not necessary for those producers to be successful, then it might not be a viable business entity in its current form.

ABM is a modeling tool that is well-suited to capturing the complexity of supply chains, particularly the behaviors, decisions, and interactions of the autonomous, intelligent, and interconnected actors (e.g., food producers, food hub managers, and customers) that inhabit them (Choi, Dooley, and Rungtusanatham 2001; Meter 2006). Unlike traditional optimization models, ABM can accommodate heterogeneous actor objectives, behaviors, dynamic interactions, and adaptations that occur among humans over time. Some agents may be more influential than others, but none completely controls the behavior of the system (North and Macal 2007). The dynamic interactions among individual agents and the objects in their environment result in a system that exhibits behavior that cannot be predicted by examining the behavior of its individual parts (Pathak et al. 2007). Such system behavior and resultant properties are said to be emergent (Gilbert and Troitzsch 2005). ABM is therefore a particularly appropriate tool for capturing the complexity and emergent behavior that result from the stochastic and dynamic elements of regional food supply chains, as well as the sociological processes that influence farmer decision-making (Higgins et al. 2010). ABM was previously used to model the decision processes and interactions of regional food producers and consumers, in an effort to inform the food hub manager's supplier selection strategy (Krejci et al. 2015; Krejci et al. 2016).

Because humans do not behave or make decisions perfectly rationally (Gigerenzer and Selten 2002), when modeling human agents, using heuristic decision-making rules can reflect more realistic human decision-making. The heuristic learning method that is used to represent decision making in the agents in the ABM described in this paper is known as reinforcement learning. Reinforcement learning is a machine learning technique in which an agent attempts to maximize the rewards it receives over time from the decisions it makes (Sutton and Barto 1998). Agents receive higher rewards for good behavior and smaller rewards for bad behavior. The agent aims to maximize its total long-run reward, which is a sum of the individual rewards it receives from each action it chooses to take in a sequential decision process. In reinforcement learning, the agent must use a trial-and-error search to find the most rewarding actions in the long term (Sutton and Barto 1998). Agents will begin to be able to differentiate good actions from poor actions by interacting with the environment and observing state transitions and the rewards associated with them. In order to develop a good policy, the agent must both exploit the current knowledge it has from interacting with the system and explore other actions to make better decisions in the future. By exploiting its current knowledge of the system, the agent can ensure that its immediate action yields a relatively high reward, but in order to improve its understanding of the potential rewards for each action, exploration is needed. Upon convergence, the agent has developed a policy to guide its decision making in the system.

One popular reinforcement-learning algorithm is the Q-learning algorithm. The Q-learning algorithm iteratively updates Q-values, which store the values for each state *s* and action *a* as *Q(s,a)*. The Q-values are used to estimate the value of each action for an agent in a given state, where an agent prefers to select actions that yield the highest possible *Q(s,a)* (Watkins 1989). To balance exploration and exploitation, an ε-greedy method can be used (Tokic 2010). In this method, ε is a probability between 0 and 1 chosen by the modeler, which dictates how often the agent will select a random action rather than the action with the highest value (i.e., the greedy action). In order to choose the greedy action, the agent's Q-values must update as it moves through the decision process. Q-values are updated using the Q-learning algorithm, which corrects the Q-values to reflect new information gained throughout a simulation (Sutton and Barto 1998):

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha(r_t + \gamma \max_a Q(s_{t+1}, a_t) - Q(s_t, a_t)) \tag{1}$$

The algorithm shown in Equation (1) updates the Q-values by first adding the immediate reward (r_t) received by the agent upon arriving in a state to a discounted goal Q-value. The discount rate, γ, is a number between 0 and 1. The discount rate reflects how long the agent is willing to wait for a long-term reward. A small value for γ indicates that an agent will choose actions that yield high immediate rewards. Next, the current Q-value (i.e., the value of *Q* at time *t*) is subtracted. This quantity is then multiplied by the learning rate α, which determines how much weight the agent places on information recently gained. The learning rate is a number between 0 and 1. A learning rate of 0 means the agent will never learn and a learning rate of 1 means the agent will only consider the most recent information. Finally, this entire quantity is added to the current Q-value, which yields an updated Q-value. This updated Q-value replaces the current Q-value in the next time-step. The process is repeated each time an agent receives a reward/punishment for taking an action.

Reinforcement learning has previously been utilized in ABMs to inform agent decision-making. In an effort to improve decision-making for shipyard crane operators, Fotuhi et al. (2013) integrated reinforcement learning into an ABM of a shipyard to improve decision making, such that the operators choose to service trucks in a sequence that minimizes the waiting time for the trucks. Bone and Dragicevic (2010) incorporated reinforcement learning in an ABM to improve the ability of the model to determine optimal forest harvesting strategies among agents with conflicting goals. Q-learning has also been used to represent supplier decision making in a manufacturing supply chain model (Valluri and Croson 2005). In this model, the authors seek to determine how to select suppliers that consistently produce high quality goods.

This paper describes a conceptual ABM that is based on a regional food system in western Iowa and eastern Nebraska. The model was developed in NetLogo (version 5.3).

2 METHODOLOGY

The Iowa regional food supply chain has three types of actors: producers, customers, and a food hub manager. These actors are represented by agents in the ABM. The ABM is based on the assumption that decisions to sell products around the hub are always initiated by a producer (rather than a customer), in an attempt to increase their profits. In each one-week time-step, the producer agents choose what they believe will be the most profitable market channel for their products: through the hub, directly to customer agents, or some combination. By contrast, the customer agents have a passive role. They will not initiate a transaction around the hub, but if they are prompted by a producer agent, they will evaluate whether they want to accept the producer agent's offer to purchase products directly. The model environment consists of two metropolitan areas (i.e., the cities of Omaha and Des Moines) that are connected by an highway, which runs directly between the two areas.

2.1 Food Hub Manager Agent

There is one food hub manager agent in the model. The food hub manager agent purchases products from the producer agents and then aggregates and delivers these products to customer agents on a weekly basis. The food hub has the ability to detect when producer agents are disintermediating (i.e., when they sell products around the hub directly to customer agents).

2.2 Producer Agents

There are 20 producer agents in the model, which is approximately the number of regularly active producers in the Iowa regional food supply chain. Six producer agents are located in Omaha, six are located in Des Moines, and eight are located along the highway that connects the two cities. The producer agents are categorized into three sizes, based on the volume of products they have available to sell in each order cycle. The four small producers, six medium producers, and ten large producers can supply a maximum of 100, 200, or 300 units, respectively, to the food hub agent in each order cycle. There is only one type of food item sold by the producer agents and demanded by customer agents, and each item is priced the same. It is assumed that the food hub manager agent has sufficient demand to purchase all of the producers' supply in each order cycle, such that producer agents do not experience competition when selling through the hub. Further, it is assumed that the producer agents are unable to communicate with one another directly, and do not know what actions other producer agents are taking.

2.3 Customer Agents

There are 25 customer agents in the model that purchase regional food items through the food hub to sell to their own customers. These agents are primarily located within the two metropolitan areas of Omaha and Des Moines. The Omaha area includes twelve customer agents, the Des Moines area includes fourteen customer agents, and four customer agents are located near the highway. Each of the eleven small, seven medium, and seven large customer agents demands 100, 200, or 300 units of regional food, respectively, in each weekly time-step. Each customer agent is randomly assigned a value that represents its willingness to purchase directly from producer agents. The values fall between zero and nine, where values less than five mean that the customer agent is unwilling to purchase items directly from producer agents, and values that are five and greater mean that the customer agent is willing to purchase directly from producer agents, if approached.

2.4 Model Description

First, the model is initialized. The producer and customer agents are created and located in the model environment, and they are assigned a size (i.e., small, medium, or large) and respective supply/demand values. Each producer agent is given a 7x4 matrix of Q-values. All entries in the Q-value matrices are initially set to 1000 (a relatively high value) to prevent the producer agents from getting caught in a local maximum early in the simulation run. Each customer agent is assigned its buying preference value, which indicates whether the customer is willing to purchase directly from producer agents.

Producer agents have four possible actions to choose from in each time-step: selling no products around the hub, offering 25% of their products directly to customer agents (i.e., around the hub), offering 50% of their products directly to customer agents, or offering 75% of products directly to customer agents. Producer agents will choose their next action greedily 90% of the time, and the other 10% they will randomly choose between the four possible actions. If the producer agent is choosing the greedy action, it will find the highest Q-value in its matrix, corresponding to its current state. In time-step 0, the greedy action is defined as selling all products through the hub.

Next, the producer agent executes its chosen action. A producer agent that chooses to sell a portion of its products around the hub will offer the entire volume of products it has available to sell around the hub directly to the nearest customer agent, in an effort to reduce transportation costs. The customer agent then has three options: reject the offer outright, accept the entire amount, or accept a partial amount. If a customer agent has no remaining demand or is unwilling to purchase directly from producer agents, it will reject the offer outright. A customer agent will accept the offer if it is willing to purchase directly from producer agents and has remaining demand to be filled. If a willing customer agent has enough demand to accept the entire order, it will do so. Otherwise, the customer agent will accept an amount that fills its remaining demand. When a transaction occurs, the producer and customer agents update their supply and demand values accordingly.

If the producer's nearest customer agent rejects its offer, it will make an offer to the next closest customer agent, and so on, until it has no supply remaining or it has exhausted all selling options. It is assumed that the producer agents will only attempt to sell directly to customer agents that are located within 30 minutes driving distance of the producer's location. Further, it is assumed that any products that the producer agent is unable to sell around the hub will be sold through the hub.

After an action to sell items around the hub is executed, the producer agent will either be discovered or go undiscovered by the food hub manager agent. It is assumed that the food hub manager would be more likely to discover a producer who sells greater volumes around the hub. Therefore, when a producer agent offers 25%, 50%, or 75% of its products around the hub, it has a 25%, 50%, or 75% chance, respectively, to be discovered by the manager agent. Therefore, a producer agent's action will result in one of seven possible states, as shown in Table 1. For example, if a producer agent chooses the action "Attempt to sell 50% of items around hub" and its behavior is detected by the food hub manager agent, it would move to state Attempt50-Caught (State 4).

Table 1: Producer agent states.

Action	Discovered by Hub?	Resulting State	State #
Sell all items through hub	N/A	None	1
Attempt to sell 25% items around hub	Yes	Attempt25-Caught	2
	No	Attempt25-Success	3
Attempt to sell 50% items around hub	Yes	Attempt50-Caught	4
	No	Attempt50-Success	5
Attempt to sell 75% items around hub	Yes	Attempt75-Caught	6
	No	Attempt75-Success	7

A producer agent receives a reward based on the state that results from its action (i.e., a reward/punishment from the manager), as well as the customer agents' responses to the producer agent's action (i.e., deciding whether or not to purchase directly). Under the management strategy that is currently employed by the real-life Iowa food hub manager, producers are not punished in any way for getting caught selling products around the hub. Therefore, in the current version of the ABM (which reflects the status quo Iowa regional food system), producer agents do not receive a punishment for arriving in a "caught" state. For products successfully sold directly to customer agents, it is assumed that producer agents receive 100% of the profit, while they only earn 80% of the profit on items sold through the food hub, since the food hub charges a fee of 20%. However, there are transportation costs associated with selling directly to customer agents, which are proportional to the distance traveled by the producer agent. To calculate its reward, a producer agent subtracts its transportation costs from its total profit accumulated from selling to producers and the hub. It then updates its corresponding Q-value to reflect the value of the reward, using the algorithm shown in Equation (1).

After all direct sales are made, the food hub manager agent purchases any remaining products from producer agents and uses this to fulfill remaining customer agent demand. The manager agent records the number of customer agents that purchase products through the hub and the total number of items sold to those customer agents. The model is then ready to move the next time-step.

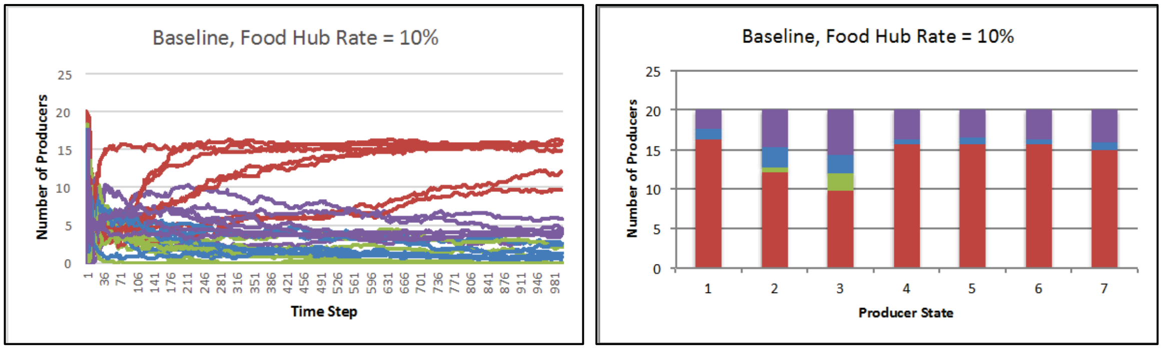
3 EXPERIMENTATION AND RESULTS

In an effort to reduce disintermediation, the manager of the Iowa food hub has considered reducing the service rate that she charges the producers. To assess the effectiveness of this approach, the ABM was run using different service rates: 10%, 15%, and 20% (the Iowa food hub's actual rate). The model was run for 1000 time-steps, and the mean number of producer agents choosing each of the four actions in each of the seven possible agent states over 10 replications are shown for the three service rates in Figures 1-3.

Before performing these experiments, the values of the producer agents' learning rate (α) and discount factor (γ) were assigned. Ideally, these parameters should reflect the abilities and preferences of the human producers in the real-life regional food system. Therefore, an iterative process was used to tune the model such that it yielded behavior that was a sufficiently realistic reflection of this system. The value of α was eventually set to 0.9 to reflect relatively rapid adaptations by producers in response to the rewards they received, such that the initial period of exploring the state space would end after several months, rather than years. This can be seen in Figures 1a, 2a, and 3a – there are large fluctuations in the number of producer agents following each policy in the early time-steps, but by time-step 36 (i.e., after 36 simulated weeks), the proportions are relatively steady.

Determining an appropriate value for the discount factor was less straightforward. When α was set to 0.9 and γ was set to 0.9, more than 50% of producer agents chose selling 75% of their products around the hub as their best action. However, in the real-life Iowa regional food system, there are some producers who do not sell products around the hub at all, nor would they ever consider doing so. Thus, a discount factor of 0.9 did not seem valid. The real-life producers are also unlikely to strictly consider their immediate reward from their selling decisions, with no regard for future rewards. To reflect this, a discount factor of $\gamma = 0.5$ was finally assigned.

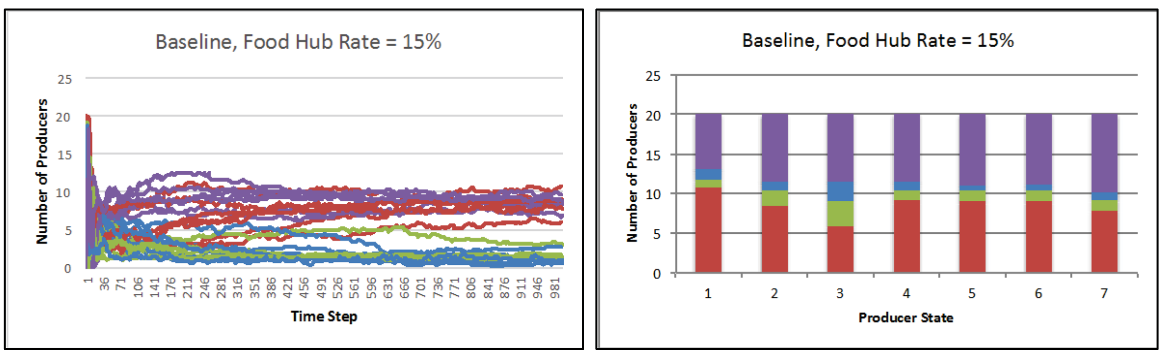
The results shown in Figures 1-3 indicate that when the food hub employs a 20% service rate (i.e., the status quo rate for the real-life Iowa food hub), at least 38% of the producer agents will select the action of attempting to sell 75% of their products around the hub, regardless of their current state. The next most popular action in all states is to sell all products through the hub, where at least 22% of producers will select this action no matter which state they are currently in. It appears that if producers decide to sell products around the hub, it is most beneficial to sell as much as they can (i.e., 75% of their products) around the hub. When the hub lowers its service rate to 15%, these two policies become split more evenly, with at least 35% of producers choosing to sell 75% of products around the hub and 30% choosing to sell no products around the hub (from all current states). It is only once the hub lowers its service rate to 10% that at least 49% of producer agents find selling no products around the hub to be the best action to take, regardless of their current state. However, for the real-life Iowa food hub, a 10% service rate is likely infeasible. The food hub must bring in sufficient revenue to cover its operating costs, and lowering the rate too far will make that impossible. The food hub manager is already experiencing problems covering transportation costs at its current service rate of 20%.



(a) (b)

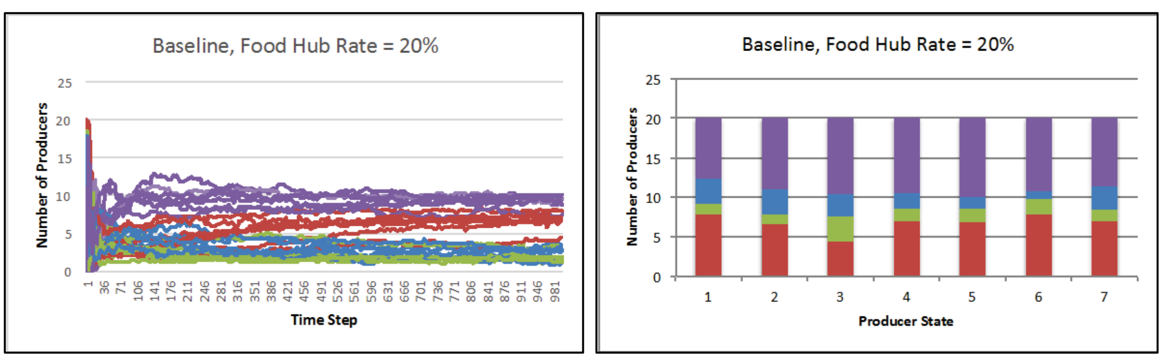
Color	#	Best Action
Red	0	Sell all items through hub
Green	1	Attempt to sell 25% items around hub
Blue	2	Attempt to sell 50% items around hub
Purple	3	Attempt to sell 75% items around hub

Figure 1: Producer agents following each policy a) over time and b) in the final time-step for a 10% rate.



(a) (b)

Figure 2: Producer agents following each policy a) over time and b) in the final time-step for a 15% rate.



(a) (b)

Figure 3: Producer agents following each policy a) over time and b) in the final time-step for a 20% rate.

In each experiment the Q-values and subsequent actions chosen will typically vary from one producer agent to the next. Since producers have different supply levels and locations, the best action for a given state will be unique to a given producer. Therefore, under these experimental conditions, there is no single policy that dominates for every producer agent. To demonstrate this, five individual producer agents having various sizes and locations were examined. Table 2 summarizes the percentage of times each of these five producer agents chose an action over the course of the experiment. As Table 2 shows, Producer 0 tended to sell all of its products through the hub (Action 0), while Producers 3 and 13 generally preferred to attempt to sell 75% of products around the hub (Action 3). Producer 15 did not have a very strong preference for any particular action, but it did not prefer to attempt to sell 25% of its products around the hub (Action 1), which was the most-preferred action of Producer 17. These differences demonstrate how a producer agent's size, its location relative to other producers and customers, and customers' buying preferences impact the decision policy that will be best for a given producer. The best policy for one producer will not necessarily be the best for another producer.

Table 2: Percentages of actions selected by five producer agents over 1000 time-steps at the 20% rate.

Producer #	**Producer Action**			
	0	**1**	**2**	**3**
0	85%	5%	5%	5%
3	8%	6%	13%	73%
13	6%	6%	6%	82%
15	26%	6%	30%	38%
17	6%	79%	7%	8%

4 CONCLUSION

Although the ABM described in this paper is a conceptual model, the results of preliminary experimentation suggest that it generally performs as expected. Over time, each producer agent develops an individual decision policy that reflects its unique location and supply volume, as well as competition with other producers for direct sales to customers. When the food hub service rate was increased, more producer agents preferred to sell products around the hub, and when the rate was reduced, more producer agents preferred to sell products through the hub. However, the results of these initial experiments suggest that, regardless of the food hub manager's approach to reducing producer disintermediation, there will likely be instances in which producers still find that the best policy is to sell their products around the hub. There will always be producers who are in a position to sell their products directly to nearby customers and will not rely on the food hub's transportation and marketing services as much as other producers do. Therefore, the food hub will need to find an appropriate balance between the number of producers it can tolerate selling products around the hub and the service rate it charges the producers.

Ongoing experimentation with the model is focused on testing the effectiveness of other management strategies on reducing disintermediation. For example, McCann and Crum (2015) recommended that food hubs stop working with producers who sell around the hub. The temporary (or possibly permanent) removal of disintermediating producer agents from the hub's pool of suppliers could serve as an effective punishment, such that choosing to sell through the hub is more rewarding in the long run. The Iowa food hub manager has also mentioned that she is considering offering profit-sharing incentives to the producers to encourage them to do more business with the hub. Such incentives could be incorporated into the model, and the behavior of the producer agents in response to different types of rewards could be observed.

Future development of the model is focused on validation through the collection of empirical data from the real-life Iowa regional food system. The conceptual model presented in this paper is based on numerous assumptions, many of which were based on conversations with the food hub manager. However, gaining a better understanding of the actual motivations and decision processes of the producers and customers in the system would greatly increase the validity of the model, as well as improving its usefulness in informing the food hub manager's strategic decision making. An interview protocol has been developed and piloted with several Iowa regional food producers and customers, and the preliminary results were enlightening – they suggest that customers play a more significant role in the disintermediation behavior than previously thought. More interviews with producers and customers will be performed, and the resulting data will serve as inputs into the ABM to justify (or replace) many of the current modeling assumptions.

The preliminary experiments described in this paper suggest that using an agent-based model that incorporates Q-learning for each producer agent is effective in generating and understanding long-term agent and overall system behavior. The Q-learning allows producer agents to use a heuristic trial-and-error approach to developing decision policies that perform well for them over time, which is likely a more realistic approach to modeling human agents than assuming that they are capable of perfectly rational decision making. ABM with reinforcement learning has the potential to be an effective tool for studying disintermediation in supply chains and guiding management strategies to avoid it.

ACKNOWLEDGMENTS

The authors would like to thank the Leopold Center for Sustainable Agriculture for supporting this research.

REFERENCES

Barham, J., D. Tropp, K. Enterline, J. Farbman, J. Fisk, and S. Kiraly. 2012. *Regional Food Hub Resource Guide.* Washington, DC: U.S. Dept. of Agriculture, Agricultural Marketing Service.

Bone, C., and S. Dragicevic. 2010. "Simulation and Validation of a Reinforcement Learning Agent-Based Model for Multi-Stakeholder Forest Management". *Computers, Environment and Urban Systems* Vol. 34, No. 2, pp. 162-174.

Choi, T.Y., K.J. Dooley, and M. Rungtusanatham. 2001. "Supply Networks and Complex Adaptive Systems: Control versus Emergence". *Journal of Operations Management* Vol. 19, No. 3, pp. 351-366.

Fotuhi, F., N. Huynh, J. M. Vidal, and Y. Xie. 2013. "Modeling Yard Crane Operators as Reinforcement Learning Agents". *Research in Transportation Economics* Vol. 42, No. 1, pp. 3-12.

Gigerenzer, G. and R. Selten. 2002. *Bounded Rationality: The Adaptive Toolbox.* Cambridge: MIT Press.

Gilbert, N. and K. G. Troitzsch. 2005. *Simulation for the Social Scientist (2nd ed.)*. Berkshire, England: Open University Press.

Higgins, A.J., C.J. Miller, A.A. Archer, T. Ton, C.S. Fletcher, and R.R.J. McAllister. 2010. "Challenges of Operations Research Practice in Agricultural Value Chains," *Journal of the Operational Research Society* Vol. 61, No. 6, pp. 964-973.

Jones, P., D. Comfort, and D. Hillier. 2004. "A Case Study of Local Food and its Routes to Market in the UK". *British Food Journal* Vol. 106, No. 4, pp. 328-335.

Krejci, C. C., R.T . Stone, M. C. Dorneich, and S. B. Gilbert. 2015. "Analysis of Food Hub Commerce and Participation using Agent-Based Modeling: Integrating Financial and Social Drivers". *Human Factors (Special Issue: Human Factors Prize on Sustainability)* Vol. 58, No. 1, pp. 58-79.

Krejci, C. C., M. C. Dorneich, and R. T. Stone. 2016. "Assessing Values-Based Sourcing Strategies in Regional Food Supply Networks: An Agent-Based Approach". *Journal on Policy & Complex Systems* Vol. 2, No. 2, pp. 21-48.

McCann, A. and R. Crum. 2015. "Avoiding Sales Around the Hub". Local Food Marketplace. Available via http://home.localfoodmarketplace.com/wp-content/uploads/2015/10/Avoiding-Sales-around-the-Hub.pdf. Accessed January 8, 2017.

Meter, K. 2006. "Evaluating Farm and Food Systems in the U.S.". In: *Systems Concepts in Evaluation: An Expert Anthology*, B. Williams and I. Imam (Eds.), pp. 141-159. Inverness, CA: EdgePress,.

Mills, J. F. and V. Camek. 2004. "The Risks, Threats and Opportunities of Disintermediation". *International Journal of Physical Distribution & Logistics Management* Vol. 34, No. 9, pp. 714-727.

North, M. J. and C. M. Macal. 2007. *Managing Business Complexity: Discovering Strategic Solutions with Agent-Based Modeling and Simulation*. Oxford, U.K.: Oxford University Press.

Pathak, S. D., J. M. Day, A. Nair, W. J. Sawaya, and M. M. Kristal. 2007. "Complexity and Adaptivity in Supply Networks: Building Supply Network Theory Using a Complex Adaptive Systems Perspective". *Decision Sciences* Vol. 38, No. 4, pp. 547–580.

Tokic, M. 2010. "Adaptive ε-greedy Exploration in Reinforcement Learning Based on Value Differences". In *Annual Conference on Artificial Intelligence*, pp. 203-210. Berlin: Springer.

Sutton, R. S. and A. G. Barto. 1998. *Reinforcement Learning: An Introduction.* Cambridge: MIT Press.

Valluri, A. and D. Croson. 2005. "Agent Learning in Supplier Selection Models". *Decision Support Systems* Vol. 39, No. 2, pp. 219-240.

Watkins, C. 1989. *Learning from Delayed Rewards.* PhD thesis, Cambridge University.

AUTHOR BIOGRAPHIES

TERI J. CRAVEN is a graduate student in the Department of Industrial and Manufacturing Systems Engineering at Iowa State University. Her research interests include modeling and simulation for sustainable supply chain management. Her email address is tjcraven@iastate.edu.

CAROLINE C. KREJCI is an Assistant Professor in the Department of Industrial and Manufacturing Systems Engineering at Iowa State University. She holds a Ph.D. in Industrial Engineering from the University of Washington. Her research is focused on the development of quantitative methodologies for the analysis and sustainable management of sociotechnical systems, including supply networks and production systems. Her email address is ckrejci@iastate.edu.