JOURNAL ON POLICY AND COMPLEX SYSTEMS

Vol. 3, No. 2 • Fall 2017

Edited by Mirsad Hadžikadić & Liz Johnson

Policy and Complex Systems
Volume 3, Number 2 - Fall 2017

CAPS Special Issue Table of Contents

Editors & Editorial Board

Editor's Letter

Welcome to the special issue of the Journal on Policy and Complex Systems (JPCS) on complexity, policy, and social good. This issue includes updated papers presented at the inaugural conference on Complexity-Based Analytics and Policies for Social Good, held in Washington, DC on April 12–14 at the University of Maryland's Washington, DC campus. This conference was organized by the JPCS and co-sponsored by the University of Maryland's Robert H. Smith School of Business, Computational Social Science Society of the Americas, Cisco, and the University of North Carolina at Charlotte. Even though it is unusual for journals to organize conferences, we thought that it was imperative for the JPCS to help establish and promote a new application domain of policy and complexity, ***social good***.

This issue of the journal is loosely divided in four sections. Section 1 is focusing on ***social conflicts***, covering conflict dynamics, civil revolutions, and genocide. Section 2 includes papers on ***social structures***, containing papers on intergovernmental networks, an education system intervention framework, censorship, novel approaches to foster care, and combating diseases. Section 3 covers papers on ***social behavior***. They include papers on models for detecting human trafficking flows, farmer behaviors under various policy schemes, and a citizen-centric government services ecosystem. Finally, Section 4 includes one speculative paper on democracy and analytics, labeled as ***future directions***.

We will continue this practice of initiating and/or supporting new disciplines and application domains related to policy and complex systems. Please let us know if you have ideas for future offerings and special issues.

Mirsad Hadzikadic

Editor, JPCS

Scenarios of Social Conflict Dynamics on Duplex Networks

Miron Kaufman[A], Sanda Kaufman[A], Hung The Diep[B]

Abstract

Social conflicts among large groups of people can have serious and costly consequences that can exceed the time and space boundaries of the contentious groups. They require carefully thought-out strategies for addressing the contested issues. However, the complexity of social conflicts poses obstacles to their resolution. Not only is it difficult to find clear cause-and-effect relationships, but the conflict dynamics impede prediction of outcomes. For any group in conflict, strategizing to find a way out requires an approach that allows testing of the range of consequences of various strategies. Thus in complex situations, where predicting how the opponent will respond to a strategy is difficult, a party to conflict might engage instead in anticipation, generating and preparing for a range of possible scenarios. We draw from duplex networks modeling to further analyze a recently proposed scenario-generating model of conflicts and illustrate its application with two examples.

Keywords: social conflict, complexity, network dynamics, and social physics.

[A] Cleveland State University, [B]University of Cergy-Pontoise

doi: 10.18278/jpcs.3.2.1

摘要

人群间的社会冲突能产生严重且成本不低的后果。这些后果超越了斗争团体间的时间限制和空间限制。为解决斗争问题，需要详细严密的策略。然而，社会冲突的复杂性为解决问题造成了阻碍。不仅很难找出明显的因果关系，而且难以预测结果。对冲突中的任何一个团体而言，都需要寻求方法以测试不同策略的结果，才能找到解决办法。因此在复杂情境中，由于很难预测对手如何回应策略，斗争的一方可能需要对一系列可能发生的情况进行预期和准备。本文对双工网络进行建模，进一步分析近期提出的冲突情境产生（scenario-generating）模型，同时用两个例子描述了该模型的应用。

关键词：社会冲突, 复杂性, 网络动态, 社会物理学

Resumen

Los conflictos sociales en grupos grandes de personas pueden tener consecuencias serias y costosas que pueden exceder los límites de tiempo y espacio de grupos polémicos. Requieren estrategias cuidadosamente pensadas para abordar los problemas controvertidos. Sin embargo, la complejidad de los conflictos sociales pone obstáculos en su resolución. No solo es difícil encontrar relaciones de causa y efecto claras, sino que también la dinámica del conflicto impide la predicción de resultados. Para cualquier grupo en conflicto, formular estrategias para encontrar una salida requiere un acercamiento que permita poner a prueba el rango de consecuencias de varias estrategias. Por lo tanto, en situaciones complejas, donde es difícil predecir cómo el contrincante responderá a una estrategia, una de las partes en el conflicto podría en vez de esto ocuparse por anticipado de generar y preparar para una variedad de posibles situaciones. Utilizamos la modelización de redes dúplex para analizar más profundamente un modelo recientemente propuesto que genera situaciones de conflicto e ilustramos su implementación con dos ejemplos.

Palabras clave: conflicto social, complejidad, dinámica de redes, física social

Social Conflict Scenarios

Social conflict (Rubin, Pruitt, & Kim, 1994) occurs among large groups of people who differ on specific issues that each group holds important, with interests driven by differing values and identities. It can be latent but frequently bursts in the open around joint decisions. At such decision points, conflicting social groups contend with each other. Conflicts around politics and consequential policies at the federal, state, and local levels fall in this category, as do international disputes. The waging of conflict can be pursued peacefully through processes such as voting or negotiations; at times, it can also descend into violence, both within and between countries. The peaceful decision processes are by no means necessarily friendly. For example, members of the clashing groups can actively engage in persuading each other to their point of view in order to secure favorable outcomes.

Conflicts tend to be very different from each other despite some general similarities. For example, although environmental conflicts usually revolve around environmental issues, their location, time, political and social contexts, specific issues, and participants make them both complex and rather unique. At the eye level of group members, interveners or observer, patterns and future directions are rather resistant to prediction and may appear chaotic. It is also difficult to foresee the outcomes resulting from the numerous decisions of interacting parties in time. The chaotic appearance at one observation level may look different at a higher level. At such a level, we may no longer discern the detailed interactions but, as with other with other complex phenomena, we may distinguish patterns. Then it is possible to derive a range of possible social conflict trajectories, or scenarios.

When two groups are engaged in conflict, the web of interactions inside each group and between the groups can be represented as two interconnected networks. Inside each network and between them, the groups' members are nodes and the inter-node links are their interactions. This approach, used in physics to study complex phenomena, is applied here to the study of complex social interactions. The utility of applying the network modeling approach to the study of social conflicts and generation of scenarios resides in the following notion: stakeholders to social conflicts need to foresee the unpredictable (Bonabeau, 2002; Lempert, Popper, & Bankes, 2002).

To engage in a joint decision—where the outcome for all depends on the choices of many—conflicting parties need to prepare strategies contingent on their opponents' reactions. However, in complex situations with numerous individuals making interacting choices, such reactions are difficult to predict reliably enough to prepare responses. Therefore, it may be wiser for each party to abandon the quest for prediction and switch to anticipation. Instead of aiming to figure out what the opponent will do (equivalent to a point prediction), it is more feasible and helpful to anticipate a range of possible opponent reactions and prepare for that range instead of the point estimate (Lempert et al., 2002).

Underlying the anticipation approach is the idea that social conflicts occur within complex, interrelated physical and social systems. Decisions in such contexts are fraught with uncertainty. Unlike simple one-cause–one-effect relationships that can be understood and managed, social systems may yield some of the results sought, but often also a host of unforeseen and negative side effects. To compound the difficulty, some decision consequences accrue quickly while others take longer time to show results. For example, a change in tax policy at the federal or state level may yield some economic results quickly. Also, education policy changes may take longer to show results. Finally, some environmental policies may not show their consequences during the lifetime of the decision makers.

Planners and ecologists faced with complex interacting social–ecological systems use scenarios to explore possible consequences of planning or management decisions (Butler, Corvalan, & Koren, 2005; Cobb & Thompson, 2012; Kriegler et al., 2012). Physicists use toy models to represent and study complex interactions (Marzuoli, 2008). One method for generating possible conflict outcomes—or scenarios—is to construct a model that can be queried in what–if fashion to explore consequences of different conditions or of adopting various strategies. We combine the network approach from physics with the scenario technique from planning to examine the range of possible outcome of social conflicts. We begin with two groups and represent their interactions using a duplex network (two interrelated networks) model. We generate and explore scenarios of the interactions that differ in the assumptions about the values of a small number of parameters. Any of the two groups could then use the model to ask what–if questions that can help the group select a strategy that might be wise for a range of scenarios instead of just one predicted possibility. Conflict strategies that cover a range of possible outcomes are considered robust (Lempert et al., 2002; Lempert, Groves, Popper, & Bankes, 2006), as opposed to fragile ones relying on point predictions.

We propose to illustrate the potential of network models to anticipate outcomes of social conflict dynamics through scenario generation. We begin by outlining a duplex network model, drawing the correspondence between elements and relationships of a two-group social conflict and some physics concepts that can be studied on networks. Then we show how the duplex network model can be applied to explore various conflict scenarios, using two recent social conflict confrontations—the 2016 "Brexit" referendum and the 2016 presidential elections in the United States. In our examples, the outcome is already known. However, when the model is applied to a conflict before the outcome has materialized, it can provide scenarios that a group can use to devise or alter its strategy in response to the dynamic at work and anticipated outcomes. The advantage of such outcome scenarios over other analyses that can inform strategy is that it parsimoniously offers a global perspective on possibilities.

Model

We consider two groups in conflict. In each group, each individual has a preference regarding how the conflict should be resolved. Each individual has an attitude *s* (corresponding to spin in physics) with respect to a specific conflict. In group 1, individuals' attitudes range from −M1 (very open to negotiating some agreement) to M1 (inclined to protracted conflict due to adherence to extreme positions consistent with one's ideology). Similarly, in group 2, individuals' attitudes range from −M2 to M2. Here for illustrative purposes we consider M1 = M2 = 3.

Members within each group are networked: they interact with each other. Each individual acts with a certain intensity to persuade others in the group to his/her point of view and is subject to others' persuasion efforts. We assume that each individual interacts with all other individuals inside his/her own group as well as with individuals in the other group. The pattern of interconnections corresponds to two Renyi–Erdös equivalent neighbor networks, a configuration with links of equal strength among all nodes.

Our model of the individuals' interactions within each group and between the groups yields averages of the individual preferences at any time *t*. Each network (group) has its own average value: *s*1 and *s*2, respectively. The in-group intensity (energy) of advocacy of an individual from group 1 is –*J*1×*s*×*s*1 while the corresponding energy of an individual in group 2 is: –*J*2×*s*×*s*2, where s1 is average of all s in group 1 and *s*2 is the average of all *s* in group 2. To reflect the effects of mutual persuasion, the inter-group intensity of interaction (energy) of any individual is taken to be proportional to the product between the individual's preference s and the mean value of the preferences of the other group's members: K12×*s*×*s*2 for an individual in group 1, and –K21×*s*×*s*1 for an individual in group 2.

Given the individuals' preferences at time *t*, the Boltzmann probability distribution of the preferences s at time *t*+1 for each group is proportional to the exponential of the intensity of interactions (energy):

$$\rho 1(s, t+1) = \frac{e^{s(J1s1_t + K12s2_t)}}{\sum_{s=-M1}^{M1} e^{s(J1s1_t + K12s2_t)}}$$

$$\rho 2(s, t+1) = \frac{e^{s(J2s2_t + K21s1_t)}}{\sum_{s=-M2}^{M2} e^{s(J2s2_t + K21s1_t)}}$$

Since persuasion is not instantaneous, we assume the interactions between the individual preferences *s* at time *t* + 1 and the collection of individuals from the two groups to be lagged, with the averages s1 and **s**2 evaluated at an earlier time t. ρ1(*s*, *t*) and ρ2(*s*, *t*) yield the fractions of all individuals who have preference *s* at time *t*, in group 1 and group 2, respectively. We (Diep, Kaufman, & Kaufman, 2017) have proposed this model and presented some of its predictions concerning the time evolution of the mean attitudes *s*1 and *s*2. Here we concentrate on the time evolution of the distribution of attitudes.

Examples

To illustrate how this model can offer some insight into the dynamics of social conflict, we show a couple of implementations that attempt to model two relatively recent events with a two-group structure: "Brexit" and the 2016 U.S. elections. In-depth political analyses are intellectually satisfying in that they seem to link causes and effects that "make sense" to us. However, as in the two examples we describe, such analyses can miss the big picture. We would argue that both types of analysis are needed and that our simple model can contribute valuable input to strategists on each side.

Brexit Case

In the Brexit case, Group 1 was composed of the British government and supporters of continued membership in the European Union (EU); group 2 contained individuals who wanted to exit the European Union. The case was characterized by time oscillations, with the pro-EU group and the pro-Brexit group alternating in leading in polls at different times. We chose relatively high values for J1 and J2, with K12 <0 and K21 > 0 to obtain similar oscillations sustained in time. In terms of the interactions between the two networks, K12 < 0 reflects that the extreme (uncompromising) wing of the pro-Brexit Group 2 encourages pro-EU Group 1 members to be accommodating, while the moderate (compromising) wing of

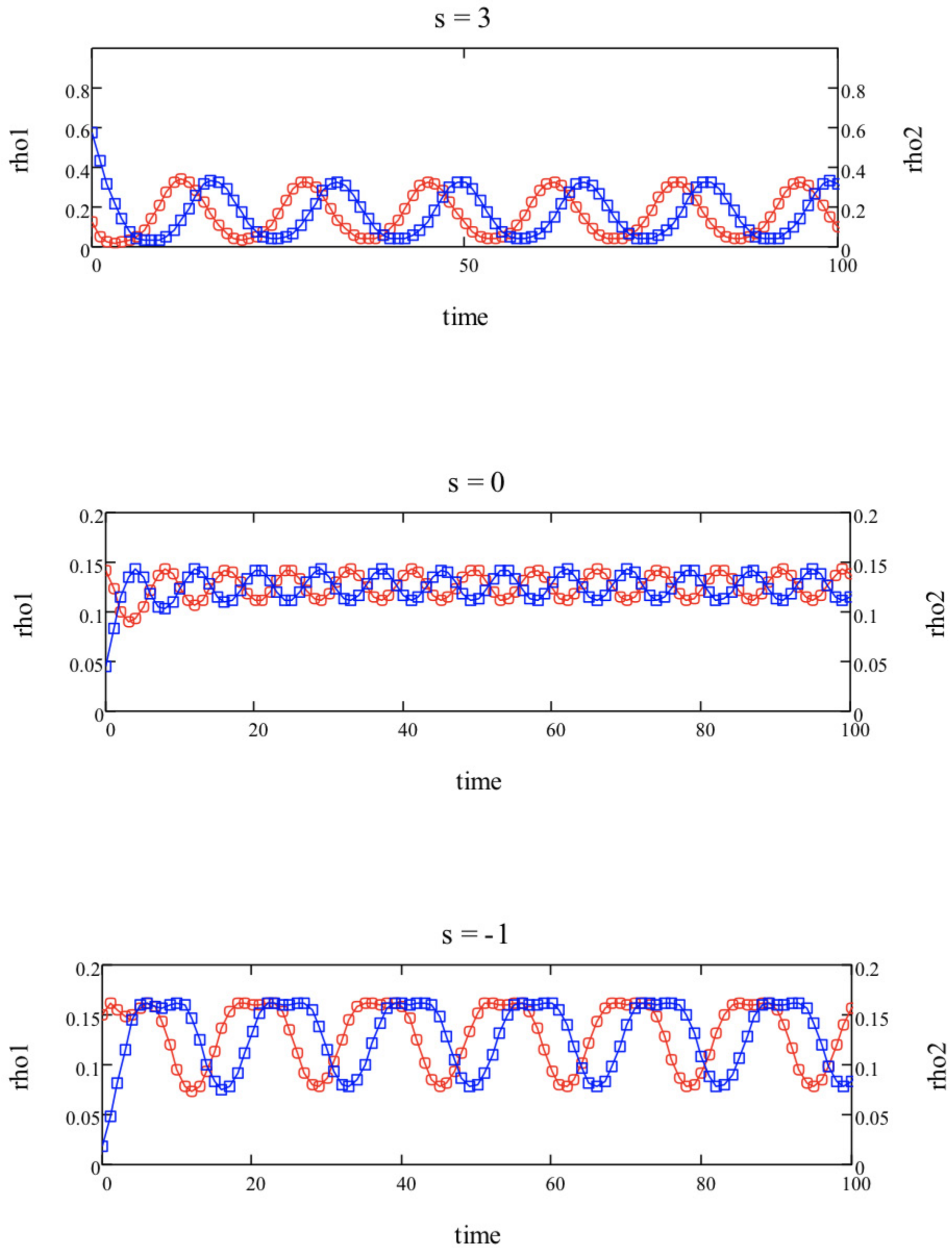


Figure 1. Time oscillations of the attitudes $s = 3, 0, -1$.

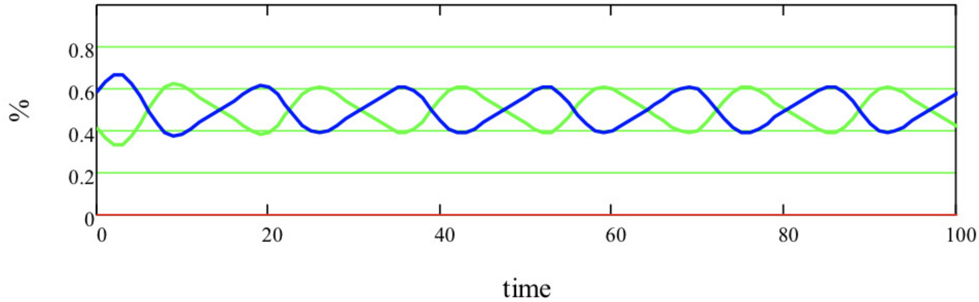


Figure 2. Time oscillations of fractions of EU supporters and Brexit supporters. In this scenario everyone votes.

Group 2 fuels the extreme wing of Group 1. K21 > 0 means that extreme wing of Group 1 strengthen the extreme wing of Group 2, while the moderate wing of Group 1 helps Group 2 moderates. This is shown in the graphs of Figure 1, where: J1 = J2= 0.25, K12 = −0.1, K21 = 0.1.

Before the Brexit referendum, we

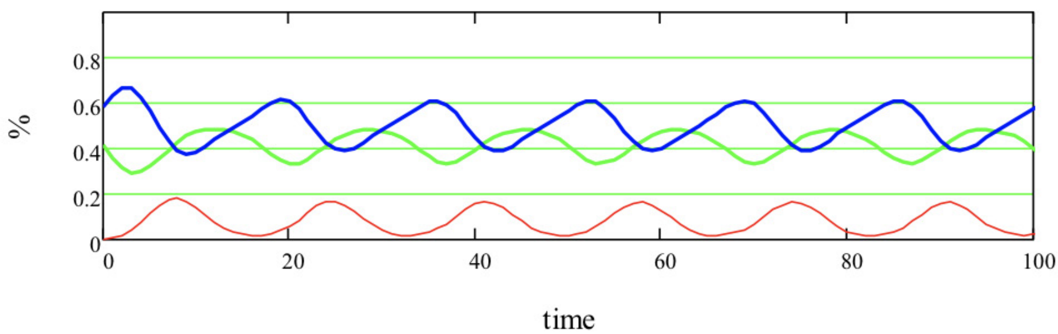


Figure 3. Time oscillations of fractions of EU supporters and Brexit supporters. Lowest curve corresponds to the non-voters.

could explore various outcome scenarios that could result from these patterns. For example, we assume that those who prefer to remain in the EU for staying in Europe are members of Group 1 whose preferences s range from 3 to −2, as well as those in Group 2 who are most accommodating, with s = −3. The rest (Group 1 members with s = −3 and Group 2 members with s from 3 to −2) would rather break away from the EU.

In another possible Brexit scenario, Group 1 members with preferences s from 3, to −2 prefer to stay in the EU; Group 2 members with preference values s between 3 and −2 would rather break away; members of both groups whose preference value s = −3 are disengaged and do not vote. Here is the outcome in this scenario:

United States Presidential Elections Case, 2016

Next, we simulate the dynamics leading to the outcome of the 2016 U.S. elections. Republicans are Group 1 and Democrats are Group 2. In our model, Democrats are more cohesive than

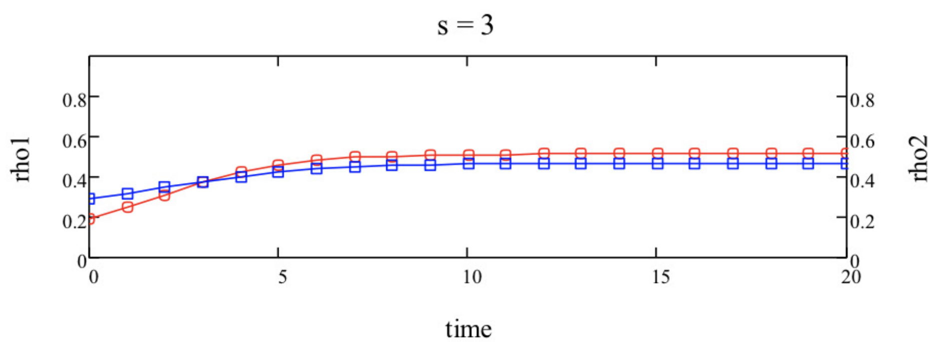


Figure 4. Time evolution of the attitudes $s = 3$ and −3.

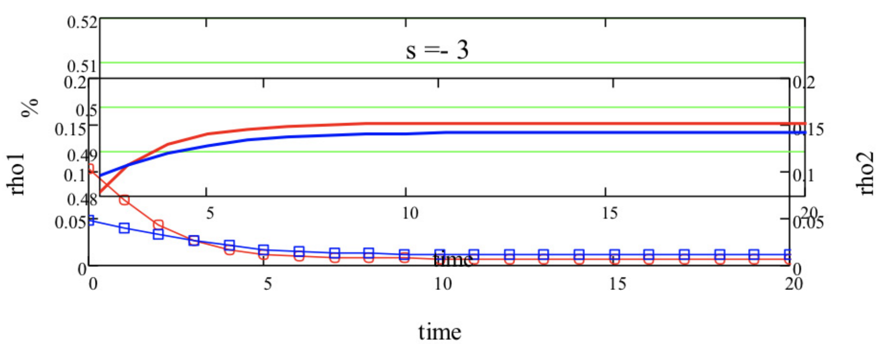


Figure 5. Time evolution of fractions of democrats and republican supporters. In this scenario the attitude $s = -3$ in both groups either do not vote or vote for third party.

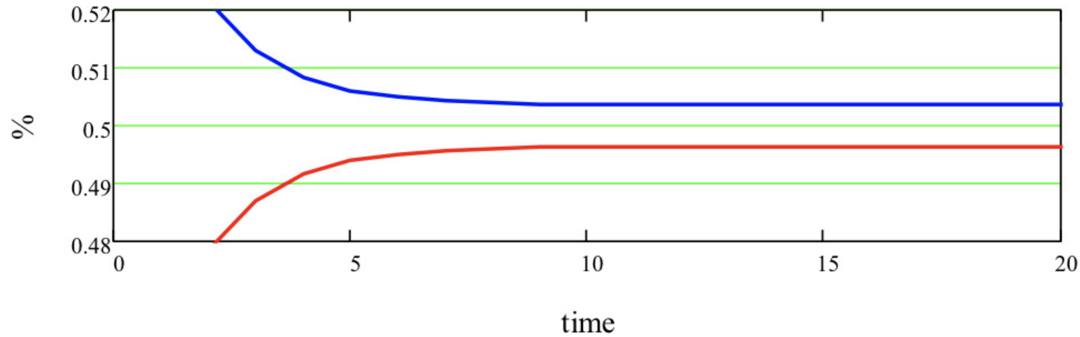


Figure 6. Time evolution of fractions of democrats and republican supporters. In this scenario the attitude $s = -3$ in both groups votes for the democrats.

Republicans are: $J2 > J1$; they influence Republicans more than Republicans influence them: K12 > K21. The results below were obtained for: $J1 = 0.25$, $J2 = 0.3$, K12 = 0.1, K21 = 0.02. In Figure 4 we show the time evolution of the non-compromising attitude $s = 3$ and of the compromising attitude $s = -3$.

We consider a couple of scenarios. In the first scenario (Figure 5), those in Group 1 whose preferences s range from 3 to −2 are likely to vote for the Republican candidate. Similarly, in Group 2 those with preferences s ranging from 3 to −2 vote for the Democratic candidate. In both groups, those whose preference $s = -3$ either do not vote (they are not sufficiently engaged) or they vote for a third candidate. In this scenario, the Republican candidate wins the elections.

In the second scenario (Figure 6) those in both groups with preference $s = -3$, whom we could view as independents groups, vote for the Democratic candidate. As a result, the democrats win the election.

Conclusions

We have proposed a parsimonious model for generating anticipatory scenarios, to be used by parties in two-group social conflict to devise strategies. We have shown how distributions of attitudes inside each group and their time evolutions can be derived, to gain insight into possible group choices. We have illustrated with two examples—2016 "Brexit" referendum and US elections—how such a model can be used to analyze real two-group conflicts. It has the potential to be enriched with several realistic features to describe other conflicts.

In the model's current form, for illustrative purposes, we have selected the values of the intra-group interactions J and the inter-group interactions K to match the example narratives regarding these conflicts. We intend to explore ways to determine these coupling constants endogenously. For example, the cohesion couplings J may increase when the two-group dynamics moves towards conflict, i.e. the mean attitude s of group members approaches the value 3, representing dominant intransigence. Similarly, *J*s may decrease when the dynamics evolve away from conflict toward a settlement, i.e. the mean attitude s of group members tends to a conciliatory s value of −3. The model will then become non-linear, and consequently may exhibit chaotic dynamics.

Currently our model is symmetric around 0 with respect to the range s to $-s$ of individual attitudes from belligerence to conciliation. In the next model iteration, we will explore the effects of breaking this symmetry. We will also examine this model's behavior when more than two groups are involved in social conflict.

The model we proposed here represents a micro-level approach, in that it takes into account the behaviors of individuals within the conflicting groups. In a previous macro-level model (Kaufman & Kaufman, 2015) we investigated the conflict trajectories of two conflicting groups, attempting to anticipate whether the two-group system would move in time toward settlement, continued conflict, or impasse. The micro- and macro-level models can be connected to offer insights about the conflict trajectories when driven by individual choices.

Finally, we are interested in continuing to pit the model against real conflicts. To this end, we will apply the mathematical framework developed here, with appropriate adjustments, to an expanded set of real conflicts, such as those described by Von Briesen, Bacaksizlar, and Hadzikadic (2017).

Acknowledgements

We are grateful for useful comments received from Professor Mirsad Hadzikadic.

References

Bonabeau, E. (2002). Predicting the unpredictable. *Harvard Business Review*, *80*(3), 109–116.

Butler, C. D., Corvalan, C. F., & Koren, H. S. (2005). Human health, well-being, and global ecological scenarios. *Ecosystems*, *8*(2), 153–162.

Cobb, A. N., & Thompson, J. L. (2012). Climate change scenario planning: a model for the integration of science and management in environmental decision-making. *Environmental Modelling & Software*, *38*, 296–305.

Diep, H. T., Kaufman, M., & Kaufman, S. (2017) Dynamics of two-group conflicts: A statistical physics model. *Physica A*, 469, 183–199.

Kaufman, S., & Kaufman, M. (2015). Two-group dynamic conflict scenarios: "Toy model" with a severity index. *Negotiation and Conflict Management Research*, *8*(1): 41–55.

Kriegler, E., O'Neill, B. C., Hallegatte, S., Kram, T., Lempert, R. J., Moss, R. H., & Wilbanks, T. (2012). The need for and use of socio-economic scenarios for climate change analysis: A new approach based on shared socio-economic pathways. *Global Environmental Change*, *22*(4), 807–822.

Lempert, R., Popper, S., & Bankes, S. (2002). Confronting surprise. *Social Science Computer Review*, *20*(4), 420–440.

Lempert, R. J., Groves, D. G., Popper, S. W., & Bankes, S. C. (2006). A general, analytic method for generating robust strategies and narrative scenarios. *Management Science*, *52*(4), 514–528.

Marzuoli, A. (2008). Toy models in physics and the reasonable effectiveness of mathematics. In *Deduction, computation, experiment* (pp. 49–64). Milan: Springer.

Rubin, J. Z., Pruitt, D. G., & Kim, S. H. (1994). *Social conflict: Escalation, stalemate, and settlement.* Mcgraw-Hill Book Company.

Von Briesen, E., Bacaksizlar, G., & Hadzikadic, M. (2017). Modeling Genocide at the System and Agent Levels, CAPS 2017 paper.

Modeling Social Media Effect on Civil Revolutions

Nazmiye Gizem Bacaksizlar, Mirsad Hadzikadic[A]

Abstract

This study is a modeling experiment of the social media effect on civil revolutions with preliminary results. The purpose of this study is designing a minimal and most general possible protests model to allow people to test their perspectives. The literature review on modeling civil revolutions presents three interrelated aspects. First, an agent-based modeling (ABM) approach is mentioned, as there is increasing interest in computational modeling of human systems. ABM helps us to assess the dynamics of civil revolutions. We review ABM methodology in terms of its benefits for understanding micro- and macro-level behaviors of the problem. Second, we present current computational ABM models and we highlight the analysis of their comparison.

This is important for future work since the contribution of our research links to previous studies. The main research question is "How does the social media affect protesters?" Therefore, we investigate social media and its effect on protester networks. To evaluate technological improvements, the third part of the literature review uses two real events from political protests in Egypt and Spain, including the outcomes in terms of social media use. We show these three interrelated parts using a system dynamics (SD) view to combine this complex problem's micro- and macro-level dynamics. In addition, we introduce the preliminary results of the base model.

Keywords: civil revolutions, system dynamics, agent-based modeling, social media

[A] Software and Information Systems, University of North Carolina at Charlotte, Charlotte, NC, USA

doi: 10.18278/jpcs.3.2.2

摘要

针对社交媒体对国民革命的影响，本研究制订了一项模型实验，并得出了初步结果。本文目的是设计一项最简易，同时又最一般化的抗议模型，用于人们检测各自的观念。对此模型的文献评论提出了三个相互关联的方面。第一，由于人类系统的计算模型（computational modeling）受到越来越多的关注，本文提及了一项基于主体的建模（Agent-Based Modeling，简称ABM）方法。ABM能帮助评估国民革命动态。本文检查了ABM方法论的好处，以理解问题的微观行为和宏观行为。第二，本文呈现了当今的计算ABM模型，同时强调了模型的对比分析。 这对未来的工作十分重要，因为本文的研究贡献与之前的研究相关。主要的研究问题是"社交媒体如何影响抗议者？"。因此，本文调查了社交媒体及其对抗议者网络的影响。第三，文献评论以埃及和西班牙的政治抗议为例（包括社交媒体产生的结果），评价了技术上取得的进步。本文以系统动态学（System Dynamics，简称SD）的观点展现这三个相互联系的方面，从而将复杂问题的微观动态和宏观动态结合起来。此外，本文还得出了该基本模型得出的初步结果。

关键词：国民革命，系统动态学，基于主体建模，社交媒体

Resumen

Este estudio es una modelización experimental del efecto de las redes sociales en las revoluciones civiles con resultados preliminares. El propósito de este estudio es diseñar un modelo de protestas mínimas y lo más generales posible para permitir que la gente ponga a prueba sus perspectivas. La revisión completa de la modelización de revoluciones civiles presenta tres aspectos interrelacionados. Primero, un modelo basado en agentes (ABM) se menciona ya que hay un creciente interés en la modelización computacional de sistemas humanos. ABM nos ayuda a evaluar la dinámica de las revoluciones civiles. Examinamos la metodología ABM en términos de sus beneficios para entender los comportamientos a nivel micro y macro del problema. Segundo, presentamos modelos ABM computacionales actuales y resaltamos el análisis para su comparación.

Esto es importante para las labores futuras porque la contribución de nuestra investigación tiene vínculos con estudios previos. La pregunta principal de la investigación es "¿Cómo afectan las redes sociales a los manifestantes?" Por lo tanto, investigamos las redes sociales y sus efectos en las redes de manifestantes. Para evaluar mejoras tecnológicas, un tercio de la revisión total utiliza dos eventos reales de protestas políticas en Egipto y España, incluyendo los resultados en términos de uso de redes sociales. Mostramos estas tres partes interrelacionadas utilizando una visión de la dinámica de sistemas (SD) para combinar la dinámica a nivel micro y macro de este complejo problema. Adicionalmente, presentamos los resultados preliminares del modelo.

Palabras clave: revoluciones civiles, dinámica de sistemas, modelización basada en agentes, redes sociales

Introduction

Agent-based modeling (ABM) helps us build an artificial environment for modeling problems with complicated behaviors. This simulated environment can refer to the society, nature, and any type of relationships among their agents, which have real meaning. For example, we can define people, viruses, cars, and roads each as an agent in the ABM. Users decide on agent's behaviors, agent's degree of rationalities, and rules of agent's interactions with other agents, who have memory, can learn, and can evolve.

In ABM, there are micro- and macro-level dynamics. Each agent has its own attributes to decide on its actions, which create interactions in both space and time (Miller & Page, 2004). The agent's interactions have a one to one relationship in the model and create micro-level dynamics. Agent-based interactions might lead to a global emergent behavior that can affect the macro level dynamics, such as a social context with socio-economic conditions, power dynamics in the country, and political situations. These macro-level dynamics also have an effect on micro-level dynamics. It can be fair to call this a "bottom-up approach" meaning interactions happening at the bottom level influence the interactions at the top.

We apply the system dynamics approach to understand the big picture of the problem (SD). This approach helps us to link subparts of the problem and their relationships to see the whole dynamics. Each arrow represents causal relationships that mean a variable in the beginning of the arrow has an influence on other variable in the end of the arrow (Sterman, 2000). This influence can be positive (+) or negative (−). Positive sign means that an increase in the variable at the beginning of the arrow can lead to an increase in the variable at the end of the arrow. On the other hand, negative sign indicates that an increase in the variable at the beginning of the arrow can lead to a decrease in the variable at the end of the arrow. If there is an ambiguity in relationships, we leave the arrow's sign blank.

There is a feedback loop dynamics between micro- and macro-level relationships. Figure 1 shows the feedback dynamics to explain why we apply ABM and SD approaches together to work on our model. Micro-level dynamics can create emergent behaviors, which develop macro-level dynamics to influence agent's interactions.

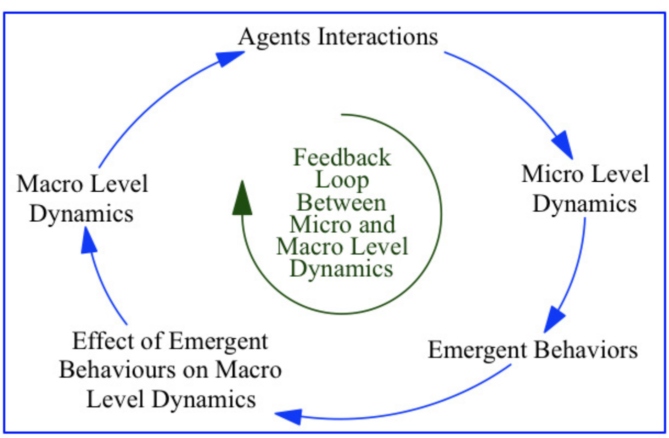


Figure 1. Feedback loop between micro- and macro-level dynamics of the ABM.

This paper presents a survey of ABM role in understanding civil revolutions dynamics. First, we introduce the benefits of ABM with human systems. Then, the second part compares computational model examples for civil revolutions. In addition to this comparison of computational models, we discuss a brief literature review on the social media effect on civil revolutions.

At the end of our research, we hope to find answers for the following questions:

- What kind of parameters do we need to consider while modeling civil revolutions?
- Are there any specific conditions that produce civil revolutions?
- What are the commonalities between different instances of successful protests or civil revolutions?
- How can we verify our model results?
- How does the structure of political protests change?
- How does the social media affect the evolution of civil revolutions?
- How does homophily affect the social media usage?

Method

Simulating Civil Revolutions with Agent-Based Modeling

In this section, we discuss in detail the benefits of applying ABM in simulating human systems. According to Bonabeau (2002), there are three main benefits of ABM. First, ABM captures emergent phenomena created by the interactions of agents (Bonabeau, 2002). Second, ABM provides a natural description of the system that makes the model closer to the real world, making it easier to interpret its dynamics compared with other traditional modeling approaches (Bonabeau, 2002). In addition, ABM can run data-driven simulations. No matter how big the data size is users can implement this data onto their model. Third, ABM is flexible in terms of changing agent's attributes and framework of the model, meaning that it is faster to create experimental results with scenarios and policy analysis in ABM (Bonabeau, 2002). With ABM software such as the free and accessible NetLogo, users can build models (Wilensky, 2016).

In this study, the main interest is modeling the effect of social media on civil revolutions. We also introduce a literature review on computational modeling of civil revolutions. Civil revolutions, in other words protests or demonstrations, come from the citizens who believe their actions are for the common good. For our research, it is important to understand the dynamics of the protests before, during, and after their occurrences.

ABM of civil revolutions. There is rich literature on ABM simulations of social civil revolutions. The first published paper about ABM of civil riots is from Epstein (Epstein, 2002). He developed a rebellion model that considers simple threshold-based rules to represent collective behavior and contagion effects. There is a threshold for agents to join a protest or riot. If other agents join the protests, its grievance level exceeds the threshold, and it turns to an active one. The active agents are in interaction with police officers, who serve the government. The police officers can arrest the protesters, who may face jail terms.

There is another parameter called government legitimacy in the model. It reflects the agents' trust of their government and its rules. Every agent wants to maximize its utility function according to its internal state, agent interactions, and environment. When its grievance level is higher than its threshold and assuming its grievance level exceeds its net risk perception, the agent wants to move and protest (Lemos, Coelho, & Lopes, 2013).

After Epstein's main contribution to the field, there are some variations of his model. Lemos et al. (2013) has a wide survey paper on civil revolutions models. Therefore, we would like to cite their work in order to show all of the former papers on civil revolts until 2012 (Lemos et al., 2013).

Here, three civil revolutions papers, published after 2012, are introduced.

Comer and Loerch (2013) suggest a new way for modeling agent's actions. In their model, agents are not moving uniformly as they were in Epstein's model. For Comer and Loerch (2013) model, agents move either asynchronously or randomly. For the asynchronous, agents' neighbors influence agents' attributes. At first, an agent moves to a state and checks its conditions by counting the numbers of active agents and cops in its vision radius. Then, according to its neighborhood, it looks at its own threshold and decides whether to act. This action is dependent on the agent's neighborhood. For the random, an agent makes decisions on whether or not its next action is active. This paper opens an additional discussion for our research: if we have a social media effect on our agents, then there will be no constraint on their visions; therefore, we can state that the social media effect will enlarge the agent's vision.

New types of agents and their roles listed correspondingly by Lemos, Lopes, and Coelho (2016). Here, the number of previously active protesters who are in jail affects an agent's grievance level. There is a feedback process in this model. If the number of jailed active agents increases, then government legitimacy decreases, and the agent's grievance level subsequently increases. For Lemos et al. (2016), the point is more realistic than Epstein's model. Instead of having a random distribution for grievance levels of the agents, they suggest an internal dynamic for an agent's activations process. We would like to add this function into our model with the help of social media.

Compared with others, Moro (2016) has the greatest contribution to the literature with his model, as it can generate real world examples of civil revolutions like the Arab Spring (Moro, 2016). His model predicted several rebellions that arose in similar ways with different scenarios. Outcomes of his model show different political conditions, such as the successful revolution in Tunisia, the failed protests in Saudi Arabia and Bahrain, and civil war in Syria and Libya. His model can create similar dynamics to these real civil revolutions.

In his model, there are three types of agents to characterize citizens: active (A), jailed (J), and revolutionary (R). In addition, there are police officers (P) who fight with the active citizens, who turn into active revolutionary citizens if they intend to kill police officers in their vision radius. Moro suggests three scenarios with different parameters for the probability of killing police: successful revolution, anarchy, and failed revolution.

When a revolutionary citizen is active, he kills a randomly selected police officer in his vision radius with a probability equal to r. For the police officers, if the randomly selected agent is a citizen, he arrests him; if he is a revolutionary citizen, he kills him with a probability equal to p (Moro, 2016). We give Moro's probabilities in Table 1.

Table 1. Probabilities in different scenarios (Moro, 2016)

Probability in different scenarios		
Scenario	***p***	***r***
Successful revolution	0.4	0.3
Anarchy	0.9	0.3
Failed revolution	0.9	0.1

Figure 2 illustrates outcomes of three scenarios. In the successful revolution (p=0.4, r=0.3), after time step 40, the number of As is higher than the number of Js, and there is an exponential decrease in the Js. For the police officers and revolutionary citizens' dynamics, initially the number of P is higher than the number of Rs. However, after approximately time step 40, the number of P decreases exponentially. On the contrary, in the failed revolution scenario (p=0.9, r=0.1), the number of Js is higher than the As. Eventually, the As become zero. P kills most of the Rs where there are not many As. Interestingly, Moro (2016) found that the civil revolution dynamics in the case of anarchy show complex behavior. As and Js create a different pattern compared with the failed revolution scenario.

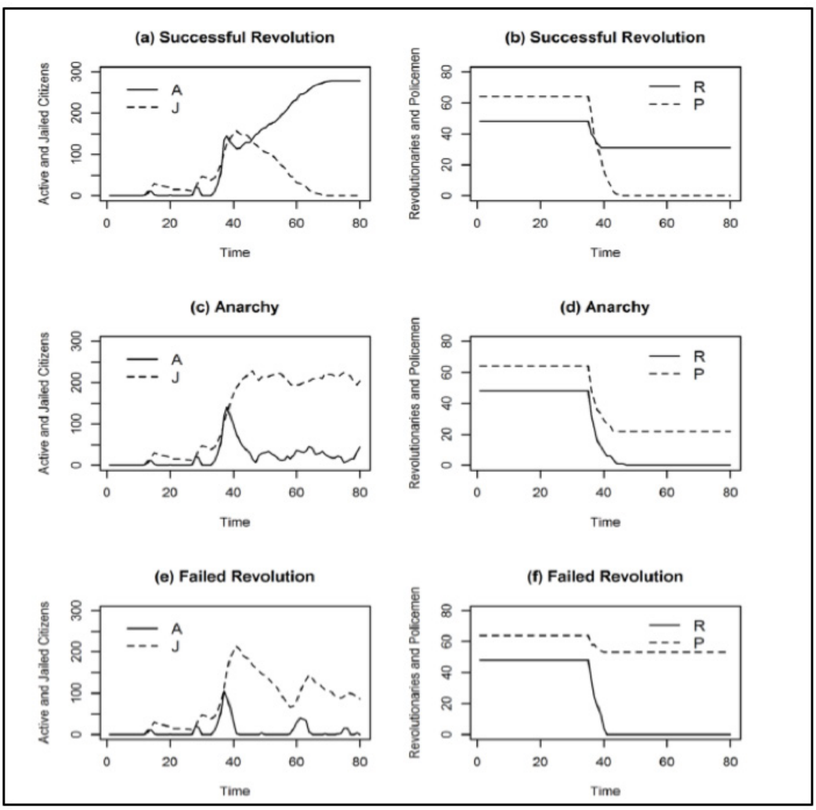


Figure 2. Time series for ABM scenarios.

Time series graph for the different model scenarios: (a) time series of the number of active and jailed citizens in a successful revolution; (b) time series of the number of revolutionaries and police officers who have survived a successful revolution; (c) time series of the number of active and jailed citizens in an anarchic scenario; (d) time series of the number of revolutionaries and police officers who have survived an anarchic scenario; (e) time series of the number of active and jailed citizens in a failed revolution; and (f) time series of the number of revolutionaries and police officers who have survived a failed revolution (Moro, 2016).

To simplify information in these four papers, Table 2 highlights their specific contributions. It shows information of author and paper's names, and their contribution to the field. As mentioned earlier, we compare these papers with Epstein's model and present with their strong points.

Table 2. Comparison of civil revolutions with ABM after 2012

Civil revolution models		
Author	**Paper**	**Contribution**
Epstein (2002)	Modeling civil violence: an agent-based computational approach	The first agent-based model for civil revolutions. Agents are moving according to their grievance levels and risk thresholds
Comer and Loerch (2013) [6]	The impact of agent activation on population behavior in an agent-based model of civil revolt	They introduce technical improvements on agents' actions
Lemos et al. (2016) [7]	On legitimacy feedback mechanisms in agent-based modeling of civil violence	Agent's grievance level is dependent on jailed citizens
Moro (2016) [8]	Understanding the dynamics of violent political revolutions in an agent-based framework	Replications of real cases form the civil revolutions are applied

Understanding social media effect on social movements. With the ease of spreading information with the Internet and smart phone use, scholars began to pay more attention to the effect of social media on current political protests. For our research, there are three questions that arise: How does social media bring people together? Does social media have an influence on participation rate to the protests? When there is Internet censorship, how can protesters keep their communication channel alive?

In the light of these questions, we introduce related works from social sciences. Centola (2013) suggests that in large groups, it is hard to find the initial source of power or specific contributor that serves as the catalyst. However, there can be a "bandwagon" effect that increases group cooperation. This means that when the first wagon starts to move, the others follow its lead. He provides examples of mass actions displaying this effect, such as large strikes, political protests, and violent revolutions. Centola's point of view is helpful in constructing the main loop of our research, which is the effect of social media on protester interactions (see Figure 3).

Figure 3 shows the feedback loop for the Homophily Effect on Protests. Homophily is the tendency for people to have social connections with people who are similar to them (Centola, 2013). If the number of people who share similar grievance levels against the current regime increases, then the number of protesters will increase. Therefore, there will be more news about protests in social and mass media. With the help of this increase in communication channels through social media, more people with common interests or views about news can more into action. Their posts and/or tweets will increase the awareness on communication channels. People with similar interests might also start following each other on social media. This process can create increased homophily in society. There is a positive feedback loop for the effect of homophily on the number of people who share similar grievance levels against the current regime and government.

The attention toward social media effect on civil revolutions has increased since the Arab Spring. This subject has become interesting for scholars especially in social sciences. Little (2015) builds

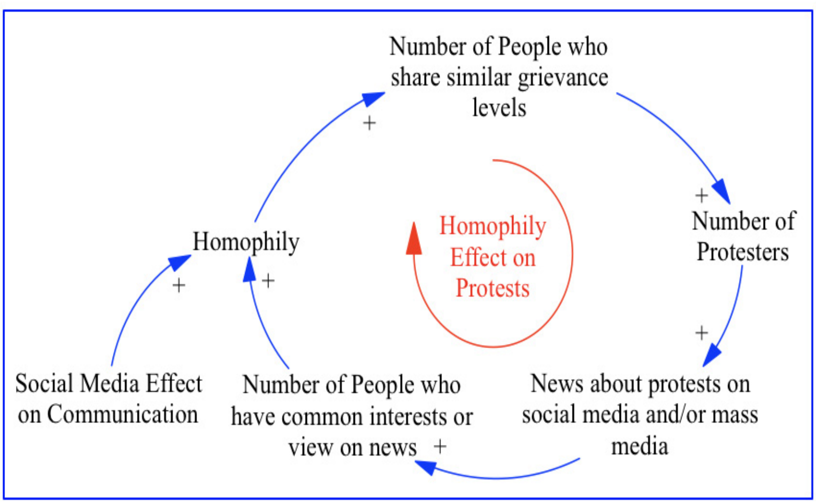


Figure 3. Feedback loop in our model for homophily effect on protests.

models for technological information channels that impact two factors: the level of dissatisfaction with the regime and logistical information about when, where, and how potential protests will occur. His models are also beneficial in drawing our main feedback loop shown in Figure 3.

According to Little (2015), public grievance levels against the regime have an ambiguous effect on protest levels (Little, 2015). There could be more or less popular effect than expected. Although sharing more complaints about the regime can increase average grievance levels against the government, there is no exact correlation between anti-regime content and protests. However, we draw this relationship with the effect of homophily as a positive relationship and an increasing factor in protest movements.

Social Media Effect on Civil Revolutions: Political Protests Examples with Two Real Cases. As mentioned earlier, since the Arab Spring uprisings in Tunisia and Egypt in early 2011, scholars have researched the Internet and social media contribution to political change in authoritarian regimes. Tufekci and Wilson's (2012) research "Social Media and the Decision to Participate in Political Protest" shows observations from Tahrir Square as events unfolded in real time (Tufekci & Wilson, 2012). This paper is evidence of how the protestors used social media and the Internet.

To understand Egypt's conditions better, Tufekci and Wilson (2012) present the timeline of the revolution. From 2005 on, there were persistent protest efforts by a small but dedicated group in Egypt. In 2009, Facebook introduced its Arabic language version. This amplified communication opportunities and in 2010, Tunisian unrest broke out. Finally, in 2011, the Egyptian revolution began after 18 days of sustained protests (Tufekci & Wilson, 2012).

For Tahrir, there was a new system of political communication due to technological improvements. This system involved three broad and interrelated components: (1) an independent TV channels (Al-Jazeera), (2) the rapid diffusion of the Internet and the rise of social media (Facebook and Twitter), and (3) the decreasing costs and the expanding capabilities of mobile phones (Tufekci & Wilson, 2012).

Tufekci and Wilson (2012) conducted a survey of media use by Egyptian protesters. They interviewed 1,050 people using both random and snowball sampling approaches. They demonstrated that "people learned about the protests primarily through interpersonal communication using Facebook, phone contact, or face-to-face conversation" (Tufekci & Wilson, 2012). According to their statistical analysis, controlling for other factors, social media use had a positive impact on protest attendance on the first day and Facebook is the most common communication source among surveyed people.

The results of their study underline the central role of social media, particularly Facebook and Twitter, which played a role in the protests leading up to the February 2011 resignation of Egyptian President Hosni Mubarak. As mentioned, although Facebook only became available in Arabic in 2009, protesters shared most of their pictures and videos about the protests through Facebook. On the other hand, protesters used Twitter and blogs to communicate text about the demonstrations (Tufekci & Wilson, 2012).

Another political protest example that we want to introduce in order to provide a worldwide perspective is the 15 M protests in Spain, held on 15 May 2011 (Anduiza, Cristancho, & Sabucedo, 2014).

They occurred at a similar time as other international uprisings such as Occupy Movements, the protests in Greece, the Arab Spring, and the Egyptian Revolution. Around 130,000 people gathered to demand democracy in Spain using the slogan "Real Democracy Now!" The protesters' grievance levels against the government had increased due to the national economic crisis, the high unemployment rate, and the government cutbacks in health care and education. These protests were different from the others that happened in Spain because of their self- organized nature. There were no political parties, unions, or large organizations behind the protests.

The authors consider the 15 M demonstrations an instance of a "self-organized connective action network with significant differences" when compared with other collective action events in Spain (Anduiza et al., 2014). They conducted interviews and postal surveys with protesters. In each demonstration, the researchers contacted the protesters randomly and interviewed them face to face. After that, they handed out a postal questionnaire. Response rates to the postal questionnaire ranged from 18% to 33%, and reached 35% for the 15 M demonstration. They had 2,265 protesters complete the postal survey.

As mentioned before, the structure of 15 M was different with respect to its mobilization processes. Almost 55% of the participants heard about the demonstration through social media and alternative media sources, and social networks informed 49% of them. In other protests, these percentages were 26% and 17%, respectively (Anduiza et al., 2014). The role of traditional media as an information channel was very limited.

Anduiza et al. (2014) show with their comprehensive comparative study on Spain demonstrations, the traditional intermediary structures, such as unions, parties, and traditional news are still important for large-scale political mobilizations. However, in 15 M case, the researchers claim that union, party, and traditional media involvement were no longer a necessary initial condition for generating high turnouts at protests because the social media network had an enormous role.

These two papers show the results of real events from the European and Middle Eastern countries. Both events occurred around a similar time, and we can use these works to derive theories about technological improvements, focusing on their huge impact on the structure of political protests. We hope to consult these studies in order to improve our theoretical and practical background in the field, and to determine the essential parameters for civil revolutions. Examples would be the nature of a regime that the protesters are for or against (democracy versus autocracy), protester types (experienced versus new protesters), and social media preferences (Facebook versus Twitter).

Overview of the Preliminary Model for Protests with the Social Media Effect with the System Dynamics View

This section introduces our preliminary work for modeling protests with social media's effect. The goal of our study is to build a model that helps us understand the relationships between the level of cops' violence and the number of protesters censored who are on social media/ communication channels.

Figure 4 displays the causal loop diagram of our work that explains the relationships on the system level. There are three main effects in the model. The first one is the homophily effect on protests, which is explained in the previous section. This effect considers grievance levels among protesters that increase with the creation of homophily using social media. We did not consider this loop in our preliminary model. Second, one examines cops' violence, which has an adverse impact on the number of protesters. If there is an increase in the agent's grievance level against government actions, then agents are more willing to protest. Therefore, the number of protesters who are against the government rises. To stop the protesters, the cops increase their violence level, which then affects agents' risk aversion. When the agents are more risk averse, they are not as eager to protest. However, the different natural characteristics of agents are of note; they have randomly distributed grievance levels, which are set initially.

The third effect explains censorship in the communication channels, which aid agents' interactions. The government does not want protesters to reach each other and share demonstration information. Thus, the government attempts to restrict communication channels such as Twitter, Facebook. Through censorship, the government cuts ties between agents and their friends who are already protesting; therefore, censorship has a negative effect on protests.

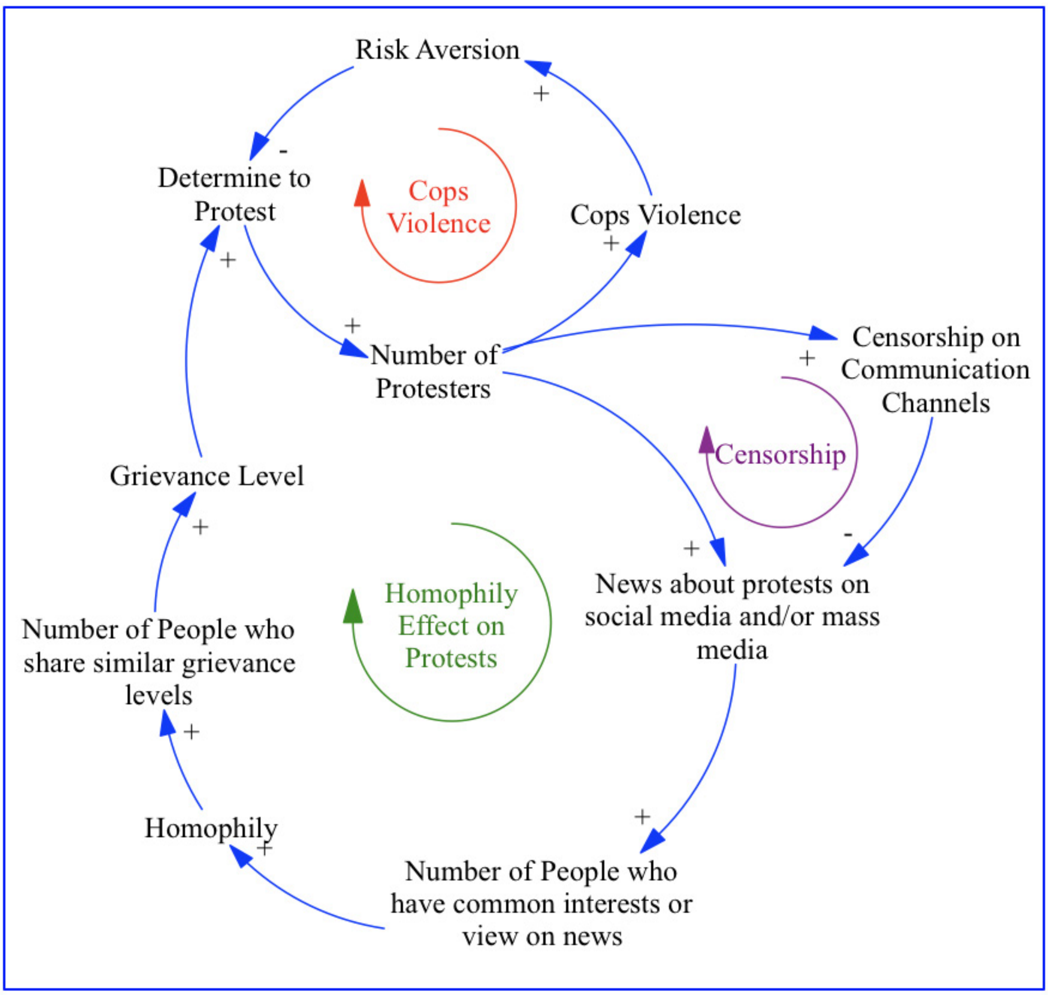


Figure 4. Causal loop diagram of the model with effects of cops' violence, censorship, and homophily on protesters.

Preliminary Model Explanation with Agent-Based Modeling

Each agent aims to protest according to their grievance levels against the government's actions and the environment, and she protects herself according to her risk aversion. Agents interact with each other if there is an available communication channel; however, cops have an aim to stop protesters by increasing their violence level. They do not kill the protesters, but by increasing their violence level, they attempt to dissuade protesters from demonstrating.

We assign each agent's grievance levels randomly and there is a diverse group of agents in terms of grievance level and risk aversion. Cops interact with the number of protesters. If there is an increase in the ratio of the number of protesters to the number of cops, then the cops' violence level rises. A cop's violence level affects an agent's risk aversion, which means that this is an internal variable for agents. When the cop's violence increases, agents become more risk averse.

We construct this model to answer two questions: What is the impact of cops' violence on the ratio of protesters among the population? What is the impact of censorship on the ratio of protesters among the population?

We build this simulation experiment of the ABM using NetLogo software. The global variables are max-cops-violence, threshold to determine protest, and censorship. The model has two main procedures: setup and go. We introduce the steps for these procedures in Table 3.

Table 3. Setup and go procedures in the preliminary model

Setup and go procedures	
Procedure	**Steps in respective procedure**
Setup	—Set the demonstration location (park) —Clear all from the previous run —Setup agents and cops
Go	—Ask agents: to determine protest (if yes, go to park), to update decision, and to give up —Ask cops: for protests with more than three protesters, go to park and increase the violence until there are no protesters

Preliminary Model Procedure Explanation

We present the model procedures for agents and cops in this section. For potential protesters, there are four procedures, which begin with the procedure labeled Go in Table 3. Movements of agents relate to the friend and cop density.

For cops, there are three procedures, which we call from Go procedure with the respective order in it (see Table 5).

Results

The simulation model is arbitrary; therefore, we base our results on early observation of this preliminary model. Figure 5 demonstrates the simulation environment for protests at the park, which we show in green. Blue stars indicate cops and agents appear as people shape of varying colors according to their protest modes (refer to Table 4 for color codes).

Table 4. Procedures for agents in order for the preliminary model

Procedures for agents in order	
Procedure	**Steps in respective procedure**
Setup-agents	—Create initial number of agents with color yellow —Assign a random grievance level and risk aversion (normally distributed) —Assign a vision for reaching protester friends
IF *(1 − censorship + no. of protester friends/no. of all possible protester) × (grievance level − risk-aversion) > threshold*	
Determine- protest	—Go to park, if condition is hold.
Go-to-park	—Go to park to protest —Turn the agents color red
IF *(1 − censorship + no. of protester friends/no. of all possible protester) × (grievance level − risk-aversion) < threshold*	
Give-up	—Give up and come back with color pink, if condition is hold.

Table 5. Procedures for cops in order for the preliminary model

Procedures for cops in order	
Procedure	**Steps in respective procedure**
Setup-cops	—Create initial number of agents with blue star shape —Assign a random violence according to the max-cops-violence —Assign a vision for reaching protesters
IF *(no. of protesters)/(no of cops) > 3*	
Go-to-park-cops	—If condition is hold, cops go to the park to fight with protesters.
Increase-violence	—If condition is hold, increase the violence level to the max-cops-violence until protesters leave the park.

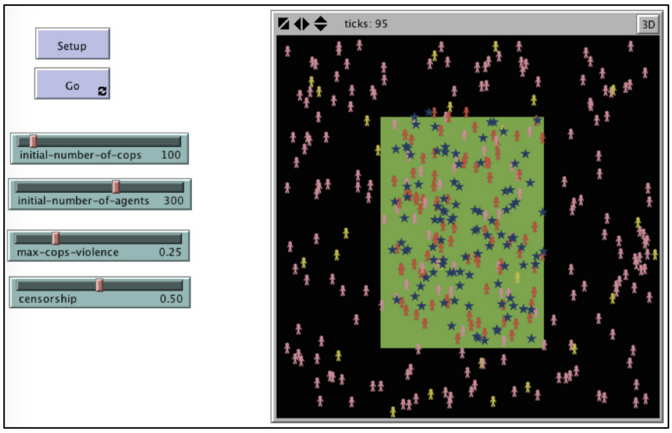


Figure 5. ABM environment for base run in the preliminary model.

Outcome One

We analyze the ratio of protesters under several scenarios with different agent's grievance level distribution and different threshold levels for the ratio of protesters to cops. We test randomly uniform, randomly normal, and randomly exponential distributions during the experimental phase. Figure 6 and Table 6 show the protester ratio among all civilians under each scenario. In most of the scenario runs, agents experience oscillatory behaviors due to interactions among agents; however, the amplitude of the oscillations differs in some cases such as case *i* and *c*. The case with normally distributed grievance levels (c) has a noisier graph under the highest threshold compared with the case with exponentially distributed grievance levels (i). Exponentially distributed grievance levels can create more realistic results (in terms of stability), but it should be considered that the threshold level might also have an effect on these oscillatory behaviors. When the threshold is at its lowest value, which means cops move as soon as they see one protester, agents become more risk averse and are not as eager to protest. Therefore, for these cases (*b, e, h)*, the ratio of protesters among civilians is lower than in other cases.

Table 6. Different agent's grievance level distribution and different threshold levels for the ratio of number of protesters over the number of cops

	Threshold		
Distribution	3	1	5
Random-normal	a	b	c
Random-uniform	d	e	f
Random-exponential	g	h	i

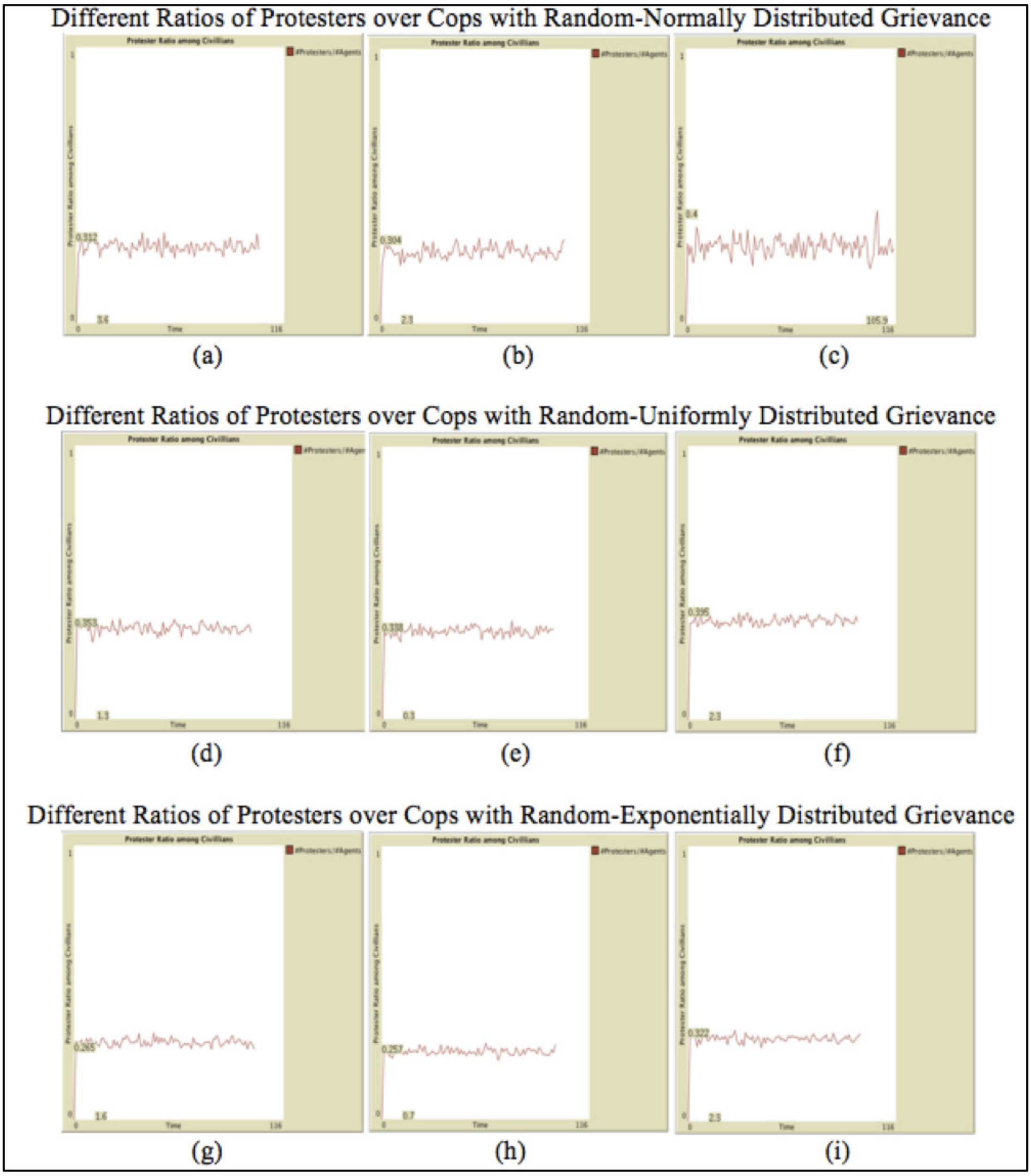


Figure 6. Protesters ratios under different scenarios.

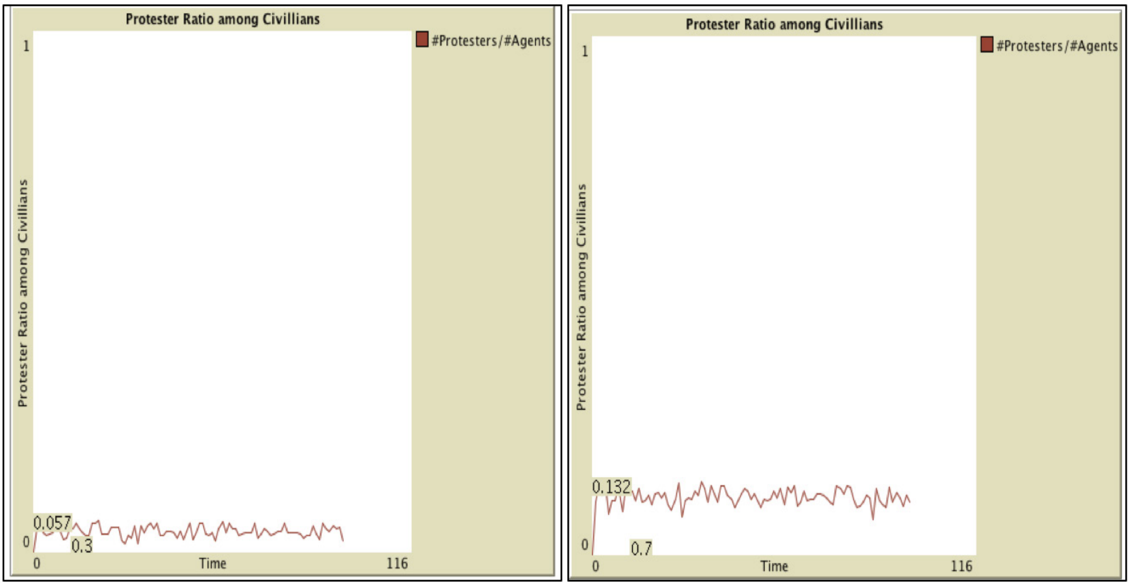


Figure 7. Protesters ratios under the scenarios with censorship and without censorship with the same cops' violence levels.

Outcome Two

Figure 7 displays the scenario for low censorship levels on social media and a corresponding high level of cop violence. In this scenario, the protester ratio among civilians decreased with an increase in cop violence. However, this ratio is higher than the one for protests after an increase in censorship.

Conclusion and Future Work

This study has two goals. The first is to discuss current studies about civil revolutions represented in the literature, which is vast and growing. It helps us to investigate other studies and to understand the ways for modeling social media's effect on civil revolutions with the ABM approach. The second goal is to introduce our preliminary model for protests with social media's effect presented at the micro-level with ABM and at macro level with System Dynamics view conceptually.

Epstein's model continues to play a significant role in civil revolution modeling. As a base model, we want to apply his ideas for protesters grievance levels, risk thresholds, and government legitimacies, however; the tendency towards using social media is included. Other variations of Epstein's model are helpful in improving the literature and investigating the missing aspects of rebellion models. These aspects are changing over time because of the unpredictable and complex nature of humans and the continual rise of technological improvements.

In our model, there is a relationship between social media and agent interactions. We assign agent's grievance levels randomly, however; internal dynamics should affect these levels, and we will add this effect in our model. Additionally, the enlargement of agent's visions occurs due to the effect of social media.

We will also build our model with the SD approach to see the micro-level dynamics effect on the overall dynamics and to compare the results with ABM outcomes. Furthermore, we believe that SD outcomes will help us to derive an accurate analysis in our ABM approach because of its holistic power in modeling causal relationships. Our research will add value to the computational modeling area of civil revolutions. We will consider the social media's effect on the relationship between grievance levels and risk threshold levels in detail and, by doing so, improve the results of our models.

References

Anduiza, E., Cristancho, C., & Sabucedo, J. M. (2014). Mobilization through online social networks: The political protest of the indignados in Spain. *Information, Communication & Society*, *17*(6), 750–764.

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the USA*, *99*(3), 7280–7287.

Centola, D. M. (2013). Homophily, networks, and critical mass: Solving the start-up problem in large group collective action. *Rationality and Society*, 25, 3–40.

Comer, K. W., & Loerch, A. G. (2013). The impact of agent activation on population behavior in an agent-based model of civil revolt. *Procedia Computer Science*, 20, 183–188.

Epstein, J. M. (2002). Modeling civil violence: An agent-based computational approach. *Proceedings of the National Academy of Sciences of the USA*, *99*(3), 7243–7250.

Lemos, C., Coelho, H., & Lopes R. J. (2013). Agent-based modeling of social conflict, civil violence and revolution: State-of-the-art-review and further prospects. *CEUR Workshop Proceedings*, 1113, 124–138.

Lemos, C., Lopes R. J., & Coelho, H. (2016). On legitimacy feedback mechanisms in agent-based modeling of civil violence Carlos. *International Journal of Intelligent Systems*, 31, 106–127.

Little, A. T. (2015). Communication technology and protest source. *The Journal of Politics*, 78.

Miller, J. H., & Page, S. E. (2004). The standing ovation problem. *Complexity*, *9*(5), 8–16.

Moro, A. (2016). Understanding the dynamics of violent political revolutions in an agent-based framework. *PLoS One*, *11*(4), e0154175.

Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world* (No. HD30. 2 S7835 2000).

Tufekci, Z., & Wilson, C. (2012). Social media and the decision to participate in political protest: Observations from Tahrir Square. *Journal of Communication*, *62*(2), 363–379.

Wilensky, U. (1999). NetLogo (Version 5.3.1) [Software]. Retrieved from http://ccl.northwestern.edu/netlogo/

Modeling Genocide at the System and Agent Levels

Elizabeth M. von Briesen, N. Gizem Bacaksizlar, Mirsad Hadzikadic[A]

Abstract

This research works to bridge the gap between the knowledge of social scientists in the field of genocide studies and that of computational experts. A thorough system-level description of the problem informs the implementation of an agent-based model (ABM) that simulates local interactions in a society composed of two identity groups. This model explores how dynamics between agents with different identities, ideology, influence, susceptibility, and threshold to act against those of a different group lead to the emergence of genocide. Early results indicate the model's usefulness in revealing the underlying factors and thresholds of importance. Next, this study presents steps toward the implementation of a formal system dynamics model, and describes how its results may complement those of the ABM, leading to more comprehensive and effective policy recommendations. Given further refinement, this research has the potential to produce a useable, generic model for social scientists and genocide researchers.

Keywords: genocide, system dynamics, agent-based modeling, complex systems

[A] Complex Systems Institute, University of North Carolina at Charlotte

doi: 10.18278/jpcs.3.2.3

摘要

本文致力搭建科学家和计算专家对种族灭绝研究认识差异间的桥梁。通过对问题进行全面彻底的系统性描述，实施了一项基于主体的建模（Agent-Based Modeling，简称ABM），该模型模拟了由两种身份团体相互影响而组成的社会。模型探索了不同主体间动态如何导致种族灭绝的发生（这些主体拥有不同的身份，观念，影响力，易感性，以及对抗另一团体所需的条件）。早期结果发现，该模型有助于揭示重要性的潜在因素和阈值。之后，本文呈现了实施正式系统动态模型所需的步骤，同时描述了该模型结果会如何对ABM结果进行补充，从而促成更多全面且有效的政策建议。考虑到本研究的进一步改进，社会科学家和种族灭绝研究者将有可能使用可行的一般性模型。

关键词：种族灭绝，系统动态学，基于主体建模，复杂系统

Resumen

Esta investigación busca llenar el vacío entre el conocimiento de los científicos sociales en el área de los estudios de genocidio y el de los expertos computacionales. Una descripción detallada de los niveles del sistema del problema informa la implementación de un modelo basado en agentes que simula las interacciones locales en una sociedad compuesta por dos grupos de identidad. Este modelo explora cómo la dinámica entre agentes con diferentes identidades, ideologías, influencias, susceptibilidades y límites para actuar contra aquellos de un grupo diferente, conducen al surgimiento del genocidio. Los resultados preliminares indican la utilidad del modelo para revelar factores de causa y límites de importancia. Después, este estudio presenta pasos hacia la implementación de un modelo de dinámica de sistema formal y describe cómo sus resultados pueden complementar aquellos de un modelo basado en agentes, lo que lleva a recomendaciones más comprensivas y efectivas para la política. Si se le da un refinamiento adicional, esta investigación tiene el potencial de producir un modelo genérico útil para los científicos sociales e investigadores de los genocidios.

Palabras claves: genocidio, dinámica de sistemas, modelización basada en agentes, sistemas complejos

Introduction

The phenomenon of genocide is complex, as is the task of deriving its definition. According to Verdeja (2010), it will be difficult to compare theories on genocide until researchers decide on the definition of genocide. While the above task of definition primarily falls upon the shoulders of social scientists and the entities with which they work, it is important in the context of this computational effort to have clear direction. Verdeja (2010) skillfully surveys a number of related works and specifies what he and others see as flaws in the United Nation's definition of genocide—the primary issue being its reference only to ethnic, racial, national, and religious groups. He points to the lack of recognition of political, economic, and other groups as one of many weaknesses of the definition of genocide; with the lack of a threshold being another weakness in the definition of genocide because there is no criteria to differentiate mass killings from genocide.

Computational modeling typically requires a simple foundation, and given the lack of consensus, we default to Straus' definition of genocide (Straus, 2012b). According to Straus (2012b), genocide is any large-scale territory-specific organized form of destruction that targets specific social groups. The above is a general statement, but its simplicity is appropriate for the domain of computational research. Creating a robust and flexible model that produces interpretable results requires stripping away details and focusing only on the most basic building blocks necessary. Given this definition, the next task is to derive a satisfactory framework from which to build a simulation, and a system (or macro) level perspective is a logical starting point.

This research attempts to gain new insights and understanding into the fundamental drivers of genocide in order to inform preventative policy decisions. Complexity theory guides the analysis, which begins by exploring the problem at the macro level. This approach produces a broad picture of the large-scale dynamics of genocide. However, given complexity theory's focus on local interactions when examining adaptive systems, the model implementation uses a simplified framework in an agent-based model (ABM) for initial exploration of the problem.

Method

System-Level View

A survey of the work of major scholars in the fields of genocide studies shows that the complexity of the problem framework increases with the depth of the research. The causal loop diagram from the field of System Dynamics (SD) provides a useful structure for visualizing genocide's system-level framework.

Figure 1 is a *simplified* causal loop diagram created with Vensim PLE Software (Version 6.3G; Ventana Systems, Inc., 2006). It shows a number of relationships and feedback mechanisms. Each variable has a descriptive name and is connected to others by arrows of specified direction and polarization. A positive sign (+) indicates a positive relationship. This means that an increase in the variable at the beginning of the arrow leads to a corresponding increase in the variable at the end of the arrow, and vice versa. A negative sign (−) denotes the opposite, with an increase at the beginning of the arrow corresponding to a decrease at the end of the arrow, and vice versa. A positive/negative sign (+/−) indicates uncertainty in the nature of the variables'

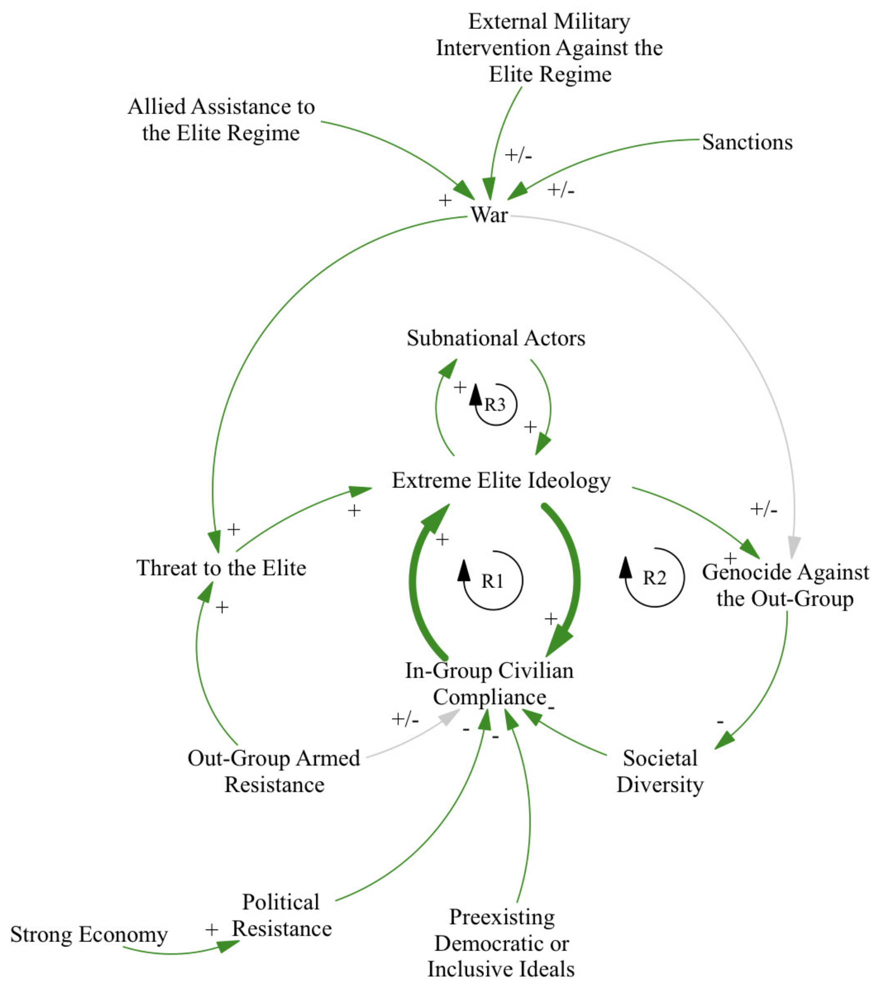


Figure 1. Simplified diagram for effect of extreme ideology on Genocide Against the Out-Group.

relationship. Feedback loops within the diagram show the dynamics of change over time between sets of connected variables.

Extreme Elite Ideology and In-Group Civilian Compliance

The central loop (R1) connects two sets of primary actors. The elite represent the ruling class of the dominant societal group, and the in-group represents civilian actors who are of the same identity as the elite. The dynamic shown between the two is specific to the incidence of genocide. While there is not complete consensus among scholars that the presence of an extreme ideology on the part of the elite is a required condition for genocide to occur, it is often acknowledged by domain scholars to have critical importance (Straus, 2012a; Verdeja, 2002; Waller, 2001).

As shown in Figure 1, this dynamic relationship is characterized by In-Group Civilian Compliance in the presence of an Extreme Elite Ideology through a positive effect in both directions.

This creates a reinforcing feedback loop (R1). While not yet fully validated in this research, it is logical to conclude that in order for an elite regime to commit genocide, they require either the active participation of their civilian population, or its tacit approval. The internal dynamics of the civilian group are not clearly distinct in one direction. Waller (2001) and Straus (2012b) both point to the possibility of support or restraint originating from the in-group. The above variation is captured in the diagram through the rise and fall of In-Group Civilian Compliance.

Genocide Against the Out-Group and Societal Diversity

Further extending the diagram is Genocide Against the Out-Group, where the out-group is of a different identity than the elite and its civilian population. An increase in extremist elite ideology leads to an increased likelihood of genocide, and by extension decreases Societal Diversity. Cox (2017) states that genocide is "…enduringly destructive: The annihilation, partly or wholly, of a group diminishes the diversity and richness of the human race. When a people is eradicated or dispersed and its traditions and culture erased, all of human civilization loses much that can never be regained" (p. 189). While this research does not yet include a full validation of the connection between a decrease in diversity and an increase in civilian compliance, it is probable that its effect is important and should be included such that it closes another reinforcing loop (R2).

Subnational Actors

Connected to the elite are subnational actors (Straus, 2012a). Due to the theoretical importance of subnational interaction with elite actors, they are included in the diagram through a reinforcing feedback loop (R3). Further research results from Straus and others will help better validate and define what is likely to be an important feedback mechanism.

War

The next extension of the diagram is through the inclusion of factors such as restraint and escalation (Straus, 2012b). The majority of scholars in this domain agree that genocide predominantly occurs in the presence of war or other armed conflict. According to Verdeja (2002), rapid and catastrophic social transformation, such as war, military coup, revolution, or economic collapse, are structural elements that often precede genocide. These conditions greatly increase instability in a system, which the elite can then exploit to drive divisive and genocidal agendas (Straus, 2012b).

Threat to Elite and Genocide Against the Out-Group

In the diagram, War has a positive effect on Threat to the Elite, which is in turn has a positive effect on Extreme Elite Ideology. While this research has not yet fully validated the above, the existence of a threat to the elite regime is likely to be an important factor of escalation. The ambiguity of the diagram's labeled effect of War on Genocide Against the Out-Group highlights the possibility that armed conflict may have a direct positive or negative effect on genocide. Territorial advances by external actors could reduce the incidence of genocide by providing safe havens for out-group civilians, although the opposite is possible in the presence of territorial advances made by the extreme elite regime. The rise and fall of Nazi Germany in

WWII, and the corresponding effect on its genocidal activities, provides an obvious example. Further research will help better define, characterize, and validate these dynamics.

War—Other Factors

There are three additional variables in the diagram affect War. It is logical to conclude that Allied Assistance to the Elite Regime has a positive effect, however, in the cases of External Military Intervention Against the Elite Regime and Sanctions, the relationships are more difficult to define. Either could be a factor of escalation or restraint depending upon conditions unique to the circumstances under examination, and this is an area that requires further research.

Internal Societal Factors

The scope of the diagram extends to include several variables internal to the society in question. All are connected directly to In-Group Civilian Compliance, as they have specific mechanisms that directly affect this set of actors. Each is characterized below, and future research here will more thoroughly explore all variables and their relationships.

Out-Group Armed Resistance

In addition to its ambiguous effect on In-Group Civilian Compliance, Out-Group Armed Resistance has a direct and positive effect on Threat to the Elite as shown in the diagram. This relationship, in turn, affects the level of extremism in elite ideology and other connected variables according to their polarity. As indicated by the ambiguous polarity (+/−) shown in the diagram, is less clear how Out-Group Armed Resistance affects In-Group Civilian Compliance. The early theory of this research is that if the resistance is effective, civilians may be less likely to accept the extremist ideology of the elite, and vice versa. This relationship may be difficult to validate, but is certainly worthy of discussion.

Strong Economy and Political Resistance

Straus (2012b) discusses the dynamics of mass violence and economics, and includes in his analysis a focus on the middle class. In the diagram, Straus' theory provides initial justification for a Strong Economy to affect Political Resistance, which in turn may have a negative effect on In-Group Civilian Compliance. Again, more research is required here, but Straus' work is compelling in highlighting the importance of considering these dynamics.

Preexisting Democratic or Inclusive Ideals

Finally, the diagram includes Preexisting Democratic or Inclusive Ideals as a variable that has a negative effect on In-Group Civilian Compliance. Straus (2012b) explores both positive and negative cases of genocide in sub-Saharan Africa from 1960 to 2008. A negative case is specified by circumstances that made genocide appear to be a likely outcome, but the situation resolved without its occurrence. According to Straus (2012b), genocide does not occur in established democracies. The initial hypothesis of this research is that a preexisting set of democratic ideals restrains civilian compliance. If these types of ideals are firmly entrenched in a society, they may cause its people to be less compliant in the face of extreme ideologies both from within their ranks and from the elite.

This completes the initial overview of the scope of the problem of genocide as viewed from the system level. However, any study of a human social system will include nonlinearity by default. This research narrows its focus on the problem through the lens of complexity theory, using the causal loop diagram of the system to determine an optimal starting point.

Agent-Level View

A complex system is one composed of interacting parts in which those interactions sometimes lead to non-linear, emergent, system-level properties. When examining such systems, the researcher must select an appropriate level of granularity. In the case of genocide, one could model only the internal psychological nuances of an individual person in the face of extremism, and at the other end of the granularity spectrum, one could model the dynamics of states as they interact with other states. Regardless of the frame of reference, local interactions are the primary focus in the study of a complex system. Emergent properties are a critical observation, but understanding their causes requires a thorough grasp of micro-level interactions and dynamics.

Genocide emerges from a system of people interacting with other people. The system-level view as shown in Figure 1 informs the next step, which is an implementation of an ABM. In this case, the system could be anything from a nation-state to a local population of a village, but the more important point of focus is on the nature of human interactions.

Figure 2 shows the inner feedback loops of the diagram in Figure 1. This section of the diagram helps visualize a very basic set of dynamic interactions and effects. The ABM implemented in this research simulates the dynamics between elite and civilian actors with respect to ideology and compliance in order to understand genocide's lowest level origins.

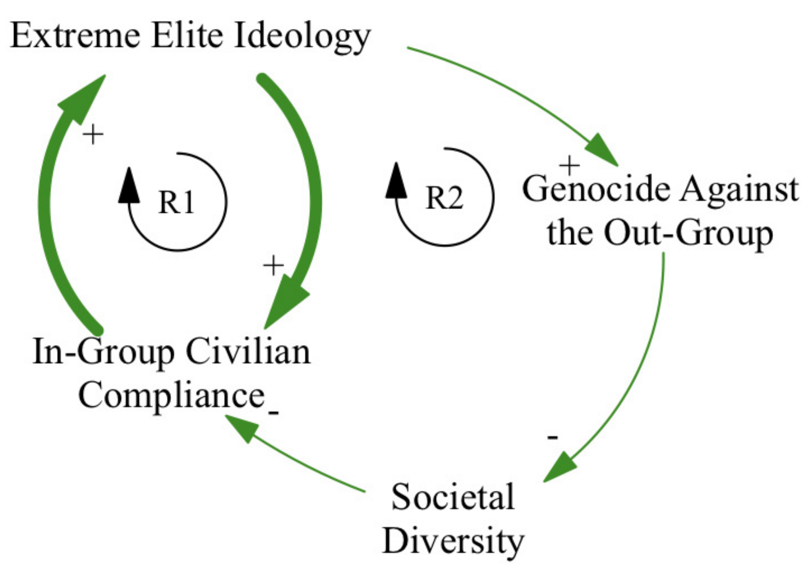


Figure 2. Focal points for an ABM.

Base Model Implementation

This model is implemented using NetLogo (Version 5.3.1; Wilensky, 1999). Structuring an ABM requires clear definitions of the environment, agent attributes, and rules for interaction. Each is defined below.

Environment

As stated above, the environment of the model is a closed system in that there are no external forces in the simulation to which agents respond. All agents have both random initial positions and subsequent movement within the environment.

Agent Attributes

Agents are endowed with six attributes:

1. *Identity* (*Group A or B*): To simulate identity-based conflict, agents are divided into two groups and have a corresponding Identity attribute. Two distinct groups represent the lowest level of societal division, making this the optimal starting point.
2. *Ideology* (*dynamic, 0 – 1, floating-point, random uniform initialization*): A higher Ideology level indicates that an agent has a more extreme and divisive view of agents from the opposing group. A review of genocide studies literature indicates that an ideological division of a society within the minds of its members is a necessary precursor to genocide. At initialization, the model does not specify an out-group or an in-group, and if all other attributes and environmental conditions are held equal, an agent's Ideology setting simply simulates its perception of the other. No group starts with an advantage or a position of dominance.
3. *Influence* (*static, 0–1, floating-point, random uniform initialization*): In the model, the elite are simulated through the Influence attribute. According to the rule structure outlined below, an influential agent has a higher likelihood of affecting an increase in its neighboring agents Ideology levels. This is a highly simplified representation of the elite in a society, and the initial model gives all agents the same random chance of being an influential, and therefore elite, actor.
4. *Susceptibility* (*static, 0–1, floating-point, random uniform initialization*): This attribute determines to what degree an agent updates its Ideology in the presence of a more influential agent with the same Identity. Its random initialization allows a uniform distribution throughout the agent population, simulating the wide variety of personality differences possible in human societies.
5. *Threshold to Act* (*static, 0–1, floating-point, random uniform initialization*): The threshold attribute sets the level beyond which an agent acts against other agents in its vicinity with a different Identity. In the historical record, not all genocides have involved a high level of civilian participation. However, as discussed above, their compliance is necessary for the elite to further a genocidal agenda. In the model, acting against another agent can simulate active participation or mere compliance in the face of genocidal acts of others from the same identity group.
6. *Radius of Sight* (*static, 1–50 patches, global setting*): In the environment, agents are located on patches, and agents of both groups have the same level of

local visibility from their position as determined by this attribute setting.

To avoid bias, the model has the same rule structure for all agents. While it is possible to add a wide variety of interaction rules, the model includes only the simplest possible set of rules, leaving layers of complexity to future implementations.

Interaction Rules

These rules form the foundation for agent adaptation. Their simplified nature makes it possible to identify micro-level sources of emergent properties. In every simulation cycle, agents individually follow the procedure below:

Agent *n* identifies the most Influential agent with its same Identity in its radius of sight.

If there exists such an agent, and its Influence is higher than that of Agent *n*, it is identified as Agent *x*. Agent n updates its Ideology according to the following formula:

$$\textit{New Ideology}_{\textit{Agent n}} = \textit{Old Ideology}_{\textit{Agent n}} + (\textit{Ideology}_{\textit{Agentx}} - \textit{Old Ideology}_{\textit{Agent n}}) * \textit{Susceptibility}_{\textit{Agent n}}$$

This means an agent's change in Ideology in the presence of a more influential agent from its own group, is directly proportional to its Susceptibility.

1. Next, Agent *n* checks to see if ***New Ideology***$_{\textit{Agent n}}$ **> *Threshold to Act***$_{\textit{Agent n}}$, and if this is true, Agent *n* kills all agents from the other identity group in its radius of sight.
2. Finally, every agent has an equal probability of reproducing in each model cycle. A new agent attributes are initialized according to the same structure specified above except for the Identity attribute, which is inherited from its parent agent.

These attributes and rules are the minimum set required to simulate genocide in an ABM in the context of the system view outlined in Figure 2. While other starting points were possible, the works of Waller (2001) and Straus (2012a, & 2012b) greatly informed the selected initial model components. Straus helped define the system system-level understanding, and Waller (2001) provided support for key elements of Straus' model, such as the presence of an extreme ideology. His focus on the psychological dynamics in the environmental presence of an extreme ideology with respect to the other guided the selection of an ABM with the above parameters for this initial implementation.

Early Results—ABM

Experiment 1—Baseline Model Behavior

Given the above initialization, the first experiment determines the baseline behavior of the model. The analysis below is for the Base Model as described above, along with the following initial settings:

Populations: 50 agents each from Group A and Group B.

Radius of Sight: Five patches for all agents.

Reproduction: An agent has a 1/10 probability of reproduction if it has less than 10 agents of its own identity within its Radius of Sight.

In all runs using the above conditions, one agent group eventually dominates the other such that only its agents survive. There is a great deal of volatility in agent group populations, both numerically and spatially. Generally, there is a loose clustering of the minority group as members that are more isolated die before they can reproduce. Because this is not a result of localized agent decision-making, it is not an emergent property, but is rather a result of the globally implemented rule structure. A typical environment is shown below:

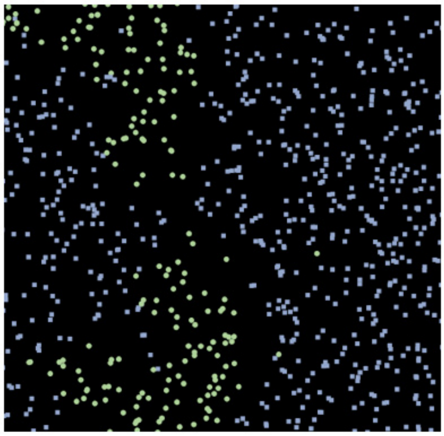


Figure 3. Base model environment.

Early fluctuations in each group's population as a percentage of the total system population have the general pattern shown below in Figure 4, with volatility and oscillations developing for simulations that run for extended periods:

500 runs of this model provide the following statistics:

1. Genocidal outcome: 100% of runs.
2. Average run length: 1,117 ticks.
3. Standard deviation: 748 ticks.

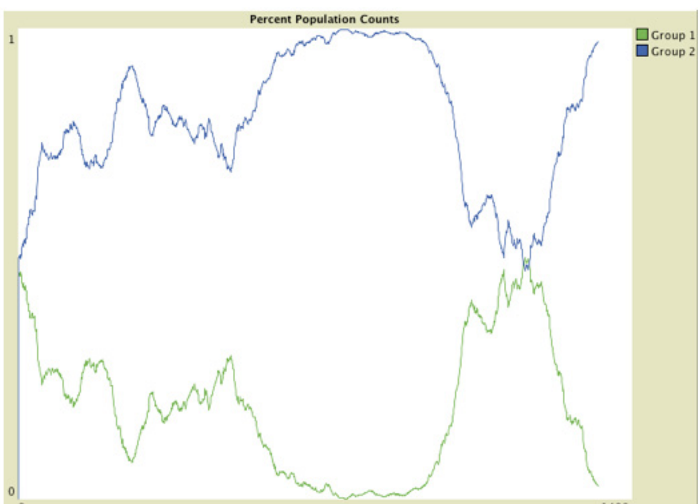


Figure 4. Base model population percentages in time.

The large standard deviation indicates that although the model invariably results in genocide, the length of time required varies considerably. This is likely due to the stochastic nature of the model, with its random uniform distributions of agent attribute variables and random movement.

This model effectively simulates genocide because of adaptation of agent Ideology in the presence of a local and more influential agent. While volatile, it provides a base from which one can build in order to study mechanisms that restrain or escalate genocide.

Experiment 2—Validation

The 1994 Rwandan genocide provides a reasonably straightforward example from which to attempt validation of this model. Prior to the genocide, Rwanda possessed a population greater than 7 million people, with ethnicities represented as follows: 85% Hutu, 14% Tutsi, and 1% Twa (United Nations, 2016). The genocide lasted for 100 days, and approximately 800,000–1,000,000 Tutsis and moderate Hutus were killed (Survivors Fund, n.d.). Allowing for a small number of the initial population of Tutsis to survive, the model calibrated for the Rwandan genocide has the following settings:

Populations: 850 Hutu/In-Group agents, 150 Tutsi/Out-Group agents.

Radius of Sight: Five patches for all agents.

Reproduction: No reproduction due to the limited timeframe.

Hutu Ideology: Random uniform distribution from 0 to .24.

Tutsi Ideology: Random uniform distribution from 0 to .04

Influence: Random uniform distribution from 0 to 1 (floating-point number)

Susceptibility: Random uniform distribution from 0 to 1 (floating-point number)

Threshold-to-Act: Random uniform distribution from 0 to 1 (floating-point number)

Population Cutoff: Model stops running when 45 Tutsi agents remain

Of 500 runs, 79% lasted less than 200 ticks/days. Of that 79%, the average run length was 67 ticks with a standard deviation of 42. As such, the model is not suitable to simulate a scenario such as that found in the Rwandan case, but it can provide a very rough validation. Beyond simple Hutu population dominance, the agent attribute that allowed this loose approximation of the actual event was Ideology. Hutu Ideology against Tutsis has a wider distribution range, which simulates their dominant and aggressive position. However, Tutsis do have a very small Ideology against Hutus, which can simulate factors such as retaliation or assistance from other parties.

This model may be more appropriate for simulating longer-term scenarios such as the Bosnian genocide or the Holocaust. In these cases, reproduction could justifiably be introduced in order to refine the model's calibration, and a longer timeframe would allow a wider and more realistic range of agent Ideology for both groups. However, these cases introduce a great deal of complexity with respect to third parties and exogenous factors, which would then require a more advanced model in order to produce a meaningful validation.

Experiment 3—Uniform Global Threshold to Act

In this experiment, all settings of the Base Model are identical to those listed in Experiment 1, except for Threshold to Act. Rather than every agent having this attribute randomly set to a floating-point number between 0 and 1 according to a random uniform distribution, it is given a fixed level that is the same for all agents from both groups. The Threshold to Act remains static throughout the simulation.

This experiment steps through Threshold to Act levels from .10 to .75 in increments of .05, running five replicates for each setting. To avoid simulations running indefinitely, the code is adjusted to stop the simulation after 7,000 ticks, as this is generally the timeframe beyond which the model never produces genocide. Between Threshold to Act levels .45 and .50, genocide does not occur beyond a cutoff.

Figure 5 below shows the general pattern of simulation run length as a function of Threshold to Act.

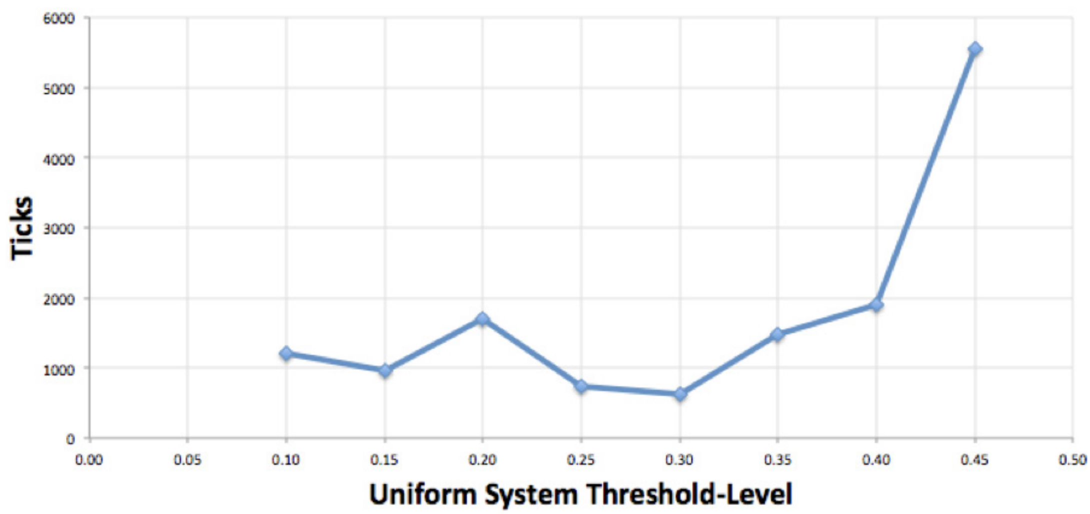


Figure 5. Simulation length versus threshold-level.

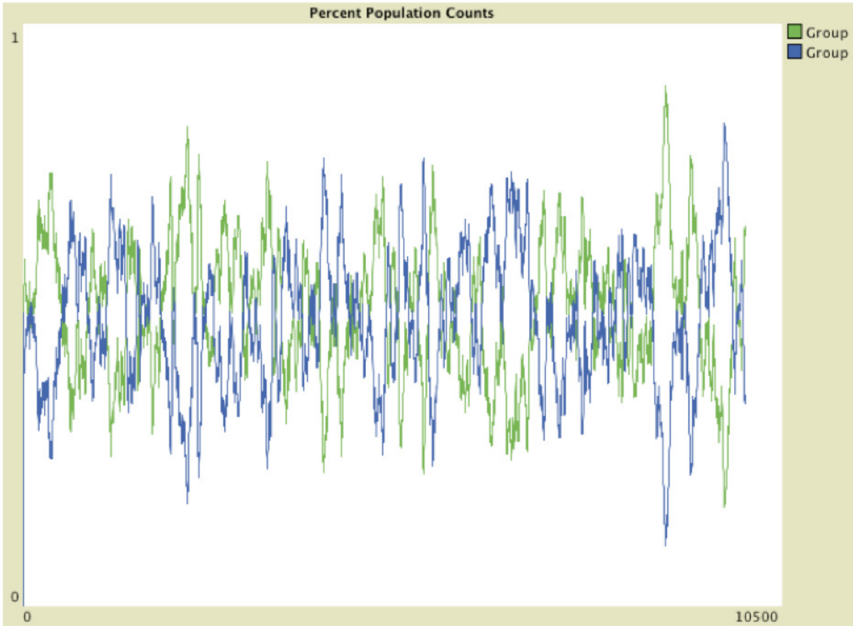


Figure 6. % population oscillation for threshold-level of .6.

These early results show a clear exponential increase in simulation run length as the global Threshold to Act level approaches .5. The first indication these results provide is that while the model is generally biased toward genocide as seen in Experiments 1 and 2, it is flexible enough to be able to avoid genocide through making the Threshold to Act a globally uniform and static value in the middle of its range. The implication of this result is that even in a divided society with identity-based conflict, a genocidal outcome is far less likely if people are unwilling to act against the opposing identity group regardless of their level of animosity or grievance against the other. Thus, policies could have greater impact if geared towards providing incentives not to act, even if they do not remove 100% of that willingness. Additionally, as noted above, endogenous factors such as a strong middle class may be a source of restraint that should not be underestimated if it indeed causes people to avoid taking action against opposing identity groups (Straus, 2012b).

This experiment yielded a second notable property. As the Threshold to Act level increased, the spread between average population counts of the two groups decreased exponentially. Running the model for 10,000 ticks with Threshold to Act levels ranging from .5 to .9 in steps of .1, and at .95 produced the following results given below. Figure 6 shows group percentage population oscillations for a Threshold to Act level of .6. Here, it is clear that while neither population can attain dominance such that it completely exterminates the other, there is a great deal of volatility in population counts, indicating high levels of violence.

Figure 7 shows the trend for the average spread between population percentages at each Threshold to Act level:

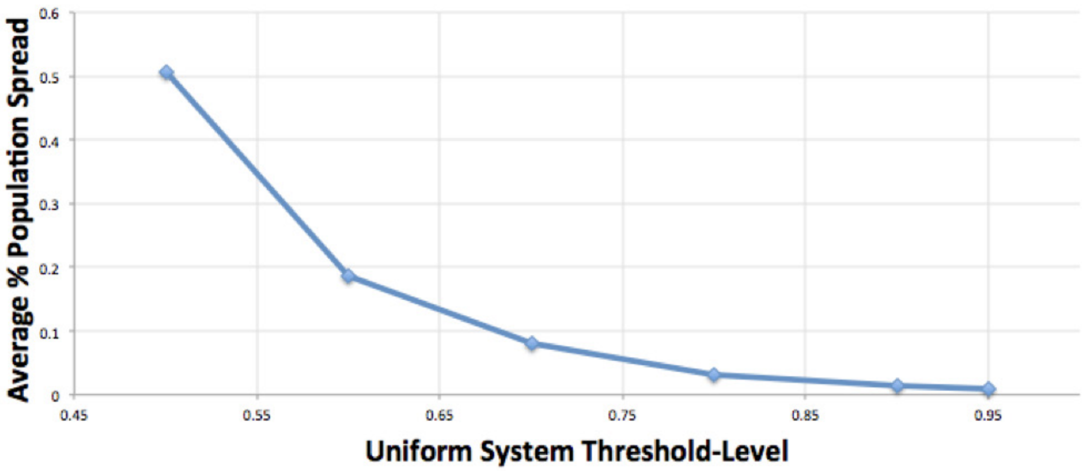


Figure 7. Average population spread versus threshold-level.

These results derive from the fact that the two groups are initialized with equal numbers of agents, and as their Threshold to Act level increases, less killing occurs, and the total populations move exponentially toward the initial condition of equal population counts. While this is in part due to the rule structure of the model, the exponential decline in the spread was unexpected. The implication of this result is that an increasing societal resistance to acting against those with a different identity has an exponentially beneficial effect as that resistance approaches a level at which no violent action is committed against other identity groups. This reinforces the point taken from this experiment's first set of results; policies geared toward incentives for people *not* to act against the other have a positive impact potential that increases exponentially as more people choose peace over violence.

SD Model Implementation

Next Modeling Steps

This section provides a brief overview of an implementation framework for a SD stock and flow model that simulates genocide. In brief, stocks are elements in the system that accumulates in time. Flows are the rates of change into and out of a stock. Different stocks may be connected to each other only through their flows. Stocks can be physical items, money, or even emotional states. To implement a running SD simulation, it is necessary to deviate in part from the causal loop diagram shown in Figure 1. This allows derivation of stocks that are less challenging to implement and validate. Figure 8 shows the simplified causal loop diagram for this model developed using Vensim PLE software (Version 6.3G; Ventana Systems, Inc., 2006). Having this diagram is the next step in the progression toward a running SD model. Here, Dominance of the In-Group increases with its percentage of the total population. As the In-Group's

Dominance increases, the rate of killing of the Out-Group (as measured by its Death Rate) increases, decreasing the total Out-Group population, increasing the In-Group Population Percentage, and thus further increasing the Dominance of the In-Group. To study the dynamics more thoroughly, In-Group Compliance and In-Group Control of Government and/or Military are included as factors of escalation that increase In-Group Dominance.

The next step in this process is to implement a working model in Vensim PLE that includes the appropriate stocks (In-Group and Out-Group Populations, and In-Group Dominance), and flows into and out of those stocks. These flows include normal birth and death rates, increasing Out-Group Death Rate due to In-Group Dominance, and increase or decrease of In-Group Dominance due to the effects of In-Group Population Percentage, Compliance, and Control of the Government and/or Military. Each effect on Dominance requires a function that determines how strongly it influences the Out-Group Death Rate. Scenario analysis is necessary in order to derive realistic functions, baseline conditions, and other relevant rates of change.

Additionally, it is possible to model both In-Group and Out-Group Dominance as separate stocks with their own unique sets of exogenous effects. This would lead to a model that simulates a more complex system and might display oscillations in genocide over time as one group loses its dominance to the other. A validated model of this nature would allow exploration into the macro-level dynamics that lead to future retaliation of former Out-Groups and potential sources of restraint.

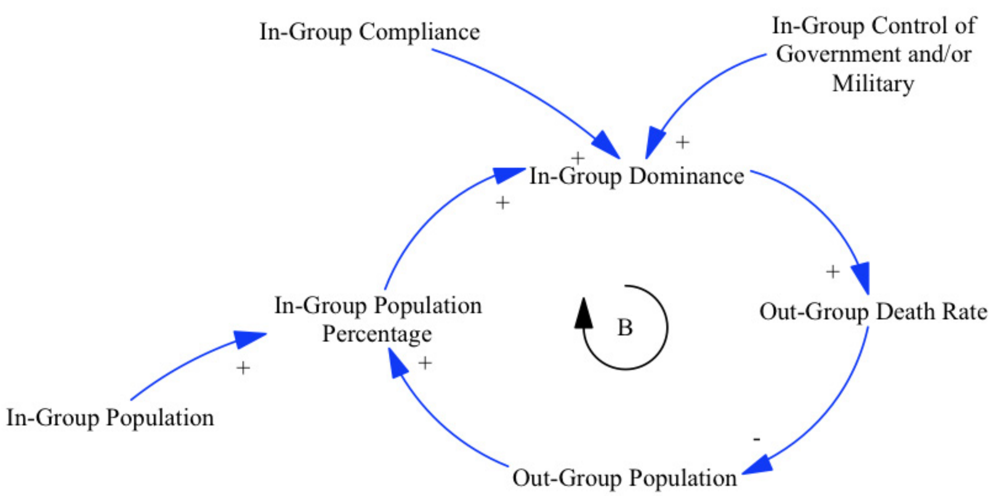


Figure 8. Diagram of population and dominance dynamics of genocide.

Connecting SD and Agent-Based Modeling

An ABM provides a window into the micro-level dynamics of agent interaction. For instance, Experiment 1 shows that completely random variation in an agent's willingness to act against others of a different identity consistently results in genocide in the presence of an identity-based conflict. However, the results of Experiment 3 show that uniformity in that willingness to act across a population, implemented as the Threshold to Act, can restrain genocide as that willingness decreases. The above result emerges despite the presence of individually varying levels of Ideology, Influence, and Susceptibility. This shows that people can be unique and varied with respect to some of their opinions and personality traits, but if they unite, even in part, in their opposition to violence against the other, it becomes possible to avoid a genocidal outcome.

A SD model provides a different perspective. Here, a model like the one outlined in the previous section can allow analysis of macro-level conditions and their impact on the likelihood of genocide. For example, a model validated on a specific historic incident can then be used to study how different variations in In-Group Control of Government and/or Military can restrain or escalate genocide. While this will not provide information about individual personalities and interactions, it does speak to the system-level dynamics. As such, an SD model can provide crucial information about macro-level mechanisms, the results of their feedback structures, and the impact on the incidence of genocide in a society.

In terms of policy recommendations and implementations, each modeling approach provides critical information that, when taken together, can result in a more comprehensive understanding of the problem, leading to policies that are more effective. For instance, an ABM similar to the one implemented in this research can be further improved in order to more accurately model and study the effects of variation of ideology with respect to different identity groups, revealing areas of sensitivity and thresholds beyond which genocide becomes unlikely. However, policy is often implemented at the macro-level, and an advanced and validated SD model can allow analysis of macro-level effects of those policies indicating the probability of their usefulness in terms of preventing genocide. In summary, an ABM can point policymakers in the right direction from the perspective of the individual citizen, and a complementary SD model can explore the potential macro-level effects of those same policies before they are implemented and filter down to the individual.

Conclusions and Future Work

In this research, the preliminary results from the ABM are encouraging. A full exploration of the model's parameters will likely reveal critical information about local-level dynamics that lead to genocide. While the model is simple, it is powerful enough to demonstrate possibilities for policy recommendations. Furthermore, the model shows that a restraining factor can have an exponentially beneficial effect as it brings forth resistance to violence in the general population.

Given the Base ABM Model's consistent genocidal outcome, future work should explore ways to mitigate this issue of bias. Questions to ask in this respect are:

- Does the addition of a third identity bring forth stability in the environment?
- How does the initial size of that third population affect its impact on the environment as a whole?
- How do the dynamics change when adding additional groups of varying sizes?
- Are there other incentives, mechanisms, or interaction rules that can be introduced to reduce bias by allowing the model to move toward or away from genocide?

In terms of policy recommendations, possible questions are:

- What incentive could agents have not to act? For example, would policies leading to a strong economy with a strong middle class provide a source of restraint yielding an effect similar to that seen in Experiment 3 with uniform Threshold to Act?
- Given an incentive for peace, such as a strong economy, what policies could be implemented in societies with identity-based divisions that improve economic conditions, and how can they be tailored such that all groups benefit?

Finally, this implementation of an ABM uses random uniform distributions for all parameters. While this has been informative, there may be other distributions, such as normal or power-law, which could produce insightful results. A full analysis of the model requires testing different distributions on every parameter and combination of parameters in order to determine differences in outcomes.

As stated above, while the broad causal loop diagram in Figure 1 and its explanation provides the framework for understanding the scope and complexity of the problem of genocide, a formal SD model implementation requires a stock and flow model in order to analyze the problem from this perspective. The causal loop diagram shown in Figure 8 provides a starting point for this model, but could be expanded to allow for greater complexity of feedback mechanisms, or exploration of areas of policy interest. Beyond the impact of the economy, the model could explore the effect of sanctions, political resistance, or armed resistance from internal or external actors. The above model can then be taken in combination with the results of ABM analysis to derive comprehensive policy decisions that account for macro- and micro-level effects.

Finally, social science researchers in genocide studies have a significant restraint on their ability to analyze, understand, and predict the problem because they must wait for an emergence, one that brings with it a great deal of human suffering, in order to validate their models. This research attempts to synthesize their knowledge, adapting it to a computational approach in order to develop models that are useful and informative. While the research and models are not fully developed, they have already produced intriguing results. As this work matures, it has a strong potential to put the power of computational modeling and simulation into the hands of social scientists, such that they can benefit from its analytic and predictive capabilities. This, in turn, increases the ability of society as a whole to more effectively address and mitigate the problems associated with identity-based conflict.

References

Cox, J. M. (2017). *To kill a people: Genocide in the twentieth century*. New York, NY: Oxford University Press.

Straus, S. (2012a). Destroy them to save us: Theories of genocide and the logics of political violence. *Terrorism and Political Violence*, 24(4), 544–560. Retrieved from http://doi.org/10.1080/09546553.2012.700611

Straus, S. (2012b). Retreating from the brink: Theorizing mass violence and the dynamics of restraint. Perspectives on Politics, *10*(02), 343–362. Retrieved from http://doi.org/10.1017/S1537592712000709

Survivors Fund. (n.d.). *Statistics*. Retrieved December 2016, from Survivors Fund. http://survivors-fund.org.uk/resources/rwandan-history/statistics/

United Nations (2016). *Outreach Programme on the Rwandan Genocide and the United Nations*. Retrieved from http://www.un.org/en/preventgenocide/rwanda/education/rwandagenocide.shtml

Ventana Systems, Inc. (2006). Vensim PLE Software (Version 6.3G) [Software]. Retrieved from http://www.temoa.info/node/3682

Verdeja, E. (2002). On genocide: Five contributing factors. *Contemporary Politics*, *8*(1), 37–54. Retrieved from http://doi.org/10.1080/13569770220130112

Verdeja, E. (2010). Genocide: Clarifying concepts and causes of cruelty. *The Review of Politics*, *72*(03), 513–526. Retrieved from http://doi.org/10.1017/S0034670510000343

Waller, J. (2001). Perpetrators of genocide: An explanatory model of extraordinary human evil. *Journal of Hate Studies*, 52, 5–22.

Wilensky, U. (1999). NetLogo (Version 5.3.1) [Software]. Retrieved from http://ccl.northwestern.edu/netlogo/

Dynamics of Intergovernmental Networks: Harnessing Agent-Based Modeling Simulations for Resilient Infrastructures

Asim Zia, Christopher Koliba[A]

Abstract

Agent-based models (ABMs) can be deployed as policy informatics platforms to track resource flows and distributions under differential configurations of intergovernmental networks and shocks to the policy systems. This study provides a detailed application of a policy informatics platform in the contested arena of transportation policy implementation networks across federal, the state of Vermont, and its regional and local governments. Through this policy informatics platform, two specific questions are addressed: (1) How the allocation and distribution of federal and state funding resources for transportation infrastructure development projects changes under different configurations of weighting state versus regional versus local government priorities? (2) How do different shocks to the intergovernmental policy system, such as increased frequency and intensity of extreme weather events and/or federally mandated funding sequestrations, influence the distribution of financial resources across regional and local governments? A pattern-oriented, ABM of a transportation governance network, calibrated for the state of Vermont including its regional and local town governments, is presented. This ABM simulates the dynamics of transportation project prioritization processes under alternate intergovernmental institutional rule structures and variable frequencies of shocks to assess their impacts on financial investment flows from federal to state, regional, and local scale governments. Multiple focus groups, individual interviews, and analysis of federal, state, and regional scale transportation project and program data informed the development of this ABM. This study demonstrates a practical and detailed application of a policy informatics platform by showing how ABM simulations may be used to evaluate the design of intergovernmental policy implementation networks under differential frequencies and intensities of shocks to policy systems and their impacts on policy outcomes.

"Implementation and evaluation are the opposite sides of the same coin, implementation providing the experience that evaluation interrogates and evaluation providing the intelligence to make sense out of what is happening" (Pressman & Wildavsky 1984, p. xv).

Keywords: institutional designs, policy computing, governance networks, intergovernmental relations, resilience, infrastructure, transport policy, agent-based models

[A] University of Vermont

doi: 10.18278/jpcs.3.2.4

摘要

基于主体的建模（ Agent Based Model ，简称ABM）可用作政策信息学平台，在政府间网络的不同配置和政策系统遭受的冲击下追踪资源流动和分配。本文提供了政策信息学平台的详细应用方式，用于联邦、佛蒙特州以及当地政府间的交通政策实施网络。 通过该平台，本文处理了两个具体问题：（1）在权衡州政府、地区政府和当地政府各自的工作重点后，联邦和州资金渠道在分配给交通基建设施发展项目时将如何变化。（2）政府间政策系统遭受不同冲击时，财政资源在地区和当地政府的分配方式将受到怎样的影响？这些冲击包括：极端天气事件发生的频率和强度增加，以及查封联邦授权的资金。 本文针对佛蒙特州及其地区政府和当地镇政府，提出了一项用于交通治理网络的，以模式为导向（Pattern-Oriented）的ABM。ABM在政府间制度规则结构（intergovernmental institutional rule structures）和冲击变量频率的情况下模拟了交通项目优先化过程的动态，以评估从联邦到州、地区和当地政府的财政资金流动所受到的影响。该ABM得以发展的原因在于多焦点小组（focus groups）、个人访谈以及关于交通项目和计划数据的分析。通过展示ABM模拟物如何在政策系统遭受不同频率和强度的冲击下，评估政府间政策实施网络的设计，本文证明了政策信息学平台的实际应用。

关键词：制度设计，政策信息学，治理网络，政府间关系，韧性，基建设施，交通政策，基于主体建模

Resumen

Los modelos basados en agentes pueden ser utilizados como plataformas de informática política para rastrear los flujos de recursos y las distribuciones bajo configuraciones diferenciales de redes intergubernamentales y los impactos a los sistemas políticos. Este estudio proporciona una aplicación detallada de la plataforma informática política en el área impugnada de las redes de implementación de política de transporte en el gobierno federal, en el del estado de Vermont y en sus gobiernos regionales y locales. A través de esta plataforma de informática política dos preguntas específicas se abordan: (1) Cómo la asignación y distribución de recursos federales y estatales para los proyectos de desarrollo de infraestructura de transporte cambia bajo diferentes configuraciones de pesaje estatal contra prioridades gubernamentales locales. (2) Cómo diferentes impactos al sistema político intergubernamental, tales como la frecuencia e intensidad de eventos de fenómenos meteorológicos extremos y/o retenciones presupuestales requeridas por el gobierno federal ejercen influencia sobre la distribución de recursos financieros en gobiernos locales y regionales? Un modelo basado en agentes (ABM) y orientado hacia los patrones de una red de gobernanza del transporte, calibrado para el estado de Vermont, que incluye su gobierno local y regional, está presentado. Este ABM simula la dinámica de los procesos de priorización de proyectos bajo estructuras reglamentarias institucionales intergubernamentales alternativas y frecuencias variables de impactos para evaluar sus impactos en flujos de inversión financiera del gobierno federal a gobiernos estatales, regionales y locales. Múltiples grupos focales, entrevistas individuales y análisis de datos de proyectos y programas de transporte federales, estatales y regionales informaron el desarrollo de este ABM. Este estudio demuestra una aplicación práctica y detallada de una plataforma informática política al mostrar cómo los simulacros ABM pueden ser utilizados para evaluar el diseño de redes de implementación de política intergubernamental bajo frecuencias diferenciales e intensidades de impactos a los sistemas políticos y sus impactos en los resultados de las políticas.

Palabras clave: Diseños institucionales, Informática política, Redes de gobernanza, Relaciones intergubernamentales, Resiliencia, Infraestructura, Política de transporte, Modelos basados en agentes

1. Introduction

This article develops a pattern-oriented, agent-based model (ABM) to simulate alternative institutional rule scenarios in order to assess the alternate outcomes of transportation funding. The ABM simulates an intergovernmental network consisting of federal resources, state decision makers, and regional and local representatives. The ABM is designed to generate experimental simulations for addressing the following two specific questions, which broadly pertain to the design and governance of intergovernmental transportation policy implementation networks and their predicted impacts on policy outcomes:

1. **Federalism versus Regionalism versus Localism Debate:** How the allocation and distribution of federal and state funding resources for transportation infrastructure development projects changes under different configurations of weighting state versus regional versus local government priorities?
2. **Resilience Assessment:** How do different shocks to the intergovernmental policy system, such as increased frequency and intensity of extreme weather events and/or federally mandated funding sequestrations, influence the distribution of financial resources across regional and local governments?

This article demonstrates that a "pattern-oriented," ABM approach can provide a viable empirical and computational methodology for effectively designing the functions and resource allocation decision-making processes of intergovernmental policy implementation networks. Through such policy informatics platforms, the emergent and self-organizing phenomena observed in policy and governance systems, as well as lags and inertia in designing and implementing programs and projects, can be more accurately modeled and simulated. The development of such policy informatics platforms requires "pattern-oriented" computational modeling of the system that "generates" observed patterns of resource flows, resource allocations, and resource distributions under an observed configuration of intergovernmental networks authorized to implement specific public policies (see Epstein (2006) for generative social science applications).

The capacity of computer models of complex governance networks to lead to accurate forecasting and prediction of particular policy outcomes is predicated on a "deep uncertainty" that characterizes our current state of understanding of complex social systems and associated wicked planning problems (Rittel & Webber, 1973, 1984). Bankes (2002, p. 7263) characterizes this deep uncertainty arising as, "the result of pragmatic limitations in our ability to use the presentational formalisms of statistical decision theory to express all that we know about complex adaptive systems and their associated policy problems."

To cope with the inherent complexity and uncertainty in the social complexity of governance networks, we undertake a variation of "pattern-oriented" modeling. Pattern-oriented models are described by Grimm et al. (2005) as "bottom-up" models that emphasize the applicability of models to real problem solving.

Grimm et al. (2005) describe pattern-oriented models this way:

> "In [this approach to modeling], we explicitly follow the basic research program of science: the explanation of observed patterns. Patterns are defining characteristics of a system and often, therefore, indicators of essential underlying processes and structures. Patterns contain information on the internal organization of a system, but in a "coded" form. The purpose of [pattern-oriented modeling] is to "decode" this information… A key idea [in these models] is to use multiple patterns observed in real systems to guide design of model structure. Using observed patterns for model design directly ties the model's structure to the internal organization of the real system. We do so by asking: What observed patterns seem to characterize the system and its dynamics, and what variables and processes must be in the model so that these patterns could, in principle, emerge?" (p. 987)

Pattern-oriented approaches are pursued because they help to focus and reduce the uncertainty found in any model of a complex adaptive system. Grimm et al. (2005, p. 990) add that,

> "[Pursuing a pattern-oriented] strategy is a way to focus on the most essential information about a complex system's internal organization. Multiple patterns keep us from building models that are too simple in structure and mechanism, or too complex and uncertain. Using patterns to test and contrast alternative theories for agent behavior or other low-level processes is a way for [modelers] to get beyond clever demonstration models and on to rigorous explanations of how real systems are organized and how they respond to internal and external forces."

In the literature, classical research methods, such as case studies, interviews, and surveys, have been used for analyzing the functions, capacities, and dynamics of intergovernmental policy implementation networks (Agranoff, 2007; Agranoff & McGuire, 2003; Rhodes, 1997). While these classical methods are useful in providing insights about analyzing governance networks, we argue in this study that computational approaches, especially "pattern-oriented" ABMs, provide a complementary powerful and evidence-based methodology to analyze the allocation of resource flows across intergovernmental policy implementation networks while accommodating for complex interactions of institutional rules, both formal and informal, at multiple levels of government. In particular, ABMs enable the modeling of emergent and self-organization phenomena, as well as lags and inertias that are typically observed in resource allocation decisions across intergovernmental policy implementation networks (Koliba, Meek, & Zia, 2010; Zia & Koliba, 2015).

Next, in Section 2, we briefly describe the federal and state transport policy context of this intergovernmental transportation policy implementation network that governs its dynamic operations in the specific context of intergovernmental transportation project prioritization. In Section 3, we present research methods that were used to model roadway project prioritization processes and to elicit decision heuristics of multilevel agents in the simulation model. In Section 4, we present the fundamental structure of the stochastic, multilevel ABM. Section 5 presents findings from experimental simulations to address the two research questions posed above. Section 6 discusses the limitations of the current simulation

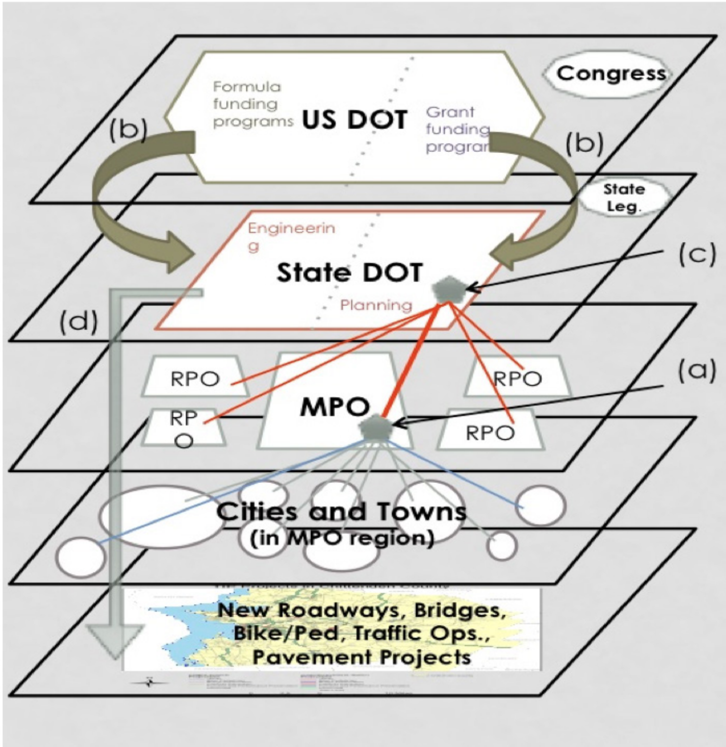


Figure 1. Intergovernmental transportation policy implementation network in Vermont.

model and possible next steps to further improve and generalize it. In Section 7, we present our conclusions and discuss the implications of our study findings for transportation policy implementation, in particular, and public policy analysis and management, in general.

2. Intergovernmental Transportation Planning and Policy Implementation Network: Legal and Policy Context for Resilient Infrastructures

Since 1991, federal transportation laws have emphasized two policy goals. First, Congress is interested in developing an intermodal transportation system where citizens can safely use multiple forms of private and public transit. Second, to qualify for federal transportation funds, projects need to be carried out with the cooperation of state and local governments: the planning process needs to be ***continuing, comprehensive, and cooperative***. This principle of ***"3C's"*** has been in place since the Federal Aid Highway Act of 1962 [Pub. L. No. 87-866, § 1, 76 Stat. 1145, 1148 (1962)]. The detailed planning system intertwines municipalities, state transportation agencies, and the US DOT. The circular framework requires a hierarchal planning system supplemented by federal funding for infrastructure development, as shown in Figure 1.

In Vermont, as shown in Figure 1, the state agency of transportation, (VTrans) is delegated the tasks assigned to the State. In addition, the Chittenden County Metropolitan Planning Organization (CCMPO), the only Vermont MPO, and

10 other regional planning commissions (RPCs) collaborate with the State to incorporate the regional transportation planning needs. The planning process and financial structure developed in ISTEA, TEA-21, and SAFETEA-LU creates both a complementary and competitive partnership between the RPCs/MPO and VTrans. Federal legislation directs the state to act as an intermediary between the MPOs/RPCs and the US DOT. ISTEA required each state transportation agency to accumulate each transportation improvement plan (TIP) from MPOs and assemble a Statewide Transportation Improvement Program (STIP). Similar to the MPO, the state must demonstrate the need for a project, the financial requirements, and the proposed funding location. In addition, each state is required to create a statewide Long Range Plan (SLRP). Congress's choice to require both the state and individual MPOs to create an LRP and TIP demonstrates its intent to commence planning at the local level while requiring involvement with all agencies and individuals with expertise in the transportation sector.

Depending on state regulations, the state completes federally funded projects by either distributing money to a municipality or completing the project themselves. The process of prioritizing transportation projects is rather more complex and relatively poorly understood. In the state of Vermont, for example, we deployed focus groups and interviews with state, regional, and local government officials to elicit the current institutional rule structure, i.e., the "rules-in-use" for project prioritization that are followed by the state and MPOs/RPCs in prioritizing transportation projects (see Ostrom (2005) for "rules-in-use"). These rules are different for six transportation project classes, which are roadways, paving, bridges, bike/pedestrian, traffic operations, and park and ride. In this article, we focus on roadway projects. An asset management system was developed in VTrans in 2006 to assign expected value scores to roadway projects for two specific criteria. Prior to 2006, transportation projects were prioritized through a political process of negotiation among the regional and state agencies. The introduction of these rules in 2006 was part of a broader "de-politicization" process, as stated by many focus group participants. For the roadway projects, VTrans evaluates each project on four criteria: highway system is assigned 40% weight, and cost per vehicle mile is assigned 20% weight. Both highway system and cost per vehicle mile for a given project are "objectively" estimated from the VTrans asset management system. Furthermore, regional priority is assigned 20% weight and "project momentum" is assigned 20% weight. Noticeably, while asset management system-based criteria (e.g., highway system and cost per vehicle mile) might protrude an air of objectivity, the criterion of "project momentum" is subjective and, according to a focus group participant, could even be used for "gaming the project prioritization system". Later, in this article, we demonstrate how changing the weights, both formal/objective and informal/subjective, provided by state agencies and RPCs can have a drastic effect on funding patterns.

Within the legal and policy context described above, the challenges posed by climate change induced extreme events, such as increasing frequency of floods, storms, and hurricanes, as well as macro-economic shocks, e.g., funding sequestration driven by political cycles (Baumgartner & Jones, 1993), periodically test the resilience of transportation systems. It is increasingly recognized that extreme events induced by global climate change pose daunting risks to transportation infrastructure across the globe (Jaroszweski, Chapman, & Petts, 2010). While a large amount of the literature on transportation and climate change focuses on mitigation (e.g., the role of transportation in carbon reduction) (Schmidt & Meyer, 2009), more research is needed on understanding the impacts of global climate change on transportation infrastructure and designing adaptation strategies to build resilient infrastructures (Hunt & Watkiss, 2011). In the Intergovernmental Panel on Climate Change's (IPCC's) Fifth Assessment Report, Revi et al. (2014, p. 562) concluded that transportation systems are "a difficult sector to adapt" due to large existing stock, especially in cities in developed countries, leading to potentially large secondary economic impacts with regional and potentially global consequences for trade and business. Emergency response requires well-functioning transport infrastructure. The IPCC synthesis team concluded (Revi et al., 2014, p. 563) that there is relatively less literature available on understanding the role of "local government decisions to include adaptation in plans and investment programs." The local government decisions are, however, typically embedded in intergovernmental networks, some of which extend beyond public–public intergovernmental configuration, into governance networks comprised of public, private, and civil society actors (Koliba et al., 2010). Coordination among national, regional, and local level agencies in such governance networks is critical for implementing adaptation strategies in the transport sector, as climate change impacts are widespread and extend across scales (Regmi & Hanaoka, 2011).

The Northeastern US has suffered from recent storms that point toward a need for improved resilience. In August 2011, Vermont was hit especially hard by Tropical Storm Irene, which caused an estimated $250 million dollars of damage to transportation infrastructure alone. In addition to the ~200 bridges that were damaged on state roads, municipalities had to repair or replace another 280 bridges and 960 culverts (Johnson, 2012). The Vermont Agency of Transportation (known as VTrans) responded quickly to rebuild, but as importantly, the disaster spurred the Agency—as a state-level organization—to formalize its stance on climate change adaptation. In a white paper released a year after Tropical Storm Irene, VTrans outlines its goals and procedures for adaptation and notes roadblocks and potential improvement actions. Among the roadblocks are regulatory constraints, forecasting difficulties, political tensions, and, above all, budgetary restrictions. Some noted opportunities for improvement include standardization and digitization of records, infrastructure resilience monitoring, the expansion of asset inventories, and updating project prioritization guidelines (Johnson, 2012). On top of climate change induced extreme event impacts, budgetary restrictions and uncertainty surrounding the funding of long-term infrastructure projects cause another set of challenges to mainstream adaptation and resilience planning in intergovernmental project prioritization processes (Schulz, Zia, & Koliba, 2017).

3. Empirical Methodology for Eliciting Observed Patterns of Roadway Project Prioritization and Funding Allocations

In addition to the analysis of legal documents, a series of in-depth interviews and two focus groups with multiple stakeholders, including local government officials, metropolitan planning organization (MPO) staff and board members, staff of other RPCs in Vermont, state department of transportation (VTrans) officials, Federal Highway Association (FHWA) representatives, and US DOT officials, federal and state senator office representatives, and local NGOs were conducted in the fall of 2010 and 2012. The focus groups and interviews were recorded, transcribed, coded, and analyzed for major and minor themes. To understand how projects were prioritized, various source documents related to project funding were examined, including major pieces of federal legislation (ISTEA; SAFTE-LU; etc.), planning and policy documents developed by the SDOT and the MPO, meeting minutes, and project databases. The project prioritization data for all transportation projects in the state between 1998 and 2011 were obtained and analyzed to derive probability distribution functions for variables of interest, as explained below. The data included the classes of project, the scoring data from the SDOT and the MPO, the location of the project, and the amount of funds that went into the project.

Table 1 below shows descriptive statistics for roadway project prioritization scores from 2007 to 2010 that were derived from VTrans project prioritization data. Note that these scores were not available prior to 2006, however, as per focus groups; these scores reflect the implicit weightings on decision criteria practiced by the state agencies and RPCs in prioritizing transportation infrastructure projects. In general, about 61 (median value) roadway projects are funded in a given year. On a 100-point scale, the expected value of these projects averages ~50 +/-15%. Projects with a higher expected value are ranked higher and prioritized for funding in a given year subject to the availability of funding and approval by Vermont's legislature and federal agencies responsible for each federal program.

Table 1. V-Trans scores for roadway projects from 2007 to 2011: descriptive statistics

	2007 **N=73**	**2008** **N=61**	**2009** **N=60**	**2010** **N=61**	**2011** **N=57**
Expected value	48.15 (15.864)	50.07 (13.103)	49.82 (13.662)	50.36 (14.709)	49.33 (15.399)
Highways system	20.808 (9.984)	20.721 (9.510)	21.050 (9.826)	20.820 (10.044)	21.474 (9.838)
Cost/vehicle mile	11.041 (5.397)	11.508 (5.793)	10.900 (5.780)	11.016 (5.895)	10.316 (5.938)
Regional priority	9.643 (6.947)	10.721 (6.802)	9.633 (7.192)	10.557 (6.714)	9.263 (7.889)
Momentum	6.658 (4.808)	7.115 (5.410)	7.066 (5.590)	6.820 (6.412)	7.228 (6.598)

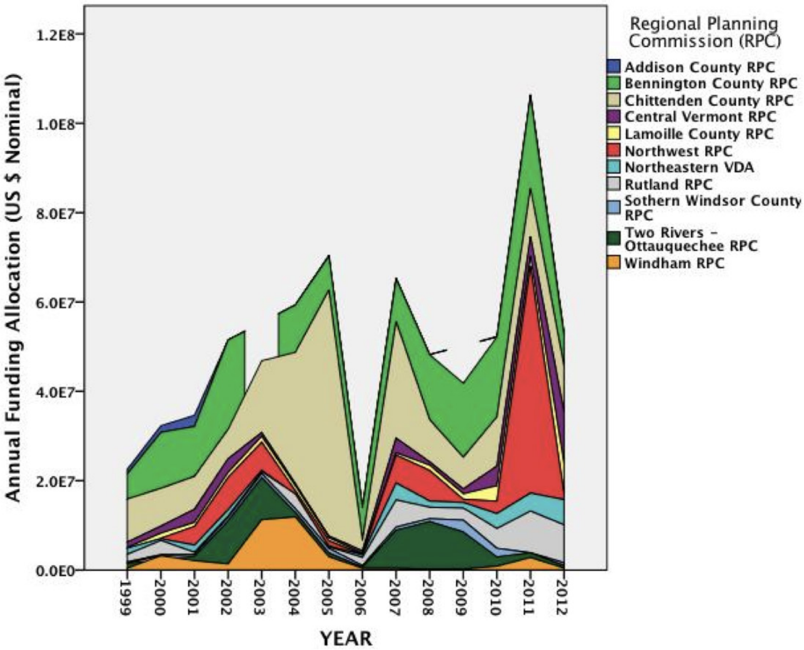


Figure 2. Observed patterns of roadway funding allocations among MPOs/RPCs in Vermont, 1999–2012 (US$, Nominal).

Remarkably, as summarized in Table 1, we found a consistent scoring pattern for the roadway projects. The statistical scoring pattern, however, belies random allocation of funding across regions and local towns, as shown in Figure 2. The observed funding patterns in Figure 2 are elicited from STIP 1999–2012 data and reveal rather a chaotic trend in terms of allocation of funding across the MPO/RPCs in the state of Vermont. The dip in total monetary allocation for roadway projects in 2006, as shown in Figure 2, is observed due to the change in the project scoring system. Another noticeable pattern in Figure 2 concerns the fact that some RPCs/MPO tend to receive more transportation funding for their projects than others over time. Furthermore, the relatively higher peak of funding allocation in 2010–2011 occurred due to the stimulus funding under the American Reinvestment and Recovery Act that is not a long-term transportation program.

4. The Structure of the Agent-Based Simulation Model

In this study, a pattern-oriented, ABM programmed in AnyLogic Professional Version is designed to simulate the intergovernmental decision-making process of the transportation policy implementation network shown in Figure 1, with an explicit focus on modeling the game-like interactions between the state, regional, and local government agencies. The ABM is designed as a stochastic multilevel model with the goal to generate observed funding allocation patterns shown in Figure 2 at the RPC level. We use Grimm et al.'s (2006) Overview, Design, and Details (ODD) protocol and present here brief structure of this intergovernmental ABM (see Figure 3). An earlier version of the ABM with a focus on ascertaining basins of attraction in resource allocations (Zia & Koliba, 2015) has been updated in this study with a calibration process implemented

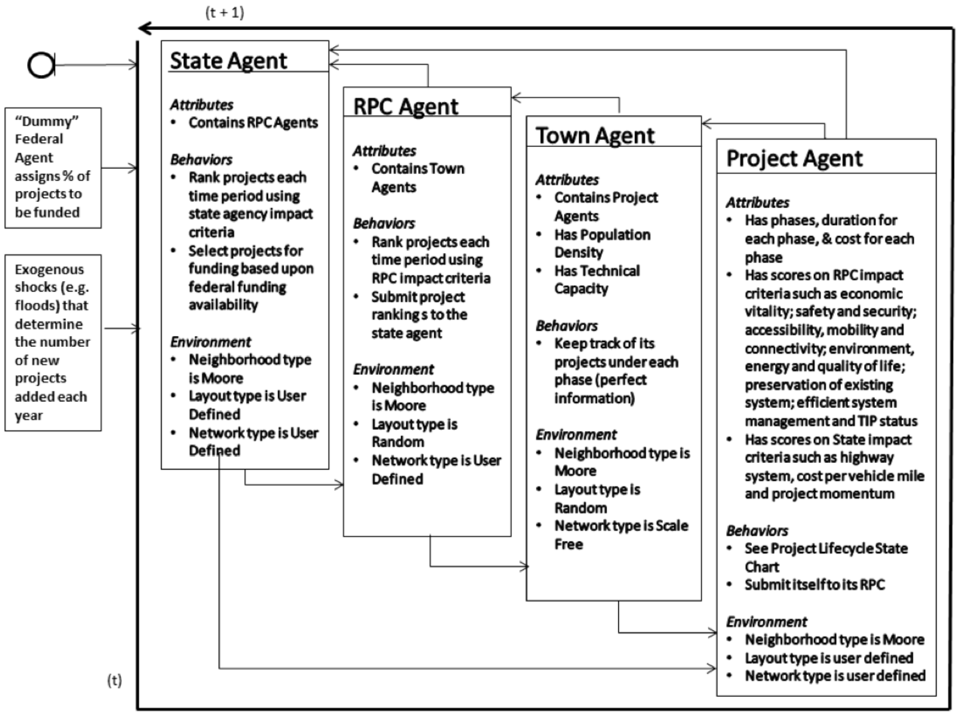


Figure 3. The internal structure of the stochastic ABM showing attributes, behaviors, and the environment of four agent classes—state agency, RPCs, local towns, and projects.

with a genetic algorithm to minimize the difference between observed and simulated funding allocations. The refined baseline scenario presented in this study presents different parametric settings compared with what was reported in Zia and Koliba (2015). Furthermore, alternate scenarios tested in this study and questions addressed are markedly different from earlier study. Finally, the simulation horizon is extended to 50 years, compared with 25 years reported in earlier work.

Structurally in the ABM, as shown in Figure 3, a state agent (i.e., VTrans) contains 11 nested RPCs and 600 local towns that are nested within RPCs. Furthermore, transportation projects are modeled at the inner most layer of the nested hierarchy. RPCs, local towns, and transportation projects are thus modeled as multilevel nested agents in this ABM, whereby project class agents are spatially situated within local town agents, local town agents are situated within RPC agents, and RPC agents exist within state agent. We have deliberately anonymized the identity of RPCs inside the ABM; whereas local towns inside the RPCs are initialized on random distribution basis (described below).

Figure 4 shows a state chart and transition functions among different states for the agent class of transportation roadway projects. During 2010 focus groups, one of the experienced participants described V-TRANS decision heuristic for funding transportation projects as a “funneling approach” that is captured for modeling the state chart of the project class in the ABM model shown in Figure 4. Every year, different agents in the policy implementation network identify a large number of transportation infrastructure problems. VTrans keeps an updated list of these problems and selects a sub-sample of these problems for undertaking feasibility

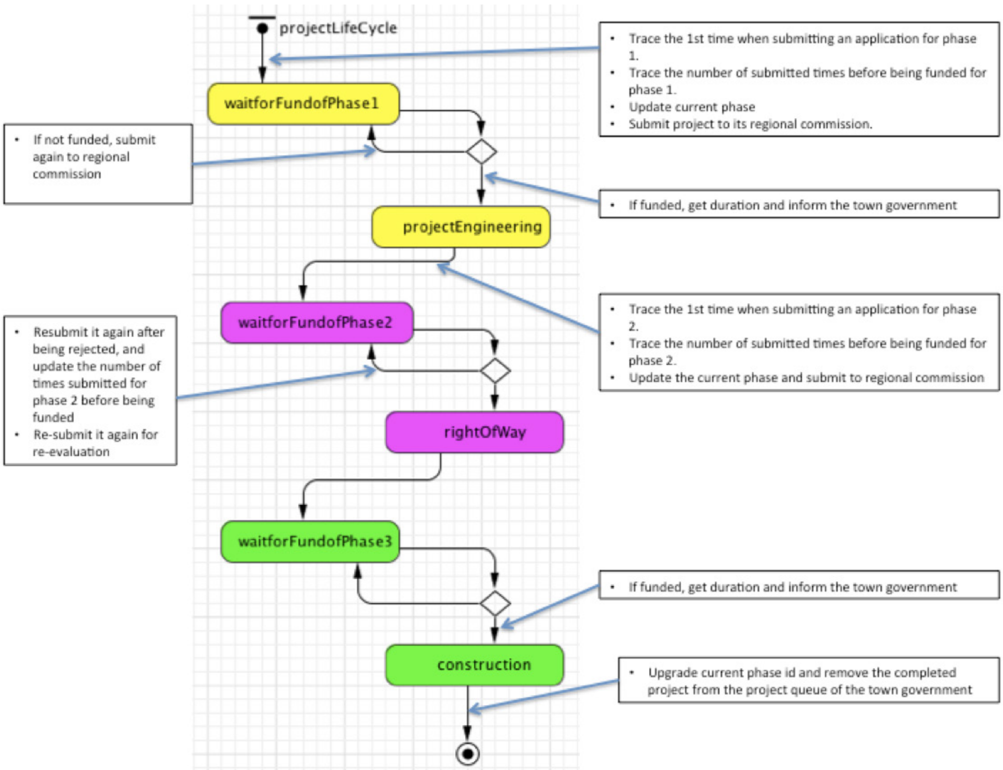


Figure 4. Transition state-chart for the agent class of "projects" (pseudo code is shown in the boxes for the key transitions between the six life cycle states).

studies (project engineering phase in Figure 4). The VTtans agent according to multiple criteria presented in Table 1 above ranks projects with completed feasibility studies. Figure 4 shows this only for roadway projects.

If a roadway project is prioritized at this stage, it goes through right-of-way (ROW) approval and then, if funding is available from relevant federal programs and if state and local government matching funds are available, the project is approved for construction phase. Overall, the ABM module of the computer simulation model for roadway (shown in Figure 4) and other project classes captures this "funneling" effect of a sub-sample of initially identified transportation projects that go through three project phases (feasibility, ROW, and construction). Duration and cost parameters for each of the three phases (feasibility, ROW, and construction), estimated from the STIP database, are shown in Figure 4. Furthermore, structurally, this project class loop in the ABM also captures the delays and the formation of old project queues that are empirically observed in the VTrans STIP database. The delays and queues across projects can be manipulated by varying the "duration" parameters in Table 3 under the project class.

The 600 local town agents in the ABM contain the environment of the projects and keep track of the transportation projects that arise over time on stochastic basis (modeled as a uniform probabilistic function to preserve the stochastic randomness). The 11 RPC agents keep track of their local town agents, as well as rank the transportation projects that are submitted to them every year based on their multiple criteria expected value functions (shown in Table 1). The parameters for the seven RPC criteria are calibrated to the observed probability distributions shown in Table 2. The RPC agents rank the roadway projects and send them to the state agent. Under the reference scenario, the state agent assigns 20% weight to the regional priority. A project that is ranked number 1 by an RPC is assigned 20 points by the state agent, a project ranked number 2 is assigned 18 points, and so on until a project that is assigned a rank of 10 or lower is given zero point by the state agent. Furthermore, the state agent calculates its expected value for each submitted project according to its default criteria (40% highway system, 20% cost per vehicle mile, and 20% to the project momentum) and ranks the projects by assigning rank no. 1 to the project with the highest expected value, 2 to the next highest, and so forth. The federal agent is not explicitly modeled as an agent in the current ABM. Rather, the federal agent is represented through a parameter that sets the available funding in a given year (shown as dummy agent in Figure 3). The ABM model user can define and vary this parameter. The default value is set to 15%, which was estimated during the calibration process of the ABM by minimizing the difference between observed and simulated funding allocations. The VTrans agent calculates the total cost of the ranked projects and selects top ranked projects for which funding is available in a given year. The state agent rejects the rest of the projects in a given annual cycle. The rejected projects are sent back to RPCs and local towns for evaluation in the next period. New projects are added (though a user-defined parameter) every year that are evaluated along with the projects that were rejected in the previous cycle. Model calibration process also showed that 60 new projects (consistent with observed data) are, on average, added every year.

Overall, there are 27 input parameters in the model as shown in Table 3. The majority of the input parameters in the model have either uniform or triangular probability distributions, which means that the ABM is stochastic at its base and each model run is a unique realization chosen from the random probability distributions. The probability distributions for the RPC parameters shown in Table 3 were estimated from the observed roadway project prioritization data (e.g., Table 2). Experimental simulations of the intergovernmental decision-making process could be run by varying the probability distributions of these input parameters, as explained in the findings section below.

Table 2. Comparison of convergent cross mapping coefficients

Outcome	Predictor	Regression beta	Correlation coefficient	CCM coefficient
CC Entries	CC Exits	0.55	0.63	0.80
CC Exits	CC Entries	0.80	0.69	0.77
FC Entries	FC Exits	0.00	0.01	0.38
FC Exits	FC Entries	0.44	0.33	0.49

Table 3. Stochastic parameters and their default probability distributions used for calibrating the ABM to the Vermont context

Agent level	Parameter	Default value
Project		
1.	Project ID	Unique Project Identifier
2.	Duration of the construction phase (years)	Uniform (1,7)
3.	Duration of ROW phase (years)	Uniform (1,3)
4.	Duration of project engineering phase (years)	Uniform (1,5)
5.	Cost of project engineering phase (dollars)	Triangular (2267, 62695, 3429757)
6.	Cost of construction phase (dollars)	Triangular (3000,350238,15199909)
7.	Cost of ROW phase (dollars)	Triangular (1000,98502, 1999884)
8.	RPC criterion of economic vitality (Scale 1–10)	Triangular (7, 9, 10)
9.	RPC criterion of safety and security (Scale 1–10)	Triangular (6, 8, 10)
10.	RPC criterion of accessibility, mobility and connectivity (Scale 1–10)	Triangular (6, 8, 10)
11.	RPC criterion of environment, energy and quality of life (Scale 1–10)	Triangular (3, 6, 9)
12.	RPC criterion of preservation of existing system (Scale 1–10)	Triangular (2, 5, 8)
13.	RPC criterion of efficient system management (Scale 1–10)	Triangular (5, 7, 9)
14.	RPC criterion of TIP status (Scale 1–10)	Triangular (3, 7, 10)
15.	V-TRANS criterion of highway system (Scale 1–100)	Triangular (25, 50, 75)
16.	V-TRANS criterion of cost per vehicle mile (Scale 1–100)	Triangular (25, 55, 85)
17.	V-TRANS criterion of project momentum (Scale 1–100)	Triangular (5, 35, 65)
Local town		
18.	Local Town ID	Unique Local Town Identifier
Regional planning commission (RPC)		
19.	Regional commission ID	Unique regional commission identifier
20.	Local town distribution in RPC	Triangular (30, 60, 80)
State agency (SA)		
21.	Percentage of projects to be funded each year	0.15 (15%)
22.	RPC distribution in the state	11
23.	Weight on regional priority	0.2
24.	Weight on highway system	0.4
25.	Weight on cost per vehicle mile	0.2
26.	Weight on project momentum	0.2
27.	New projects added each year	60

The ABM calculates annualized flow of financial resources from the state government to regional and local town jurisdictions that is contingent upon the project prioritization decision-making undertaken by the intergovernmental network of regional and state agents. Approvals of projects for different regions are characterized on an annualized basis and take into account delays and queues inside the system. Observed project cost distributions for each phase of the project in Vermont were used to calibrate the computer simulation model. The findings from the experimental simulations presented in this study, however, are generated from the model that is calibrated to reflect the institutional design structure of intergovernmental decision-making in the state of Vermont, its RPCs, and local towns. Through a user interface for the ABM, the users (e.g., researchers, policymakers, managers, and other stakeholders) can run reference and alternate scenarios by defining the parameters for different scenario runs as a policy informatics platform.

5. Findings from Experimental Simulations

All the results reported below are averages of 1,000 realizations for each model run scenario. Since the ABM is initialized with random parameters as shown in Table 3, 1,000 realizations for each scenario model run provide a robust set of findings. The model interface also allows decision makers to run up to 5,000 realizations for each scenario but that requires higher computational capacity. There was not much difference between 1,000 and 5,000 realizations of each of the scenarios reported below; hence, we decided to report 1,000 realization results that could be replicated on any modern computer. Furthermore, there are many possible scenario runs by varying the 27 input parametric values. Here, we focus on reporting the findings from the scenario runs that address the two research questions posed above. In particular, we focus on reporting the funding allocations at RPC and local town level for each scenario run. All scenarios are constrained to a 50-year simulation horizon that could be reduced or expanded if desirable by the stakeholders. Fixed seed runs were imposed to compare the simulation outputs across the scenarios. The reference scenario is run with the default parametric values reported in Table 3. Figure 5a shows the allocation of funds at the RPC level and Figure 5b shows the allocation of funds at the local town level under the reference scenario. Noticeably, this reference scenario is calibrated against baseline weights between regional and state project prioritization criteria, as reported in Table 1. Another important component of the model calibration process is driven by the observed probability distributions for the project class on the input parameters reported in Table 3 (e.g., probability distributions for the state agent criteria and RPC criteria assessments of the projects submitted by local towns).

The reference scenario, and all other scenario run outputs discussed below, presents three types of output information: first, as shown in Figure 5a, the time stack charts show simulated allocation of US $ (Nominal) for all 11 RPCs in the annualized time-step of the 50-year model simulation horizon. Second, the time plot in Figure 5a (and all other scenario runs in the Results section) presents the mean percentage of total projects funded per RPC per year. More successful RPCs in a given year have a 15% or higher success rate for funding the projects, whereas less successful RPCs have a relatively lower (<15%) success rate, i.e., less than 15% of their submitted projects get funded and implemented. Second, as shown in Figure 5b, the spatial grids present four distinct states of local towns at the end of the 50-year simulation horizon: (a) yellow pixels represent towns that have successfully funded and implemented all of their submitted projects, (b) red pixels represent towns that keep on waiting to get even a single project funded during the simulation horizon, (c) black pixels represent towns that have not submitted even a single project during the simulation horizon, and (d) green pixels represent local towns that have at least one of their submitted projects funded and under construction. We have the ability to replace this hypothetical grid-based representation of the local towns with actual local towns for more practitioner-oriented ABMs; however, the underlying stochastic nature of this ABM implies that the forecast for each local town might not be as precise as would be expected in more process-oriented deterministic simulation models. Our focus here is rather to ascertain the big picture project funding and implementation patterns across local towns, for which purpose such hypothetical grids are traditionally used in the ABM community.

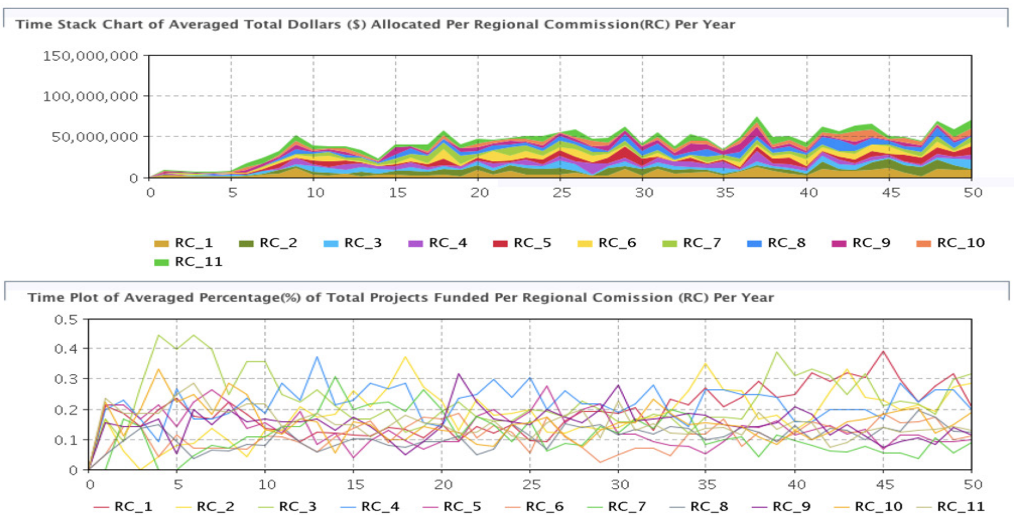


Figure 5a. Simulated allocation of roadway project funding across RPCs under reference scenario (years on X-axis).

Figure 5b. Simulated success rates of roadway projects across local towns under reference scenario.

5.1. Federalism Versus Regionalism Versus Localism Debate

To highlight how intergovernmental authority plays a role in the distribution of transportation project, Figure 6a and b reports the simulated allocations if the intergovernmental policy implementation network was to raise the regional priority weight from 20% to 60%, while the criterion of "project momentum" is reduced from 20% to 10%, the weight on highway system is reduced from 40% to 20% and cost per vehicle mile is reduced from 20% to 10%. Remaining parameters are similar to the reference scenario. A comparison of simulated allocations between Figures 5a and 6a scenarios demonstrate no specific difference in allocation of resources across RPCs. More emphasis on regionalization, however, changes the success rate of funding by different towns (as can be seen by comparing Figure 5b with 6b).

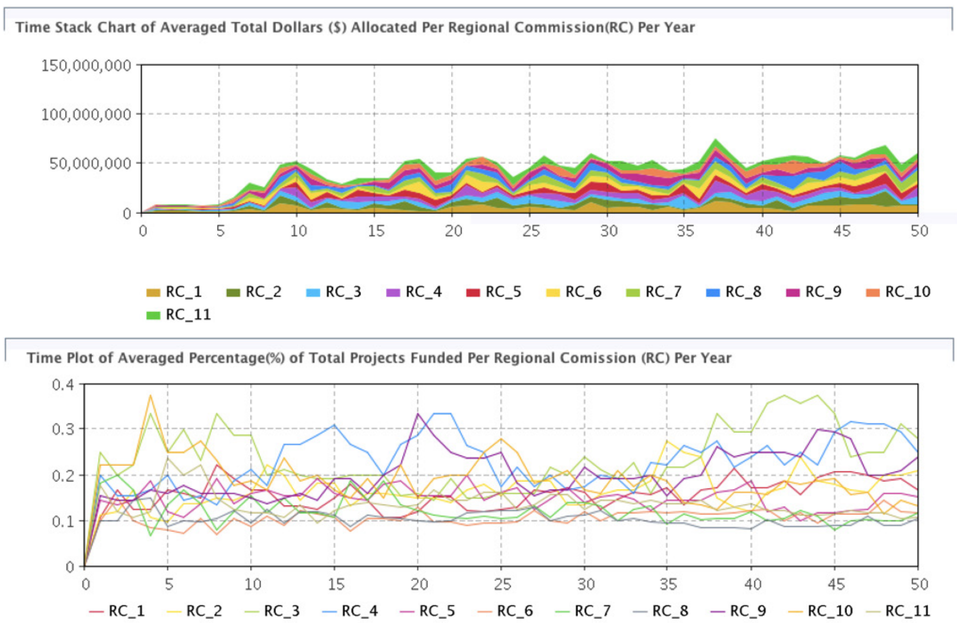


Figure 6a. Simulated allocation of roadway project funding across RPCs under increased regional priority.

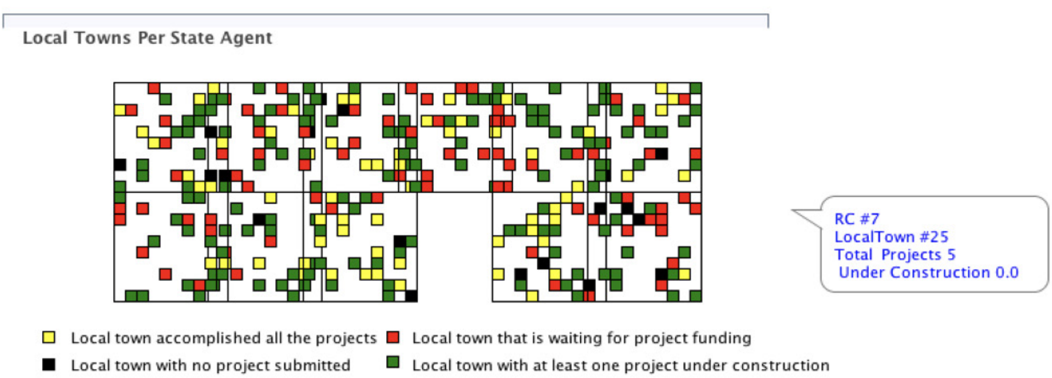


Figure 6b. Simulated success rates of roadway projects across local towns under increased regional priority.

5.2. Resilience Assessment

While there are numerous sources of potential external and internal shocks to test the resiliency in the intergovernmental policy implementation network, we single out two potential sources of potential shocks: (i) increased extreme events leading to the societal demand for funding higher number of projects and (ii) funding sequestration shocks leading to a decreased rate of projects funded each year. Here, we demonstrate the ability of this ABM to quantify the effects of these two shocks in the system. Figure 7a and b shows simulation output for a scenario where we increase the number of new projects to be added each year from 60 to 120, while the remaining parameters are same as reported for the baseline scenario. There are two counterintuitive findings with this scenario: first, despite the success rate of the projects being kept at 15%, the resource allocation levels almost double from \$50 million steady state (Figure 6a) to \$100 million (Figure 7a). Second, the variability in the success rate of projects decreases under the scenario shown in Figure 7a compared with 6b.

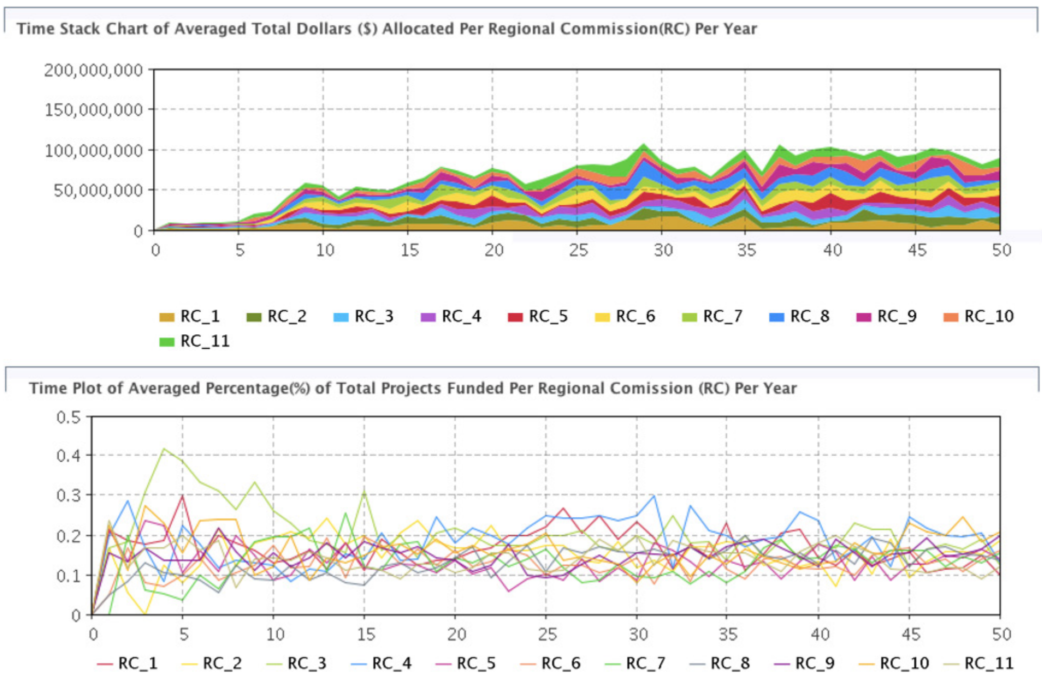


Figure 7a. Simulated allocation of roadway project funding across RPCs under increased number of new projects needing funding every year.

Figure 7b. Simulated success rates of roadway projects across local towns under increased number of new projects needing funding every year.

In Figure 8a and b, we add sequestration effect on top of the increased number of projects (raised from 60 to 120 new projects per year), which implies that the average number of projects funded each year reduces from 15% to 5% during the simulation horizon. Another way to assess the resilience in the system is to model the changes in the federal funding allocations in the system. Figure 8a and b represents the reference scenario project prioritization criteria, except that the parameter on the Percentage of Projects to be Funded Each Year is reduced from 15% to 5% and the number of new projects added each year is increased from 60 to 120. This leads to a much higher number of red and black dots in Figure 8b, i.e., much lower number of towns get at least one of their submitted projects approved for funding. While the funding allocation stabilizes at $50 million per year, experimental simulations point out to a convergent pattern of decreased variability in the success rate of projects across RPCs.

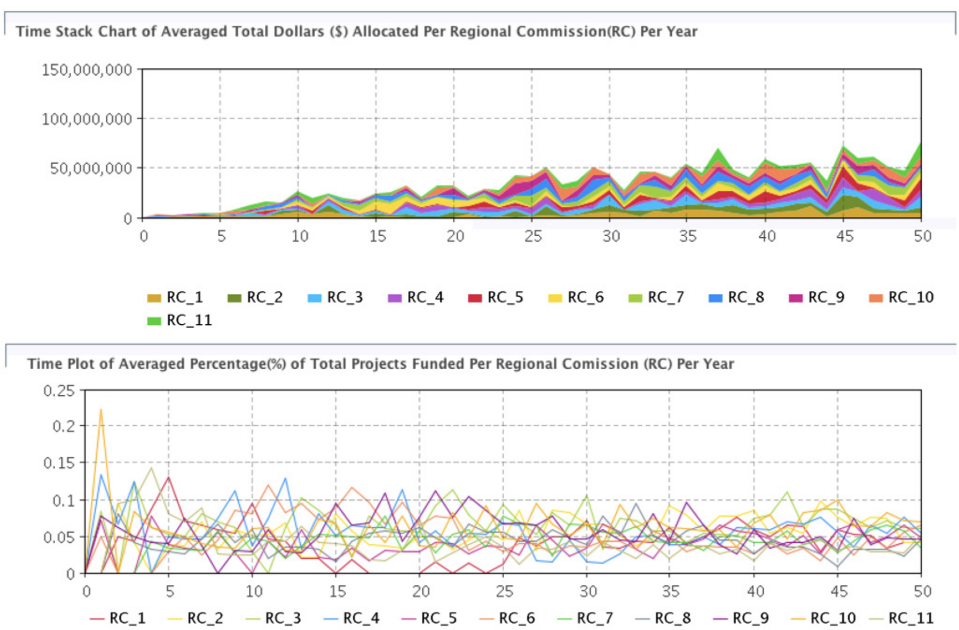


Figure 8a. Simulated allocation of roadway project funding across RPCs under higher number of new projects and decreased funding availability.

Figure 8b. Simulated success rates of roadway projects across local towns under higher number of new projects and decreased funding availability.

6. Limitations of the Current ABM and Next Steps

All models are abstractions of reality. It is the function and purpose of designing computer simulation models that broadly define the boundaries of the dynamic policy and governance system that is abstracted out of the complex reality. The main purpose of designing the ABM presented in this study is to simulate the decision-making dynamics of project prioritization processes among multilevel government agencies. The baseline institutional designs and rules that govern the interactions of intergovernmental governance dynamics are the focus of this study, yet the system boundaries of the ABM presented in this study led us to simplify and in some cases even ignore the institutional rules that might have direct bearing on the model outcomes presented in this study. The geophysical boundaries of ABM are limited to Vermont. We intentionally excluded five other project classes and just focused on roadways. While this simplification allowed us to focus on modeling the institutional rules across different levels of government, the competing dynamics that occur in terms of allocation of funding across different project classes (i.e., between bridges and roadway and/or bike/pedestrian and park and rides) have been ignored in setting up the system boundaries.

There are many possible ways, both vertically and horizontally, to extend and generalize the ABM presented in this study. Within the current transport policy domain, the ABM could be made spatially explicit by adding Geographic Information System (GIS) layers. Furthermore, the ABM could be extended to all six transport project classes. Explicit rules of federal programs, such as Surface Transportation Program (STP), Interstate Maintenance (IM) and others, could be incorporated in an upward expansion of the model. Similarly, in a downward expansion of the model, the complex network dynamics that occur in local towns and their planning commissions/boards could be explicitly captured using social network analysis techniques (Wasserman & Faust, 1994). Furthermore, even at a finer grain level, the outcomes of the funded projects could be ascertained by coupling integrated land-use transportation models such as UrbanSim (Waddell, 2002), with the policy implementation network simulation model presented in this study. Finally, as well known in the policy implementation literature since Pressman and Wildavsky (1984), financial capital flows that ensue from federal legislative actions to states and local governments typically entail development and execution of discrete projects. Since projects are the fundamental unit of analysis in the ABM presented in this study, similar context-specific ABMs could also be developed in other policy domains, e.g., energy, health, watershed, education, environment, and international development.

7. Conclusions

In this study, we introduced pattern-oriented agent-based modeling as a complex adaptive system-based policy informatics tool to compare alternate institutional designs of intergovernmental decision-making processes and assess transportation system resiliency under different shocks to the policy system. An ABM of transportation governance network in the state of Vermont has been presented to demonstrate an application of intergovernmental policy implementation network analysis in real-world public policy settings. This ABM enables simulation of the dynamics of transportation project prioritization processes under alternate intergovernmental institutional rule structures and variable frequencies of shocks to the policy system. Furthermore, this ABM enables decision makers to visualize the impacts of alternate intergovernmental institutional rules on the emergent patterns of financial investment flows from federal to state, regional, and local scale governments. The ABM presented in this study could be extended to test system-wide effects of alternate institutional designs on the differential emergence of project prioritization patterns by adding and/or re-weighting transport project prioritization criteria. Furthermore, if the goal were to extend the ABM to other states or countries, the probability distributions for these parameters would need to be empirically estimated for calibrating ABM to that particular geographical scope. Such computational models could also be developed to simulate public policy implementation processes in other policy domains for informing the governance of policy implementation networks. The ABM model parameters could be re-defined and re-calculated to adapt the model design to any intergovernmental policy implementation network in the United States that, for example, governs energy or food policy implementation systems. The computational modeling of intergovernmental policy implementation networks can improve our understanding about the apparently chaotic emergent patterns that arise due to overlapping and sometimes competing interests among different levels of the government, vertically; and among governments, markets, and societies, horizontally. Experimental simulations generated through such computational models could be harnessed as policy informatics tools in a wide variety of policy domains across multiple scales of governance and can be potentially used to inform the design of resilient infrastructures in the face of climate change induced extreme events and funding uncertainties under different political cycles.

Acknowledgments

We acknowledge funding support from NSF grant # EPS-1556770. Findings reported in this paper are solely the responsibility of the authors and not the funding agencies.

References

Agranoff, R. (2007). *Managing within networks: Adding value to public organizations.* Washington, DC: Georgetown University Press.

Agranoff, R., & McGuire, M. (2003). *Collaborative public management: New strategies for local governments.* Washington, DC: Georgetown University Press.

Bankes, S. (2002). Tools and techniques for developing policies for complex and uncertain systems. *Proceedings of the National Academy of Sciences of the USA*, *99*(3), 7263–7266.

Baumgartner, F. R., & Jones, B. D. (1993). *Agendas and instability in American politics.* Chicago, IL: The University of Chicago Press.

Epstein, J. M. (2006). *Generative social science: Studies in agent-based computational modeling.* Princeton, NJ: Princeton University Press.

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., … Huth, A. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, *198*(1), 115–126.

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S. F., … DeAngelis, D. L. (2005). Pattern-oriented modeling of agent-based complex systems: Lessons from ecology. *Science*, 310, 987–991.

Hunt, A., & Watkiss, P. (2011). Climate change impacts and adaptation in cities: A review of the literature. *Climatic Change*, *104*(1), 13–49.

Jaroszweski, D., Chapman, L., & Petts, J. (2010). Assessing the potential impact of climate change on transportation: The need for an interdisciplinary approach. *Journal of Transport Geography*, *18*(2), 331–335.

Johnson, I. (2012). *Adapting Vermont's transportation infrastructure to the future impacts of climate change: VTrans climate adaptation white paper.* VTrans. Retrieved from http://vtransplanning.vermont.gov/sites/aot_policy/files/documents/planning/Climate%20Change%20Adaptation%20White%20Paper.pdf

Koliba, C., Meek, J., & Zia, A. (2010). *Governance networks in public administration and public policy.* Boca Raton, FL: CRC Press/Taylor & Francis.

Ostrom, E. (2005). *Understanding institutional diversity.* Princeton, NJ: Princeton University Press.

Pressman, J. L., & Wildavsky, A. (1984). *Implementation: How great expectations in Washington are dashed in Oakland: Or, why it's amazing that federal programs work at all, this being a saga of the Economic Development Administration as told by two sympathetic observers who seek to build morals on a foundation of ruined hopes.* Berkeley, CA: University of California Press.

Regmi, M., & Hanaoka, S. (2011). A survey on impacts of climate change on road transport infrastructure and adaptation strategies in Asia. *Environmental Economics and Policy Studies*, *12*(1), 21–41.

Revi, A., Satterthwaite, D. E., Aragón-Durand, F., Corfee-Morlot, J., Kiunsi, R. B. R., Roberts, D. C., & Solecki, W. (2014). Urban areas. In C. B. Field, V. R. Barros, D. J. Dokken, K. J. Mach, M. D. Mastrandrea, T. E. Bilir, … L. L. White (Eds.), *Climate Change 2014: Impacts, adaptation, and vulnerability. Part A: global and sectoral aspects. Contribution of Working Group II to the fifth assessment report of the Intergovernmental Panel on Climate Change* (pp. 535–612). Cambridge, UK: Cambridge University Press.

Rhodes, R. (1997). *Understanding governance: Policy networks, governance, reflexivity and accountability.* Buckingham, UK: Open University Press.

Rittel, H. W. J., & Webber, M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, 4, 155–169.

Rittel, H. W. J., & Webber, M. M. (1984). Planning problems are wicked problems. In N. Cross (Ed.), *Developments in design methodology* (pp. 135–144). Chichester, UK: John Wiley and Sons.

Schmidt, N., & Meyer, M. D. (2009). Incorporating climate change considerations into transportation planning. *Transportation Research Record*, 2119, 66–73.

Schulz, A., Zia, A., & Koliba, C. (2017). Adapting bridge infrastructure to climate change: Institutionalizing resilience in intergovernmental transportation planning processes in the Northeastern USA. *Mitigation and Adaptation Strategies for Global Change*, *22*(1), 175–198.

Waddell, P. (2002). UrbanSim: Modeling urban development for land use, transportation, and environmental planning. *Journal of the American Planning Association*, *68*(3), 297–314.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications.* Cambridge, UK: Cambridge University Press.

Zia, A., & Koliba, C. (2015). The emergence of attractors under multi-level institutional designs: Agent-based modeling of intergovernmental decision making for funding transportation projects. *AI & Society*, *30*(3), 315–331.

Education System Intervention Modeling Framework

Pratik Mital[A], Roxanne A. Moore[B], Donna C. Llewellyn[C]

Abstract

Most educational policymakers and reformers do not rely on quantitative models when making resource allocation decisions for K-12 school interventions. However, schools and school districts are complex, dynamic systems affected by numerous factors. While no single model is applicable across a broad spectrum of school contexts, a unified framework can be used to build unique models for different school settings. In this work, the Education System Intervention Modeling Framework (ESIM) is presented—an approach for quantitatively modeling K-12 school interventions. In this article, techniques from such as agent-based modeling and social network analysis are used to model the bottom-up mechanisms of intervention implementation in schools. Variable screening methods reveal that there are several common factors likely to influence intervention sustainability, even with significant differences in the school environments. Some factors, such as turnover of teachers or administrators, can even have either a significantly positive or a negative impact, depending on the context. The framework can be used to assess risk factors and to guide decision-making prior to and during school interventions.

Keywords: education, interventions, complex systems, modeling, sustainability

---

[A] Revenue Analytics

[B] Georgia Institute of Technology

[C] Boise State University

doi: 10.18278/jpcs.3.2.5

摘要

大多数教育政策制定者和改革者在为K-12学校干预做资源分配决定时并不依赖定量模型。然而，学校及其地区是复杂的动态系统，该系统受多重因素影响。尽管在不同学校背景下无法适用单一模型，却可以使用一项统一的框架为不同学校背景建立唯一的模型。本文为此提出了一项教育系统干预建模框架（Education System Intervention Modeling Framework，简称ESIM），用于K-12学校干预的定量建模。为对自下而上的干预实施机制建模，本文运用了包括基于主体建模（agent-based modeling ，简称ABM）和社会网络分析（Social Network Analysis， 简称SNA）在内的技术。变量筛选方法显示，有好几个共同因素可能会影响干预的可持续性，就连不同学校环境存在显著差距时也是如此。在不同背景下，一些因素（例如老师或管理员的出勤率）可以产生显著的积极影响或消极影响。ESIM能用于评估风险因素，并在学校干预之前和干预期间指导决策。

关键词：教育，干预，复杂系统，建模，可持续性

Resumen

La mayoría de los reformadores y creadores de política de la educación no se apoyan en modelos cuantitativos al momento de tomar decisiones acerca de la asignación de recursos para las intervenciones en escuelas de niños menores de 18 años. Sin embargo, las escuelas y distritos escolares son sistemas complejos y dinámicos que se ven afectados por muchos factores. No hay un modelo único que se pueda aplicar a todos los diferentes tipos de contextos escolares. Sin embargo, un marco unificado puede ser utilizado para construir modelos especiales para diferentes tipos de escuelas. En este estudio se presenta el marco de modelización de intervención de sistemas educativos (ESIM) – Un método para la modelización cuantitativa de las intervenciones a escuelas de niños menores de 18 años. En este artículo, técnicas como la modelización basada en agentes (ABM) y el análisis de redes sociales (SNA) se usan para la modelización de mecanismos ascendentes de la implementación de intervenciones en las escuelas. Los métodos de examen variables revelan que hay varios factores comunes que tienen una alta probabilidad de influenciar la sustentabilidad de las intervenciones, inclusive con diferencias significativas de ambientes escolares. Algunos factores, como el tiempo de permanencia de maestros y administradores, puede incluso tener un efecto significativamente positivo o negativo, dependiendo del contexto. Este marco se puede usar para evaluar factores de riesgo y guiar la toma de decisiones antes de y durante las intervenciones escolares.

Palabras clave: educación, intervenciones, sistemas complejos, modelización, sustentabilidad

1. Introduction and Background

Every year, millions of dollars are spent on education interventions, and every year, many of those initiatives prove to be unsuccessful, unsustainable, or not scalable. From federal policies such as No Child Left Behind (2001), to research grants, to efforts of individual teachers and administrators, it is difficult to measure the outcomes of educational interventions and often more difficult to understand why they fail or are not sustainable (Fullan, 2000). In spite of the complexity of school systems, most educational policymakers and reformers do not rely on models to make decisions about where to intervene and how to manage resources during interventions. Such decisions often have unintended consequences or outcomes (Lubienski, 2005). Models of schools themselves, when designed carefully and used appropriately, can provide valuable insights about factors affecting intervention success and sustainability.

Education researchers, public policy experts, and complexity experts have studied different aspects of education systems; however, these three disciplines have rarely collaborated in order to get a more complete picture. Education research as a field has broadened from simply designing and evaluating interventions to studying implementation within the school system (Fixsen, Naoom, Blase, Friedman, & Wallace, 2005). This work has identified several factors that affect intervention implementation and outcomes, including professional development (PD), leadership, organization, school structure, resources, and support (Billig, Sherry, & Havelock, 2005; Blumenfeld, Fishman, Krajcik, Marx, & Soloway, 2000; Ely, 1990; Elmore, 1996; McLaughlin & Talbert, 2003; Spillane, Reiser, & Reimer, 2002). Social network analysis (SNA) has been applied to understand how a school's social structure and teacher networks affect intervention implementation (Coburn, Russell, Kaufman, & Stein, 2012; Daly, 2010; Moolenaar, 2012; Moolenaar & Daly, 2012). Moreover, experts in educational and public policy have studied reform agendas within social and organizational contexts (Borman, Carter, Aladjem, & LeFloch, 2004; Crawford & Ostrom, 1995; Weaver-Hightower, 2008). While these approaches aid in understanding education interventions and policies in general, they do not capture the underlying mechanisms and feedback loops affecting intervention implementation and sustainability.

Meanwhile, engineers and complexity experts have only recently turned their attention to the education realm. Recent efforts have applied system dynamics (SD) and agent-based modeling (ABM) to understand student interest and selection in STEM (Science, Technology, Engineering and Mathematics) (Allen & Davis, 2010; Sanchez, Wells, & Attridge, 2009). ABM has emerged as a popular technique for modeling social systems because it captures emergent behaviors, is a natural description of a system, and is flexible enough to accommodate different scales temporally and spatially (Bonabeau, 2002). Other researchers have promoted the use of SD for studying education as a complex system (Groff, 2013). These models rely on survey data to formulate causal relationships, but often lack a mechanism for distinguishing correlation from causation (Pedamallu, Ozdamar, Akar, Weber, & Özsoy, 2012; Pedamallu, Ozdamar, Ganesh, Weber, & Kropat, 2010). To understand the effects of an intervention on a particular school system, systems engineering, policy, and education research approaches need to be combined, leveraging SD, ABM, and SNA where appropriate (Maroulis et al., 2010; Sarkis, 2012).

No single model is applicable across a broad spectrum of schools or interventions, and attempting to develop a one-size-fits-all model would marginalize the nuances of different schools and the dynamics of the combination of a school and a particular intervention (Barab & Squire, 2004). In this paper, we present and demonstrate Education System Intervention Modeling Framework (ESIM), a framework that can guide model development for specific intervention implementations. By applying the framework to interventions across diverse school settings, variables can be identified that are likely to affect intervention implementation and sustainability. The resulting models allow decision makers to make informed resource allocations during interventions, assess risk factors in different school settings, and ensure that interventions are sustainable and scalable.

2. Modeling Sustainability of Interventions

The intent of this modeling effort is to understand whether an education intervention is sustainable. Specifically, can the intervention be carried forward by a school without additional resources from the interveners? Many interventions occur with the aid of a federal or state-level grant, which brings significant personnel and financial resources to bear. However, grants have finite lifespans. If we can identify school partners more intelligently and understand the factors most likely to affect the sustainability of the intervention, resources can be allocated strategically. To quantify a sustainable intervention outcome, a conceptual study from the educational public policy literature is utilized (Blumenfeld et al., 2000). This study presents three gaps affecting intervention implementation: policy management (PM), capability (Ca), and cultural (Cu) gaps. In this work, we assign appropriate variables to each gap and assess the difference between the actual state of an organization and the "ideal" state. In each of these three dimensions, we define an acceptable "tolerance" with respect to the ideal state that constitutes the "acceptable zone." The acceptable zone bounds the set of end states in which the education researchers and practitioners believe that sustainability can be achieved. We can then simulate different school settings with different initial conditions, assess the likelihood of sustainability after a grant period, and determine the factors having the biggest impacts on the school's intervention capacity.

3. Education System Intervention Modeling Framework

The ESIM framework has four stages: model definition, model design, model analysis, and model validation. We present the ESIM framework stages in conjunction with a case study intervention called "Science Learning: Integrating Design, Engineering, and Robotics" (SLIDER), an eighth grade physical science curricular intervention promoting inquiry learning using LEGO® robots (Usselman, Aguilar, Llewellyn, & Ryan, 2009). This NSF-funded 5-year intervention was implemented in three different middle schools, each exhibiting very different demographics and cultures. In this paper, we refer to these schools as "Rural School," "Suburban School," and "Lab School." Lab School is treated as an ideal reference point because the school is always within the acceptable zone. While an ideal or near-ideal school setting is useful

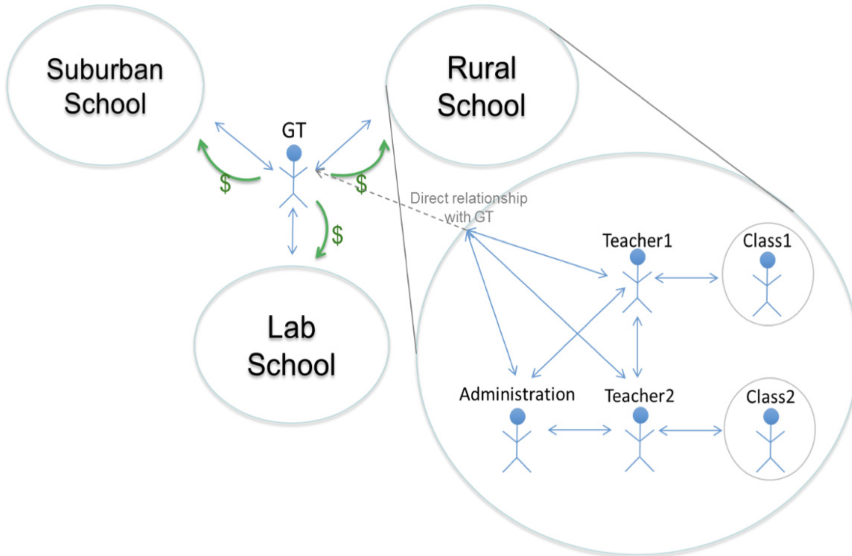


Figure 1. Agent network for SLIDER case study.

in terms of having a "control" variable, a common fallacy is to design an intervention only within this type of setting, only to have it fail in settings that are more complex. In terms of scalability, it is important to design and refine interventions in schools with more constraints. The other two schools exhibited much more variation and change during the course of the intervention. Through collaboration with the SLIDER team at Georgia Tech (GT) (educational researchers and practitioners) and public policy researchers, we apply the ESIM framework to the SLIDER case study.

3.1 Model Definition

Figure 1 shows the agent network for SLIDER for two scales of analysis simultaneously. At the macroscale, GT is interacting with three schools and allocating resources dynamically. Within each school, the agent networks contain teachers, students, and administrators. The blue arrows represent relationships, and the green arrows represent resource flows. The resource flows for SLIDER include PD (time) and supplies from GT. The attributes of the agents and school identified as potentially relevant by the SLIDER development and research team are shown in Table 1.

<u>External trends</u> are changes occurring in schools that are independent of the intervention. These trends are modeled as external forces on the school that are not impacted by the intervention. For SLIDER, there were two primary external forces considered: student–teacher ratio and test score trends. While the student–teacher ratio at Rural School was constant, the student–teacher ratio at Suburban School increased significantly after the start of the intervention, likely due to the economic recession and layoffs occurring in 2008–2010. The eighth grade science test scores at the two schools were examined for trends and were compared with district performance to assess external pressures on the schools. For Rural School, test scores were comparable with overall district performance, but at Suburban School, the school average was significantly below the district average. Because of this gap between school and

Table 1. Agents and their attributes for the SLIDER case study

Agent	Attributes	
Class	• SLIDER pre-/post-test scores • Percent free and reduced lunch	• Science test scores • Class size
Teacher	• Inquiry teaching competency • Content knowledge • Support for intervention	• Self-efficacy • Organizational citizenship
Administration	• Leadership	• Support for intervention
GT	• Total budget • PD budget allocation	• Material and supply budget allocation
School environment	• Class duration • Classroom layout/size • Supply availability • Threat rigidity	• Student–teacher ratio • Teacher preparation time • Teacher turnover rate • Administration turnover rate

district performance, threat rigidity is likely to have occurred in Suburban School. Threat rigidity (Staw, Sandelands, & Dutton, 1981) is an organizational behavior term that correlates external pressures on an organization with a reversion to tried and tested norms, generally at the expense of innovation.

For SLIDER, the acceptable zone for sustainability is calculated as the difference between the highest (best possible) and the second highest attribute levels associated with a gap when measured on a 5-point Likert scale. The attributes associated with each of the three gaps are shown in Figure 2 and were measured using survey data or subject matter experts (SMEs) familiar with the intervention implementation.

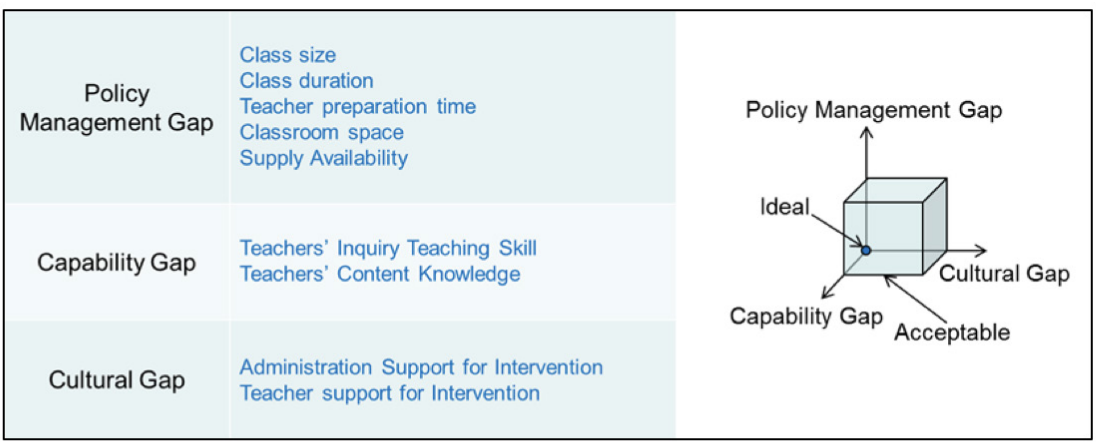


Figure 2. Acceptable zone modeled using PM, Cu, and Ca gaps.

3.2 Model Design

Changes in the agents' attributes and composite system state are modeled as a Discrete-Time Markov Chain (DTMC). It is assumed that change can be modeled as taking place in discrete time steps and that changes in the agents' attributes depend only on the current attribute level and current relationships, not on past system states. The discrete time step for this intervention is generally defined as one1 month. However, for the first three3 months of the school year starting in year three3, the time step is reduced to one1 week to accommodate increased frequency of interactions among the agents during active implementation.

The following equation represents the general structure of the changes taking place in the model:

$$p_{change}(t) = w_{internal} \cdot p_{internal}(t) \cdot f_s + w_{external} \cdot pexternal\ (t) \qquad (1)$$

where p_{change} is the overall probability of change at time t. The overall probability of change includes both an internal and external component. The internal component includes aspects of change affected directly by the intervention, and its weight is a function of the scale of the intervention. The structure of this equation is based on the assumption that every school environment has factors that lie beyond the scope of the intervention. However, the weighting allows a modeling decision to be made; if the intervention is school-wide or district-wide, w_internal will be greater than for an intervention involving only a single class or teacher. The effects of the weighting can be tested using sensitivity analysis. The probabilities are vector quantities representing the three possible state changes,$[p_{improve}, p_{stay}, p_{worsen}]$, which sum to 1. In addition, the weights, $w_{internal}$ and $w_{external}$, are non-negative and sum to 1. Finally, the internal part of the equation also includes a multiplicative factor f_s which captures an '"S"' curve pattern in learning. This curve represents cumulative adoption of change in a complex adaptive system, meaning that it is harder to enact state change when an agent is already in a very high or very low state (Rogers, Medina, Rivera, & Wiley, 2005).

The internal probability of change $p_{internal}(t)$ is further divided into two parts: *transient* and *steady state*. It is assumed that a school is in steady state prior to an intervention. As implementation of the intervention begins, the school enters a temporary transient phase where change is more likely to occur, and then again enters a steady state after a certain number of time steps. The following equation represents the structure of the internal component of the change equation:

$$p_{internal}(t)=e^{-kt} \cdot p_{transient}+(1-e^{-kt}) . \ p_{steady} \qquad (2)$$

where k is the transient parameter affecting how long the system stays in the transient phase, and $p_{transient}$ and p_{steady} are the transient and steady steady-state sub-components, respectively, of the internal component of the change probability,$p_{internal}(t)$. In the transient state, the absolute values of the agents' attributes affect change in other agents, whereas in steady state, change in an agent only occurs when there is a *change* in other agents' attributes for those agents with whom there are relationships. This modeling choice dampens the effects of positive and negative feedback loops over time. Note that:

$$\lim_{(t \to \infty)} p_i(t)=p_{steady} \qquad (3)$$

Rules: Change in Attributes

The internal change probability, $p_{internal}(t)$, for a particular agent is a function of three inputs: the current attributes of the agent, the agent's relationships with other agents, and the attributes of the agents with whom the agent has relationships. As an example, for the SLIDER case study, the internal probability of change equations for '"support for intervention"' of a teacher is the following:

$$p_{transient\_support}(t+1) = w_{teacher} \cdot \underbrace{teacher(t)}_{attributes} + w_{admin} \cdot \underbrace{relationship(t)}_{teacher:admin} \cdot \underbrace{admin(t)}_{support} \quad (4)$$

$$p_{steady\_support}(t+1) = w_{teacher} \cdot \underbrace{I_{teacher}(t)}_{attributes} + w_{admin} \cdot \underbrace{relationship(t)}_{teacher:admin} \cdot \underbrace{I_{admin}(t)}_{support} \quad (5)$$

$$\text{where } I_x(t) = \begin{cases} 1; \text{ if } x(t) - x(t-1) > 0 \\ 0; \text{ if } x(t) - x(t-1) = 0 \\ -1; \text{ if } x(t) - x(t-1) < 0 \end{cases}$$

In the above equations, $p_{transient\_support}(t+1)$ and $p_{steady\_support}(t+1)$ represent the transient and steady steady-state change components for teacher's support for intervention in period $t+1$. $I_x(t)$ is a modified indicator function with x being a given normalized attribute value; this term is used only in the steady steady-state equations and is used to capture attribute changes. The first term in both transient and steady steady-state equations reflects the teacher's own attributes, and $w_{teacher}$ is the weight for this term. All attribute values are initialized using survey or SME data collected on a Likert scale and then normalized. The second term captures the relationship of the teacher with the administration, where: w_{admin} is the weight for this term, and

$$\underbrace{relationship(t)}_{teacher:admin} \quad \text{and} \quad \underbrace{admin(t)}_{support}$$

represent the relationship strength between the teacher and administration and the administrative support for intervention at time t, respectively. Again, $w_{teacher}+w_{admin}=1$ and the effects of the weights are tested using sensitivity analysis. In the above equations and the equations presented next, the function $p(.)$ can take either positive or negative values. If $p(.)$ is positive, then it represents the probability of the attribute to improve ($p_{improve}$), and 1-$p(.)$ represents the probability of the attribute to stay the same (p*stay*). If $p(.)$ is negative, then the absolute value $|p(.)|$ represents the probability of the attribute to worsen (p_{worsen}), and 1-$|p(.)|$ represents the probability of the attribute to stay the same (p_{stay}). The above equations are representative of change equations for modeled attributes.

Rules: Effect of Resource Flows

`One resource flow from GT to the school agents is in the form of Professional Development (PD), which affects teachers' pedagogical skills with respect to inquiry-based learning and content knowledge. A teacher's inquiry-based teaching skill is a multi-faceted attribute capturing specific pedagogical practices, including student-centered structuring of lessons where the teacher serves as a facilitator rather than a lecturer (Llewellyn, 2014). A teacher's inquiry skill is also assumed to be affected by relationships with other teachers in the school receiving PD. In the example below, there are two teachers at a single school who are receiving PD from Georgia TechGT, and the equations model the change in inquiry teaching skill of teacher1.

$$p_{transient\_inq}(t+1) = w_{teacher2} \cdot \underbrace{rel(t)}_{teacher1:2} \cdot \underbrace{teacher2(t)}_{inquiry} + w_{pd} \cdot GT_{pd}(t) \cdot \underbrace{teacher1(t)}_{self-eff\,\&\,sup} \quad (6)$$

$$p_{steady\_inq}(t+1) = w_{teacher2} \cdot \underbrace{rel(t)}_{teacher1:2} \cdot \underbrace{I_{teacher2}(t)}_{inquiry} + w_{pd} \cdot GT_{pd}(t) \cdot \underbrace{I_{teacher1}(t)}_{self-eff\,\&\,sup} \quad (7)$$

In the above equations, $p_{transient—inq}(t+1)$ and $p_{steady\_inq}(t+1)$ represent the transient and steady steady-state change equation components for teacher1 inquiry teaching in period $t+1$. $I_x(t)$ is the same indicator function defined in equation 5. The first term in both equations captures the relationship between *teacher1* and *teacher2*, where $w_{teacher2}$ is the weight for this term,

$$\underbrace{rel(t)}_{teacher1:2}$$

is the relationship between *teacher1* and *teacher2* at time *t*, and

$$\underbrace{teacher2(t)}_{inquiry}$$

is the inquiry teaching level of *teacher2* at time t. The second term is for PD given to the teacher where: wpd is the weight for this term, $GT_{pd}(t)$ is the PD level at time period t, and

$$\underbrace{teacher1(t)}_{self-eff\ \&\ sup}$$

is the weighted average of the self-efficacy and support for intervention of teacher1 at time *t*. Again, $w_{teacher2}+w_{pd}=1$. The above equation illustrates the overall structure of change equations where an agent's attributes is affected by three things: its own attributes, its relationship with the other agents, and attributes of other agents. The only difference is the addition of $GT_{pd}(t)$, the resource flow, to the second term. This resource flow variable allows us to capture differences in PD quality and duration for different interventions.

<u>*Rules: Change in Relationships*</u>

In addition to attributes, relationships also change over the course of the intervention: relationships can become more positive, stay the same, or may sour. Two concepts from social network theory aid in modeling relationship changes: homophily and structural balance (Davis, 1963; Heider, 1958). Homophily assumes that individuals are more likely to form ties with other individuals who are similar to them in a variety of ways including demographics, hobbies, and interests. Structural balance assumes that individuals are more likely to form positive ties with friends of friends, negative ties with friends of enemies, and positive ties with enemies of enemies.

For the SLIDER intervention, we have explicitly modeled the change in relationships among the teacher, administration, and GT. In this case, the relationship between the school administration and GT is affected by the administration's support for intervention. The following equation represents the change in relationship between the administration and GT:

$$p_{change\_rel\_admin\_GT}(t+1) = w_{admin} \cdot \underbrace{I_{admin}(t)}_{support} \tag{8}$$

In this equation, $p_{change\_rel\_admin\_GT}(t+1)$ represents the internal change equation for the relationship between administration and GT at time *t*+1. $I_x(t)$ is the same indicator function defined in equation 5, and *x* is the support for intervention attribute. Here, w_{admin} is a weight and is equal to 1 because no other factors affect relationship quality in this case.

These equations, when extended to the other agents, attributes, and relationships, complete the model for the SLIDER intervention at the individual schools. Before moving forward into analysis, the model is validated using SMEs, including the SLIDER research team, public policy researchers, and one of the teachers involved in the intervention.

3.3 Model Analysis

In this phase of the framework, simulation results were generated and analyzed. Data collected during the intervention were used to populate the initial state of the simulation model and to validate the results of the simulation model. Data included demographics, experience, and content knowledge of teachers, classroom observations, student pre-post SLIDER test scores, class sizes, and support for intervention of the teachers and administration. Other data were acquired through the Georgia Department of Education. Any remaining data needed for these models were acquired through surveys given to the SLIDER team at GT and the teachers and administration at the schools involved in the intervention. The simulation model was built using the object-oriented programming language C# in Microsoft Visual Studio 2010 and run on a personal computer. The stochastic models were run 1000 times and averaged to approximate a deterministic output, which is needed for sensitivity analysis and variable screening methods. Complete sensitivity analysis results for the SLIDER case study can be found in (Mital, 2015), where the model was found not to be unusually sensitive to the selection of any one parameter.

3.4 Model Validation

The verification and validation steps implemented during the different phases are discussed briefly here. Some of these were compiled by Sargent in his work on validation and verification of simulation models (Sargent, 2004).

Conceptual model and face validation: Ten SMEs were used to test whether the model and its behavior were conceptually logical and whether the model's input–output relationships were reasonable. They examined the model for completeness, consistency, coherence, and correctness as described in the framework proposed by Pace (2000).

Data validation: The changes in the attributes of the agents were compared to the data collected with respect to these attributes.

Comparison to Other Models: The model results obtained are consistent with educational research studies discussed in the literature review.

Parameter variability—Sensitivity analysis: The model was run under different sets of parameter and input conditions and model outputs were analyzed.

4. Analysis and Insights

For brevity, individual attribute simulation results are not presented here, but can be found in (Mital, 2015). Figures 3 and 4 depict the changes in the gaps for Rural School and Suburban School, respectively, with the blue line representing the simulation results, the red line representing reality (as determined from survey data), and the green box representing the acceptable or sustainability boundaries as modeled. As can be seen in Figure 3, Rural School starts as a "higher risk" school in that the Ca gap is well outside of the acceptable zone, and the Cu gap is on the edge of acceptability. However, Rural School consistently moves toward the acceptable zone, and by the end of year 4, Rural School is within the acceptable zone or close to it for all three gaps. This is consistent with the survey data and reality—Rural School has maintained many aspects of the intervention since the end of the SLIDER grant. On the other hand, Suburban School, which started in a "low risk" state relative to Rural School, continues to worsen during the grant period, as shown in Figure 4. Because of external trends and decreasing administrative support, the gaps increase and it appears that the intervention will not be sustained beyond the grant period. This is also consistent with reality—no aspects of the intervention were maintained at Suburban School after the SLIDER grant completed.

One of the most important contributions of ESIM is the identification of attributes that most likely influence the intervention's sustainability. The Method of Morris (MoM) (Morris, 1991) is a one-at-a-time design of experiment technique for identifying the subset of variables having a statistically significant impact on a model outcome of interest. For SLIDER, the MoM experiment is conducted for the two schools with 43 different input variables in the initial state. The response variable used is overall fitness at the end of the intervention, an inverse linear combination of the three gaps. Figures 5 and 6 depict the means (on the *x*-axis) and standard deviations (on the *y*-axis) of the effects of different inputs on the overall fitness of Rural and Suburban School, respectively. In the figures, the dotted lines correspond to the equation: mean = ±2•SEM, where SEM is the standard error of the mean. Inputs lying outside of the "v" (or far from zero) have an impact on the response variable that is statistically different from zero. The major factors affecting a sustainable implementation of this intervention in both schools are the teacher's inquiry teaching, content knowledge, and self-efficacy, relationship between the administration and GT, class duration, classroom space, availability of supplies, teacher preparation time, and time spent by GT on PD. All of these attributes are positively correlated with fitness.

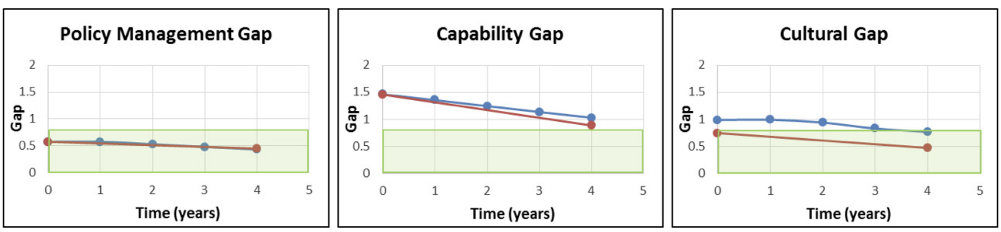


Figure 3. PM Gap, Ca gap, Cu gap in rural school. The gaps are normalized on a unitless scale where 0 is ideal and 2 is the most challenging.

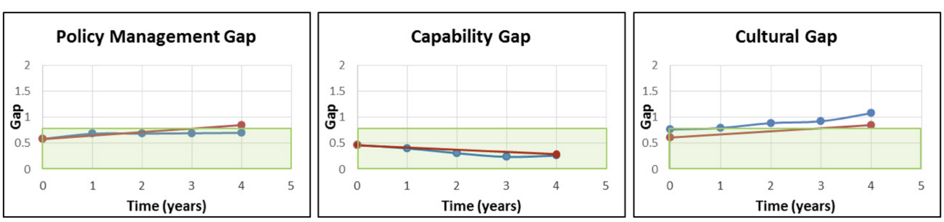


Figure 4. PM Gap, Ca Gap, Cu gap in suburban school. The gaps are normalized on a unitless scale where 0 is ideal and 2 is the most challenging.

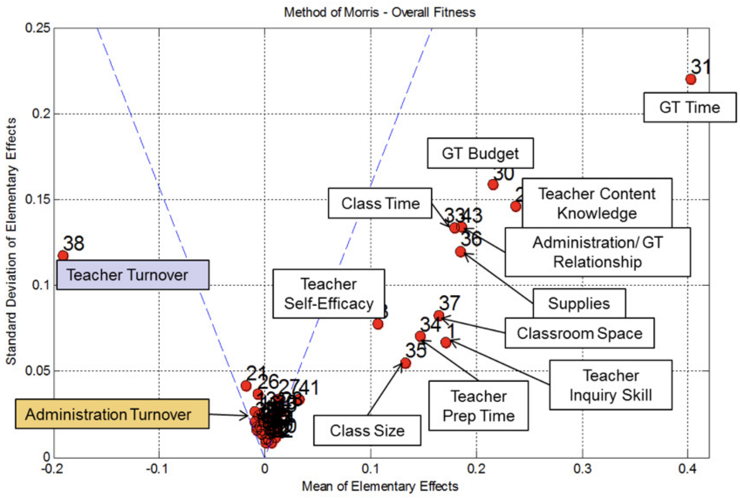


Figure 5. Barriers and enablers for rural school.

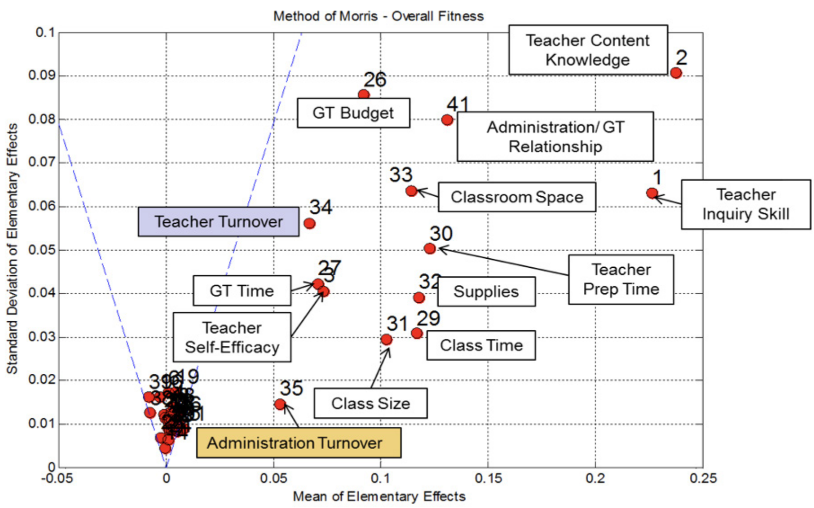


Figure 6. Barriers and enablers for suburban school.

One significant attribute that has a different effect in the two schools is "teacher turnover rate." This attribute has *negative* impact on the overall fitness in Rural School, but a *positive* impact on overall fitness in Suburban School (highlighted in Figures 5 and 6). This is a counterintuitive insight; logically, one would expect that the loss of resources invested in the teachers in a school would have a negative outcome on an intervention. In Suburban School, however, the culture attributes of the teachers are steadily declining because of the administration attributes and the environment. Therefore, the introduction of a new teacher potentially increases these attributes, *if* it is assumed that the new teachers would have comparable attribute levels to the original teachers at the start of the intervention.

These results did not align with the expectations of the research team at the beginning of the grant; without rigorous analysis of external trends and model-based thinking about attributes like turnover of teachers and administrators, Rural School seemed like a high-risk environment, while Suburban School seemed ideal. The end states were very different from where they started. Using the modeling framework may help future interveners get a more complete picture of schools that they collaborate with at the beginning of the intervention and understand the risks that they face with respect to intervention sustainability.

5. Policy Implications

There is a relatively new research methodology in educational research called Design-Based Implementation Research (DBIR), which tries to answer the questions such as: What type of interventions work, in which school settings? How can an intervention be made more sustainable? What capacities should the school system improve to facilitate the intervention (Fishman, Penuel, Allen, Cheng, & Sabelli, 2013)? Currently, educational and public policy researchers are trying to address these questions, but they may not be experts in the modeling and analysis techniques that complex systems researchers use. The questions posed by the DBIR community can be better answered if the education and public policy researchers work together with the engineers and complexity experts. Educational research provides knowledge about which factors affect change in the agents' attributes and their relationships and how to measure them. If an attribute arises as important for an intervention's sustainability, then the educational research community can focus on understanding and measuring that attribute, while safely ignoring those that are unlikely to be impactful and may be difficult or expensive to measure.

Another implication for educational and public policy researchers involved in the implementation of interventions is the characterization of the risk of intervening in a given school system. With the help of the gaps and identification of key variables influencing sustainability, the school can be categorized as either a high-risk, medium-risk, or a low-risk environment for a particular intervention. Having a diverse portfolio of risks allows an intervener to have a better understanding of what types of school settings are a good fit for the intervention and how to design the intervention to be adequately robust.

6. Conclusions and Future Work

In this paper, we have shown that quantitative models can be used to analyze interventions in K-12 school settings and can help guide decisions leading to more sustainable interventions. Through the application of the ESIM framework, key attributes likely to have impacts on the sustainability of an intervention can be identified. The SLIDER case study shows that a school's capacity for intervention is multidimensional and changes with time due to both the intervention itself and trends external to the intervention. Models may help inform school selection for future interventions; a diverse "portfolio" of different risks within schools can help test the limits of intervention impacts and sustainability. In addition, model insights may allow for more strategic allocation of resources during interventions.

In the future, the ESIM framework is well situated to be used as a policy tool by intervention implementation agencies. Currently, ESIM is being used in an NSF-funded intervention, EarSketch (Freeman, Magerko, & Edwards, 2014), which will eventually be implemented across 20 different schools in the 4-year grant period. This will allow for more rigorous testing and refinement of the framework as well as the development of simplified models that can be used prior to intervention work to assess risk factors in different school settings. In addition, the modeling effort is being applied in tandem with the intervention design and rollout, so the modeling efforts impact the thinking of the design team, rather than serving only as an explanatory mechanism after the implementation is completed (Helms, Moore, Edwards, & Freeman, 2016).

Acknowledgements

This work has been funded by National Science Foundation (NSF) award 1238089, Advanced Manufacturing and Prototyping Integrated to Unlock Potential (AMP-IT-UP). We thank Mike Ryan, Marion Usselman, David Goldsman, Gordon Kingsley, Meltem Alemdar, and all the other members of our interdisciplinary team.

References

Allen, T. T., & Davis, N. (2010). *A simple agent-based social impact theory model of student STEM selection*. In Proceedings of the 2010 Winter Simulation Conference (WSC), IEEE.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences*, *13*(1), 1–14.

Billig, S. H., Sherry, L., & Havelock, B. (2005). Challenge 98: Sustaining the work of a regional technology integration initiative. *British Journal of Educational Technology*, *36*(6), 987–1003.

Blumenfeld, P., Fishman, B. J., Krajcik, J., Marx, R. W., & Soloway, E. (2000). Creating usable innovations in systemic reform: Scaling up technology-embedded project-based science in urban schools. *Educational Psychologist*, *35*(3), 149–164.

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. Proceedings of the National Academy of Sciences of the United States of America, *3*(99), 7280–7287.

Borman, K. M., Carter, K., Aladjem, D. K., & LeFloch, K. C. (2004). *Challenges for the future of comprehensive school reform. Putting the pieces together: Lessons from comprehensive school reform research* (pp. 109–150). NCCSR. Retrieved from http://www.uky.edu/~gmswan3/609/Cairncross_Mannion_2001.pdf

Coburn, C. E., Russell, J. L., Kaufman, J. H., & Stein, M. K. (2012). Supporting sustainability: Teachers' advice networks and ambitious instructional reform. *American Journal of Education*, *119*(1), 137–182.

Crawford, S. E., & Ostrom, E. (1995). A grammar of institutions. *American Political Science Review*, *89*(3), 582–600.

Daly, A. J. (2010). *Social network theory and educational change.* ERIC. Retreived from https://eric.ed.gov/?id=ED515312

Davis, J. A. (1963). Structural balance, mechanical solidarity, and interpersonal relations. *American Journal of Sociology*, *68*(4), 444–462.

Elmore, R. F. (1996). Getting to scale with good educational practice. *Harvard Educational Review*, *66*(1), 1–27.

Ely, D. P. (1990). Conditions that facilitate the implementation of educational technology innovations. *Journal of Research on Computing in Education*, *23*(2), 298–305.

Fishman, B. J., Penuel, W. R., Allen, A. R., Cheng, B. H., & Sabelli, N. (2013). Design-based implementation research: An emerging model for transforming the relationship of research and practice. *National Society for the Study of Education*, *112*(2), 136–156.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. USF. Retrieved from http://nirn.fpg.unc.edu/sites/nirn.fpg.unc.edu/files/resources/NIRN-MonographFull-01-2005.pdf

Freeman, J., Magerko, B., & Edwards, D. (2014). *EarSketch: An authentic, studio-based STEAM approach to high school computing education* (NSF Abstract Award # 1417835).

Fullan, M. (2000). The three stories of education reform. *Phi Delta Kappan*, *81*(8), 581–584.

Groff, J. S. (2013). Dynamic systems modeling in educational system design & policy. *New Approaches in Educational Research*, *2*(2), 72–81.

Heider, F. (1958). *The psychology of interpersonal relations*. New York, NY: Wiley.

Helms, M., Moore, R., Edwards, D., & Freeman, J. (2016). *STEAM-based interventions: Why student engagement is only part of the story.* Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT), 2016, IEEE.

Llewellyn, D. (2014). *Inquire within: Implementing inquiry- and argument-based science standards in grades 3-8*. Thousand Oaks, CA: Corwin Press.

Lubienski, C. (2005). Public schools in marketized environments: Shifting incentives and unintended consequences of competition-based educational reforms. *American Journal of Education*, *111*(4), 464–486.

Maroulis, S., Guimerà, R., Petry, H., Stringer, M., Gomez, L., Amaral, L., & Wilensky, U. (2010). Complex systems view of educational policy research. *Science*, *330*(6000), 38–39.

McLaughlin, M., & Talbert, J. (2003). *Reforming districts: How districts support school reform.* Center for the Study of Teaching ad Policy. Retrieved from http://www.education.uw.edu/ctp/sites/default/files/ctpmail/PDFs/ReformingDistricts-09-2003.pdf

Mital, P. (2015). *A modeling framework for analyzing the education system as a complex system* (Doctoral dissertation). Retrieved from https://smartech.gatech.edu/bitstream/handle/1853/53557/MITAL-DISSERTATION-2015.pdf

Moolenaar, N. M. (2012). A social network perspective on teacher collaboration in schools: Theory, methodology, and applications. *American Journal of Education*, *119*(1), 7–39.

Moolenaar, N. M., & Daly, A. J. (2012). Social networks in education: Exploring the social side of the reform equation. *American Journal of Education*, *119*(1), 1–6.

Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, *33*(2), 161–174.

No Child Left Behind (NCLB) Act of 2001, Public Law Number 107–110, § 115, Statute 1425 (2002). 115.

Pace, D. K. (2000). Ideas about simulation conceptual model development. *Johns Hopkins APL Technical Digest*, *21*(3), 327–336.

Pedamallu, C. S., Ozdamar, L., Akar, H., Weber, G. W., & Özsoy, A. (2012). Investigating academic performance of migrant students: A system dynamics perspective with an application to Turkey. *International Journal of Production Economics*, *139*(2), 422–430.

Pedamallu, C. S., Ozdamar, L., Ganesh, L., Weber, G. W., & Kropat, E. (2010). A system dynamics model for improving primary education enrollment in a developing country. *Organizacija*, *43*(3), 90–101.

Rogers, E. M., Medina, U. E., Rivera, M. A., & Wiley, C. J. (2005). Complex adaptive systems and the diffusion of innovations. *The Innovation Journal: The Public Sector Innovation Journal*, *10*(3), 1–26.

Sanchez, H. A., Wells, B., & Attridge, J. M. (2009). *Using system dynamics to model student interest in science technology, engineering, and mathematics.* In Proceedings of the 27th International Conference of the System Dynamics Society.

Sargent, R. G. (2004). *Validation and verification of simulation models*. In Proceedings of the 2004 Winter Simulation Conference, 2004. IEEE.

Sarkis, J. (2012). Models for compassionate operations. *International Journal of Production Economics, 139*(2), 359–365.

Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research*, *72*(3), 387–431.

Staw, B. M., Sandelands, L. E., & Dutton, J. E. (1981). Threat rigidity effects in organizational behavior: A multilevel analysis. *Administrative Science Quarterly*, 26(4) 501–524.

Usselman, M., Aguilar, J. C., Llewellyn, D. C., & Ryan, M. (2009). *Science Learning: Integrating Design, Engineering, and Robotics (SLIDER)* (NSF Project Number: 0918618).

Weaver-Hightower, M. B. (2008). An ecology metaphor for educational policy analysis: A call to complexity. *Educational Researcher*, *37*(3), 153–167.

Censorship as/and Social Good: Modeling the Publication of Mark Twain's Autobiography

Jeremy Throne[A]

Abstract

This paper takes the publication of Mark Twain's Autobiography as a case study for exploring how information circulates through an audience. I examine how agents familiar with media coverage from the nineteenth century respond to the publication of the Autobiography under different scenarios using an agent-based model and conclude that Twain's desire to withhold the text from his contemporaries but not from later generations suggests a belief that his readers were part of a vibrant, engaged, and highly connected community that would endure well into the future. At the beginning of the model, agents are given topics of conversation based on the output of topic models of nineteenth-century media coverage preserved in the Chronicling America project at the Library of Congress. As the model progresses agents attempt to discuss these topics with their neighbors. Failed conversations lead agents to select new topics of conversation that come from either Twain or the press, depending upon the preference of the agent. I track conversation dynamics in the model in order to determine Twain's influence on the environment. The model provides an exploratory tool for constructing counterfactual literary histories. It also offers a framework within which the spread of information through an audience may be examined more generally. This framework may be of interest to policymakers seeking to evaluate the spread of information through an environment and policy critics seeking to work backward from available information in order to construct a window into the worldview behind a particular decision.

Keywords: agent-based model, information dispersal, conversation dynamics, counterfactual literary history, Mark Twain, Autobiography, nineteenth-century U.S. literatures, nineteenth-century newspapers, topic modeling, chronicling America

[A] University of California, Santa Cruz

doi: 10.18278/jpcs.3.2.6

摘要

本文将马克•吐温（Mark Twain，以下简称吐温）自传的出版作为案例研究，探索信息如何在读者中进行流通。通过一项基于主体的建模（agent-based model，简称ABM），本文检测了熟悉19世纪媒体报道的读者在不同情境下如何回应自传的出版，并从中得出结论，认为吐温不将其自传公布给同龄人，而是让后代人阅读的意图在于其相信后代读者是有活力、专注且高度连接的社区成员，他们能够适应未来。模型首先给读者不同的谈话主题，这些谈话基于国会图书馆美国编年史中有关19世纪媒体报道的各类主题。随后，读者试图和附近的人讨论这些话题。 如果对话不成功，读者则会根据自身的偏好选择有关吐温或出版社的新话题进一步沟通。本文记录了模型所收集的谈话动态，以确定吐温对当时环境的影响。该模型为建构反事实文学史（counterfactual literary histories）提供了探索工具，同时还提供了一项框架（在该框架内，检测通过读者扩散的信息可能会更具有一般性）。框架可能会引起政策制定者和政策批评者的兴趣。前者试图通过环境评价信息扩散，后者试图从现有信息中进行逆向工作，建构窗口，以理解某个决定背后的世界观。

关键词：基于主体建模，信息分散，谈话动态，反事实文学史，马克•吐温，自传，美国19世纪文学，19世纪报纸，话题建模，美国编年史

Resumen

Este artículo toma la publicación de la Autobiografía de Mark Twain como estudio de caso para explorar cómo la información circula a través de una audiencia. Examino cómo agentes familiarizados con la cobertura mediática del siglo diecinueve responden a la publicación de la Autobiografía bajo diferentes supuestos utilizando un modelo basado en agentes y concluyo que el deseo de Twain de ocultar el texto de sus contemporáneos, pero no de generaciones futuras, sugiere una creencia que sus lectores formaban parte de una comunidad dinámica, participativa y altamente conectada que duraría mucho tiempo. Al principio de la modelización, se les da a los agentes temas de conversación basados en los resultados de modelos temáticos de la cobertura mediática en el siglo diecinueve preservados en el proyecto Chronicling America de la Biblioteca del Congreso. En cuanto el modelo progresa, los agentes intentan dialogar estos temas con sus vecinos. Las conversaciones insatisfactorias impulsan a los agentes a seleccionar nuevos temas de conversación que vienen ya sea de Twain, o de la prensa, dependiendo de la preferencia del agente. Monitoreo la dinámica de la conversación en el modelo para determinar la influencia de Twain en el ambiente. El modelo provee una herramienta exploratoria para construir historias literarias ficticias. También ofrece un marco teórico dentro del que la dispersión de información a través de una audiencia se puede examinar más generalmente. Este marco teórico puede interesarle a los formuladores de políticas que buscan evaluar la dispersión de información a través de un ambiente y crítica política para trabajar de forma retrospectiva desde la información disponible para construir una ventana hacia la visión del mundo que está detrás de una decisión particular.

Palabras Clave: modelización basada en agentes, información dispersal, dinámicas de conversación, historia literaria ficticia, Mark Twain, Autobiografía, literaturas estadounidenses del siglo diecinueve, periódicos del siglo diecinueve, modelización temática, Chronicling America

Background and Introduction

Mark Twain worked on an autobiography off and on for a substantial portion of his career. Twain scholars have identified pieces of writing intended for the text beginning as early as 1870 (Twain, 2010a). All of these early drafts were set aside for one reason or another, but late in life the autobiography received a great deal of his attention. "You will never know how much enjoyment you have lost until you get to dictating your autobiography; then you will realize, with a pang, that you might have been doing it all your life if you had only had the luck to think of it," Twain wrote to his friend and editor William Dean Howells on January 16, 1904 (Smith, 2010a).

Twain's interest in dictation came after years of failing to write an autobiography that met his expectations; and, even dictation was not immediately compelling. He experimented with the approach in small doses for several more years before launching into an extensive effort to dictate the text. For a period of three years beginning in January 1906, he met nearly every day with stenographer Josephine Hobby and his official biographer, Albert Bigelow Paine, to reflect on his life and current events. By the time these meetings ended more than 250 dictations had been produced and the autobiography ran in excess of 500,000 words (Smith, 2010b). In its printed form, *Autobiography of Mark Twain: The Complete and Authoritative Edition*, published by UC Press in 2010, contains more than 2000 pages spread over three volumes and consists largely of Twain's dictations.

The appearance of the UC Press edition 100 years after Twain's death in 1910 capitalizes on a convenient marketing angle provided by none other than the author himself. Twenty-five excerpts from the autobiography were published by Twain in the *North American Review* during his lifetime; each contains an announcement that the work will not be published "in book form" until after his death (Smith, 2010c). What may have been a marketing ploy was also framed as a public benefit. Delaying publication, Twain reasoned, would allow him to produce a more honest account of himself. In 1899, Twain mused in an interview "a book that is not to be published for a century gives the writer a freedom which he could secure in no other way. In these conditions you can draw a man without prejudice exactly as you knew him and yet have no fear of hurting his feelings or those of his sons or grandsons" (Smith, 2010d). Six years later Twain took this sentiment to heart, suppressing his dictated autobiography to protect his reputation and benefit future generations:

> "I'd like to see a lot of this stuff in print before I die—but not the bulk of it, on no! I am not desiring to be crucified yet. Howells thinks the Auto will outlive the Innocents Abroad a thousand years, & I know it will. I would like the literary world to see (as Howells says) that the form of this book is one of the most memorable literary inventions of the ages. And so it is. It ranks with the steam engine, the printing press & the electric telegraph. I'm the only person who has ever found out the right way to build an autobiography." (Smith, 2010e)

The company in which Twain places his text suggests that he viewed the Autobiography as a technological advance that would improve the circulation of information in its wake. Whatever we make of the mix of public and private factors driving his decision to withhold this invention from his contemporaries, only the *North American Review* excerpts were released while he was alive. Following his death, a broader selection of excerpts appeared, but a complete version of the text was not attempted for a century. The text was known to scholars and circulated in different forms, but the unexpurgated autobiography remained largely hidden from public view.

In this paper, I treat Twain's Autobiography as a communications technology and test the impact this technology has on the topics of discussion taken up by his contemporaries. Simulating the introduction of Twain's Autobiography draws attention to the effect of his decision to withhold the text from his contemporaries and opens a window into Twain's sense of his place in the world. The model also outlines an approach that may be used to explore the movement of information through an audience and to identify assumptions upon which this movement may depend.

Description of The Model

The model world I am working with is established in NetLogo by asking 1,089 patches to sprout a turtle. Each turtle attempts to communicate with a neighbor once each tick and the results of these interactions are tracked using the variables below:

```
turtles-own [
  topic-of-conversation ;; the topic of conversation
  speaking ;; how often a turtle starts a conversation
  listening ;; how often a turtle listens
  conversationfailed ;; how often a turtle fails to start a conversation
  prior-topic ;; prior topic of conversation
  repeatconvo ;; indicates if turtle is making a second attempt at a conversation with a topic
]
```

Aggregating the turtle-specific information collected above provides an overview of the conversation dynamics unfolding in the environment.

Once the simulation starts, every turtle picks a topic of conversation from a list of available topics. Choices are made in proportion to a probability distribution obtained by topic modeling a selection of historic newspapers found in the Chronicling America Project at the Library of Congress. I used the Machine Learning for LanguagE Toolkit (MALLET) via the Software Environment for the Advancement of Scholarly Research (SEASR) to identify 30 topics and their distributions over the month of January 1906.[1] The SEASR flow I designed finds a minimum of one and a maximum of five active topics in each day of coverage, with an average of 2.28 and a standard deviation of 1.39 active topics per day. The distribution of topics in the corpus is skewed to the right: days with few topics are more common and days with multiple active topics are less common in the data. Figure 1 summarizes the distribution of topics produced using topic modeling.

This approach provides a representation of topics present in the news for each day of the month. For example, January 4 and January 6 are each described by five active topics; while coverage for

[1] For Chronicling America, see http://chroniclingamerica.loc.gov/.

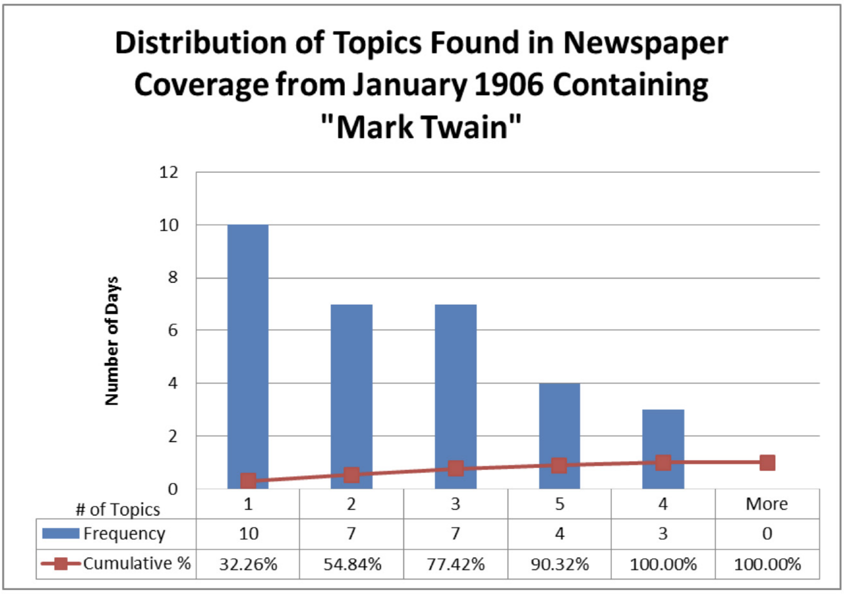


Figure 1. Distribution of topics containing "Mark Twain."

January 3 divides across two topics; and, January 2 is associated with a single topic. Looking more closely at an individual day reveals, for instance, that January 1 divides into three topics: Topic 1 describes 21.5% of the coverage; Topic3 is associated with 78.4% of the coverage; and Topic 8 accounts for <1%. January 2 is described by Topic10; and January 3 coverage is 88% Topic21 and 12% Topic13.

Turtles in my model select topics of conversation in accordance with my topic modeling data. The first cycle of the model coincides with the distribution of topics present on January 1, 1906: 21% of turtles take up an interest in Topic 1; 78% of turtles are ready to discuss Topic 3; and 1% of turtles in the environment are ready to converse about Topic 8. Turtles also adopt a color associated with their particular topic of interest so that the distribution of topics through the space may be noted by looking at an overview of the model world.

New lists of available topics are introduced into the environment every tick. The rate at which turtles take up these new topics is controlled by a percent-active-readers slider, which determines the number of turtles in the environment that draw their topic of conversation directly from the list of active conversations available at every tick of the model. Turtles identified by the percent-active-readers slider may be thought of as people that take an active interest in and keep up with news coverage every day. At 100%, every turtle in the model draws its topic of conversation from the current news coverage made available with each tick of the model. At 0%, turtles take up current news topics only when they are unsuccessful in an attempt to discuss their current topic of conversation.

For SEASR, see https://ischool.illinois.edu/research/projects/software-environment-advancement-scholarly-research-seasr and for MALLET, see http://mallet.cs.umass.edu/

Results

The 1,089 turtles attempting to converse over 31 ticks of the model produces 33,759 opportunities for discussion. When no active readers are present in the environment the total number of discussions that take place over the life of the model regularly approaches 18,000 (a success rate near 50%). Figure 2 shows the distribution of successful and failed conversations over the month of January for Random Seed 44 when no turtles take an active interest in discussing current topics from the news.

Tracking the number of conversations initiated by each agent during this run shows 54.62% of conversation attempts were successful during this run ("pctyes" in Figure 2). Looking at the performance of individual turtles we see that by the end of the simulation turtles will have between 4 and 30 conversations, with an average of 16.93 (SD of 5.422) and a mode of 18 (Figure 3).

Increasing the percent-active-reader's slider slightly depresses the level of conversation between turtles observed. In Figure 4 we see that a 5% increase in the number of turtles that keep up with current events drives the total number of discussions that occur in over the life of the simulation down to 13,198 (a success rate of 39.09%).

A large increase in the percent-active-readers slider, on the other hand, raises the level of conversation observed between turtles. In Figure 5 (below) we see that a 90% increase in the number of turtles that keep up with current events raises the total number of discussions that occur in over the life of the simulation to 20,388 (a success rate of 60.39%).

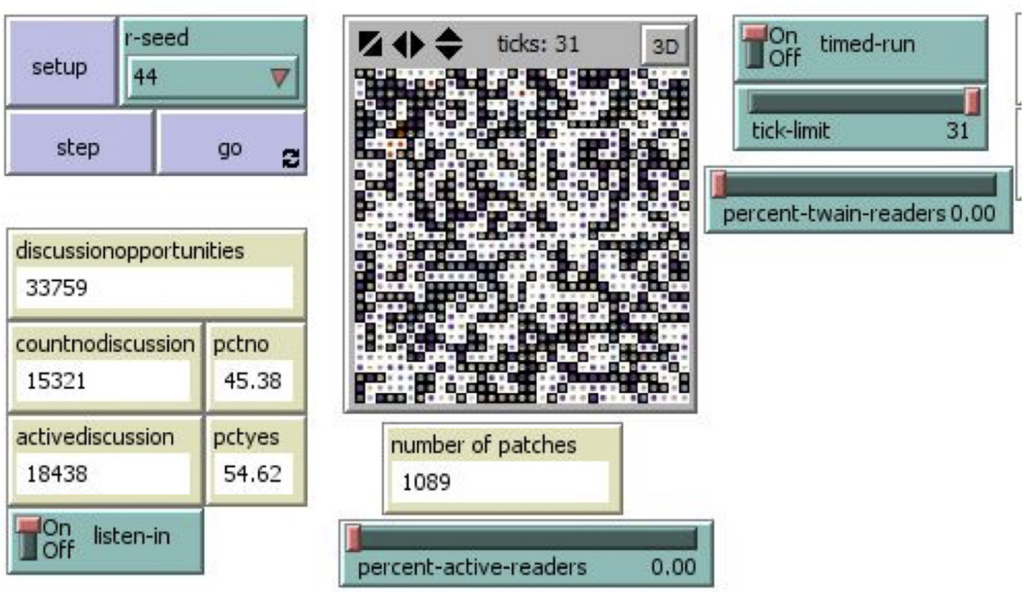


Figure 2. Distribution of successful and failed conversations over the month of January.

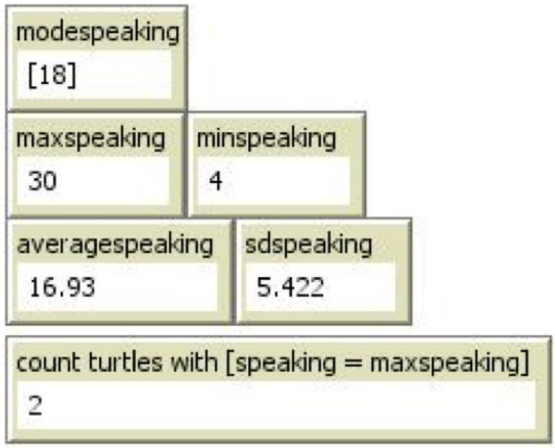


Figure 3. Performance of individual turtles.

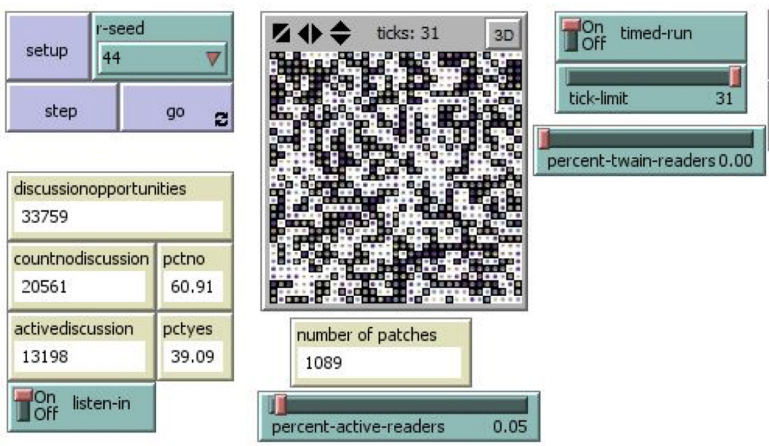


Figure 4. Results of increase percent-active-reader slider slightly.

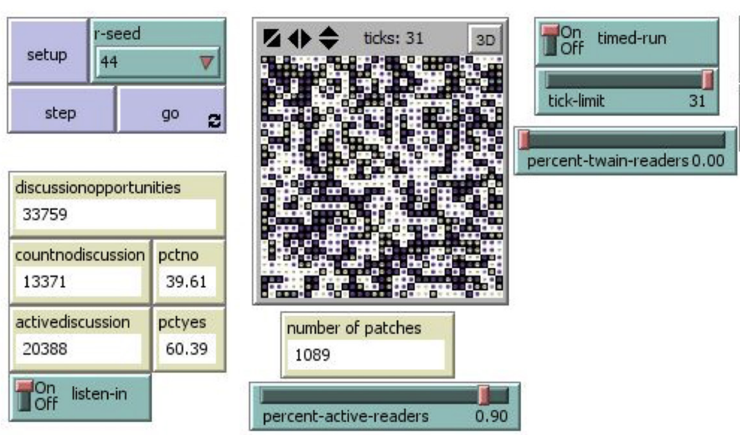


Figure 5. Large increase in percent-active-readers slide results.

Graphing conversation levels over the course of the simulation shows three different patterns of activity associated with the number of active readers in the environment. Conversation levels decrease over the course of the month when turtles do not take an active interest in current events, but more often than not turtles succeed in having a conversation with a neighbor (Figure 6).

When the percentage of active readers in the environment increases slightly, conversation gives way to an inability to communicate (Figure 7).

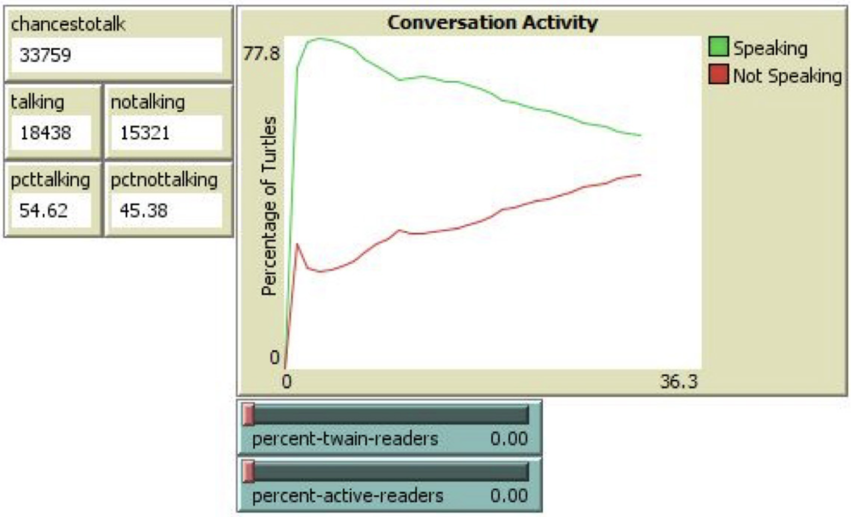


Figure 6. Results of graphing conversation levels.

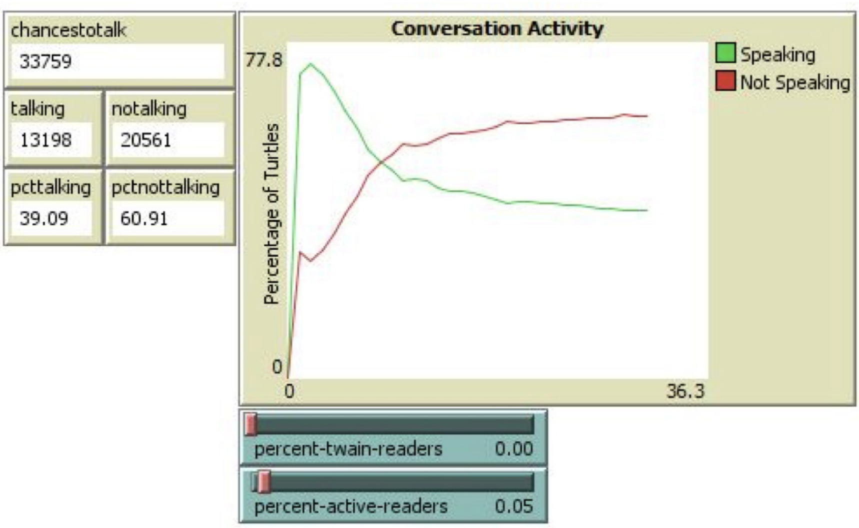


Figure 7. Results of percentage increase slightly in active readers.

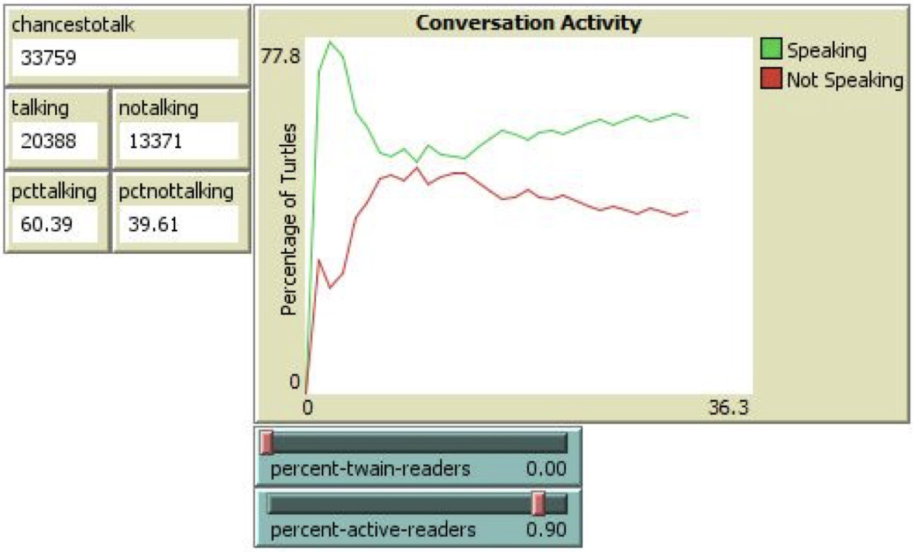


Figure 8. Results when the number of active readers in the environment is substantial.

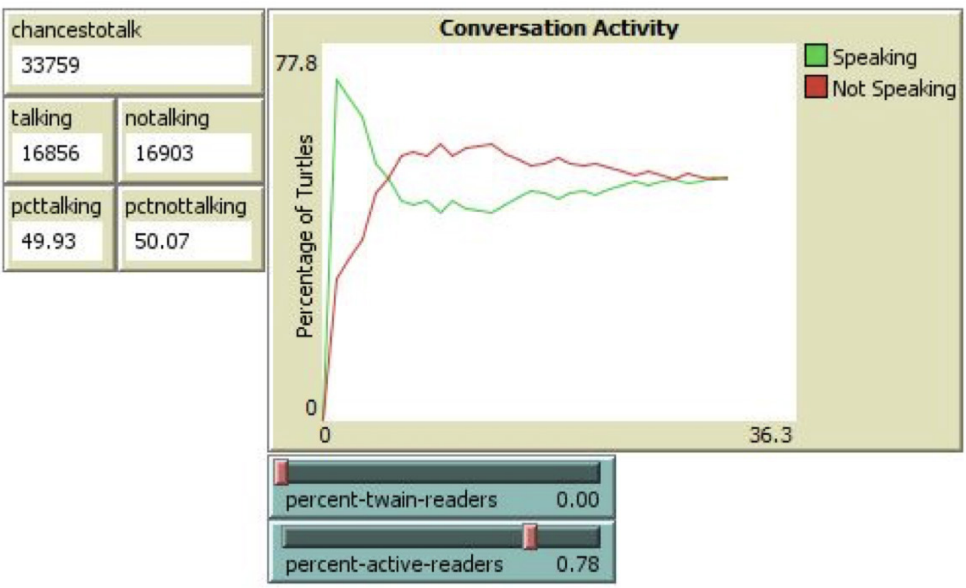


Figure 9. Results indicating conversation levels successfully rebounding.

When the number of active readers in the environment is substantial, conversation levels dip and then rebound over the course of the month (Figure 8).

The point at which conversation levels successfully rebound first occurs when approximately 78% of turtles in the environment take an active interest in keeping up with current events (Figure 9). Above this point, conversation between turtles eventually outpaces an inability to communicate; below this point, the opposite tends to hold true.

Leaving in place the assumption that 78% of turtles in the simulation have an interest in keeping up with current events, I will now turn my attention to seeing how the introduction of an appetite for Twain's text influences conversation patterns in the model.

The percent-twain-readers slider controls the percentage of turtles in the environment that will take up an interest in Twain's dictations when they are available. Unlike the news topics, which potentially change every day, Twain's dictations in January are intermittent. He produced 12 dictations in January: Jan 9–13; Jan 15–19; and Jan 23–24. In the current model, I introduce his commentary into the environment on those days. As is the case with topics drawn from the news, topics drawn from Twain are immediately available to those turtles that are interested in them and are updated every tick. One way to compensate for the intermittent nature of Twain's work may be to give his topics a longer period of activity than topics found in the news, but in the current model both sets of topics are equally transient.

The model shows the number of successful conversations in the environment begins decreasing once even a few turtles become interested in discussing Twain. Figure 10 shows the results obtained when 10% of turtles in the environment turn their attention to his text.

A similar decrease is evident when 78% of turtles in the model prefer to discuss Twain instead of current events (Figure 11).

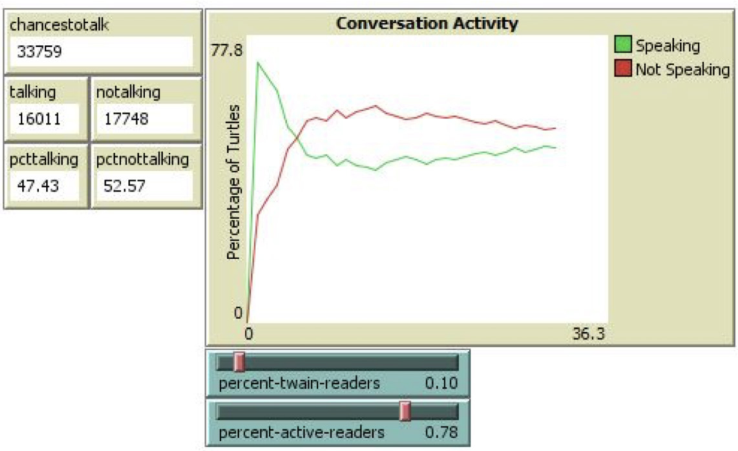


Figure 10. Results of the number of successful conversations decreasing.

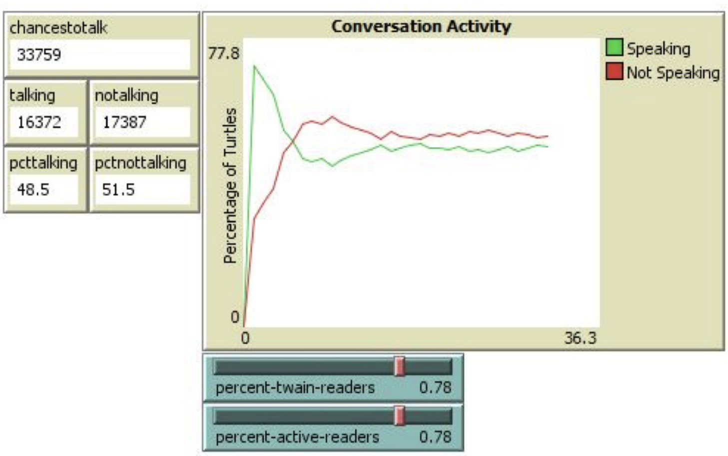


Figure 11. Results when turtles prefer to discuss Twain instead of current events.

Equilibrium between active and failed conversations returns when around 80% of agents in the model take an interest in discussing Twain (Figure 12).

The number of active conversations in the model once again outpaces the number failed attempts at conversation by the time 85% of turtles in the environment prefer Twain to current events (Figure 13).

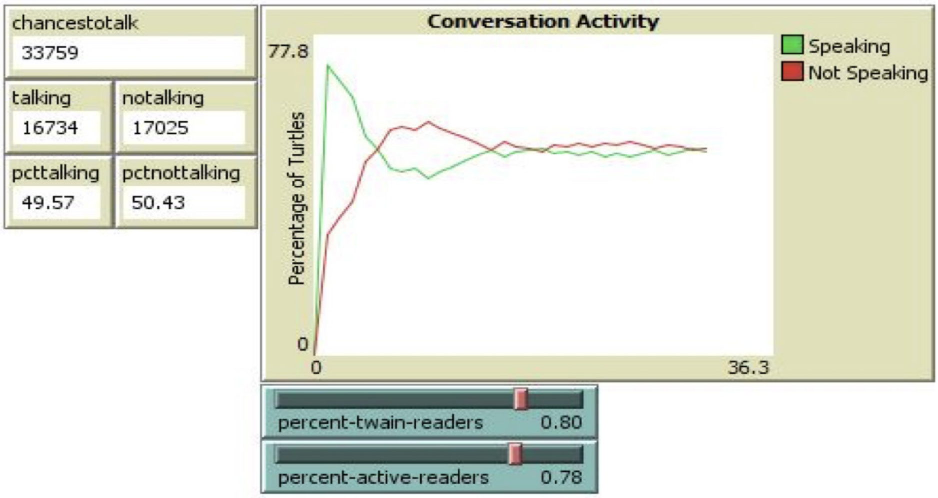


Figure 12. Results of equilibrium between active and failed conversations.

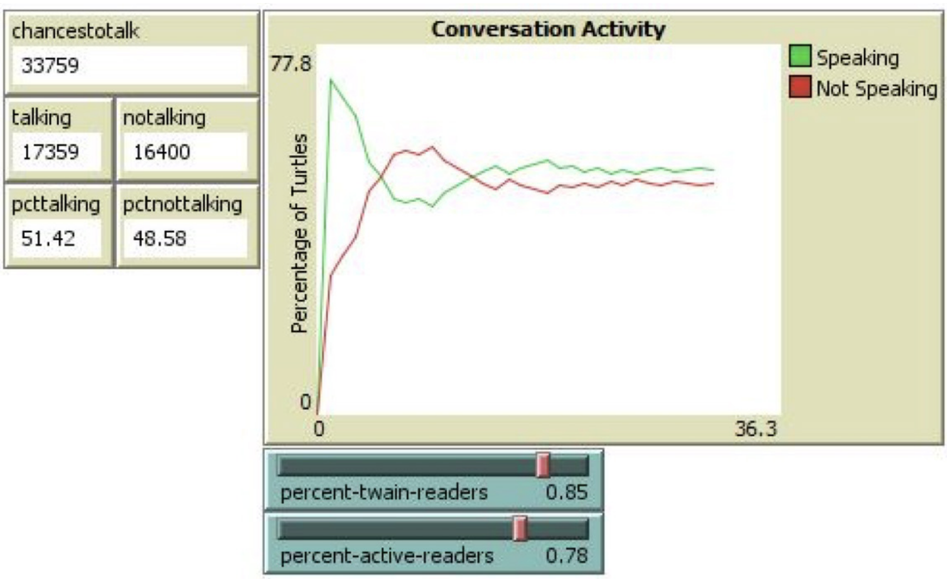


Figure 13. Active versus failed conversations.

Conclusions

Benedict Anderson assigns the press a powerful role in fostering a sense of national identity through the creation of "imagined communities" of readers (Anderson, 1983). My results imply the presence of limits that must be surpassed before these communities can take hold. When less than 78% of turtles in my model take an interest in the news, conversation between agents appears to be stifled more often than not. This result suggests that the nation building effects Anderson attributes to the mass media must truly be operating at scale before they can be said to be a unifying force in the world.

Adding Twain's voice to the model suggests that he may have understood the power of mass communication and his place within that industry very clearly. Withholding the Autobiography enhances the likelihood that turtles in my model will converse until at least 80% of readers can be assumed to take an interest in his work. If we place a premium on communication, conversation, and discussion between agents, Twain's silence can be read as an effort to benefit the public good and an indication that he viewed the high levels of audience attention necessary to enhance the level of public discourse he saw around him as unattainable at the moment. Suppressing the text, in other words, could be said to protect his reputation and benefit his society. Should his fame increase, however, publication of the text also has the potential to benefit society. In my model, publication of the Autobiography enhances levels of conversation in the model once the text captures the attention of 80% of agents in the environment.

Even for an influential writer and international celebrity of Twain's stature, capturing the attention of 80% of an audience seems like an unreasonably high bar to cross, but Twain's goals for the Autobiography were not modest: "I intend that this autobiography shall become a model for all future autobiographies when it is published, after my death, and I also intend that it shall be read and admired a good many centuries because of its form and method—a form and method whereby the past and the present are constantly brought face to face, resulting in contrasts which newly fire up the interest all along like contact of flint with steel" (Twain, 2010b). A writer seeking to influence the shape of "all future autobiographies" would be unlikely to see capturing the attention of 80% of readers as an unattainable goal.

What may first appear to be an improbable goal begins to seem much less ostentatious if we consider the conversation dynamics observed in the model from an agent's perspective. Considering activity in the model from the perspective of an individual agents shows that as agents become more aware of their neighbors, information may reach fewer turtles and still appear to permeate an environment. In my model 80% of turtles in the environment can be exposed to Twain's work when as few as 30% of turtles in the environment actively read Twain's writing when turtles are given the ability to be aware of what agents beyond their immediate neighbors are reading (Figure 14).

As familiarity between turtles and their neighbors grows, the percentage of agents Twain needs to reach to attain 80% awareness continues to fall. As few as 20% of turtles in the environment reading Twain can reach the 80% mark when turtles are given the ability to keep up with the reading habits of all turtles within a 4-unit radius (roughly 48 turtles or 5% of turtles in the environment) (Figure 15).

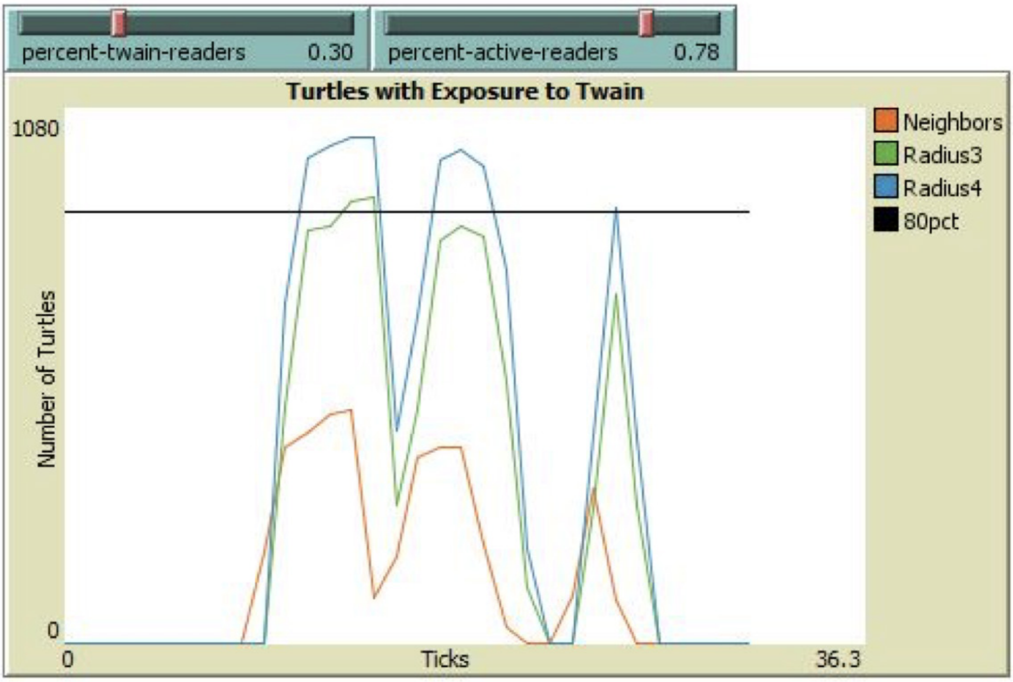


Figure 14. Results of turtles exposed to Twain's work.

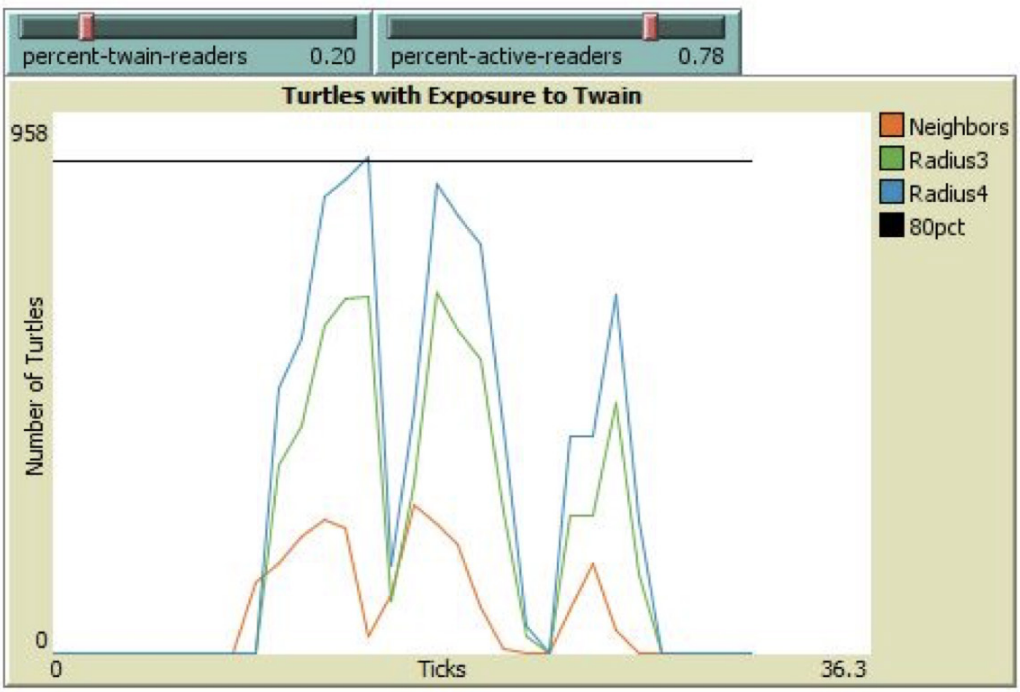


Figure 15. Results that show familiarity between turtles and their neighbors grows.

While these individual data points may be of little value in themselves, together they demonstrate that the more connected turtles are to their contemporaries, the fewer turtles need to be exposed to a topic in order to give the impression that it is a popular point of conversation. Applying this observation to Twain and his audience suggests that Twain's confidence in the popularity of his Autobiography could be attributed to the strength of the connections he sees between audience members. Twain himself was quite well connected to his contemporaries. His typical rate of letter production has been noted at between two and 10 letters a day; and, Robert Hirst notes, "anybody who wrote him tended to get a reply. He easily wrote 50,000 letters" (Griffith, 2001). The more connected Twain may have imagined his readers to be, the easier it becomes to arrive at the conclusion that his work is reaching a large audience.

Looking beyond Twain's specific situation, my model can generally be said to show that high levels of interest in a topic of conversation are more easily imagined as audience members in an environment become more connected. Even a small audience can generate the impression that large numbers of agents are interested in a topic if agents in the model are sufficiently connected. This observation points to a reading of the Autobiography in which Twain's enthusiasm for the text signals a belief in a vibrant, social, engaged society that stands in contrast to the bleak worldview more commonly associated with his later years.[2] It is also draws attention to how closely questions about the circulation of information are linked to assumptions about audience behavior in the nineteenth century, and in our own.

At the moment the United States is led by a President who claims an online audience of 100 million people while leading a nation estimated by the U.S. Census Bureau to have a population of nearly 325 million (Barber, Sevastopulo, & Tett, 2017).[3] The President routinely characterizes his online presence as an alternative to and an escape from the media, stating on Twitter December 5 of 2016 "if the press would cover me accurately & honorably, I would have far less reason to 'tweet.' Sadly, I don't know if that will ever happen!" (Trump, 2016). Many experts estimate his audience is likely to be much smaller than 100 million. Journalist Jeff Nesbit summarizes some of their objections in a recent contribution to U.S. News and World Report before concluding "Trump seems to genuinely believe that he is bypassing the media and speaking directly to large numbers of American voters through his Twitter account. But, by any reasonable measure, he is not. The irony is that his actual megaphone to a large audience may, in fact, be occurring when the news media he despises writes stories about his tweets" (Nesbit, 2017). Given the fact that the President describes his 100 million followers as an aggregate spread across Twitter, Facebook, and Instagram, I see no reason to disagree with Nesbit's criticism. What interests me in this debate are the reasons that may be marshaled to support these two very different estimates of audience size. The model I have constructed to investigate the circulation of Mark Twain's Autobiography suggests that a central point in this discussion and other discussions involving audience size and the circulation of information involves determining the model of audience behavior

[2] See, for example, the biographical portrait provided by the Mark Twain House & Museum at https://www.marktwainhouse.org/man/biography_main.php

[3] For a current estimate of the U.S. population see https://www.census.gov/popclock/.

upon which estimates of audience size are based. Establishing this model, whether it be an assumption that politics are rarely or widely discussed by a particular audience, is a step toward determining the effectiveness of communicating with a mass audience via social media; and, it also opens a window into the worldview of those who choose to communicate using these tools. Agent-based modeling seems well positioned to provide both policymakers and policy critics with an opportunity to address this issue, whether it is done while investigating questions of censorship and circulation in the nineteenth century or our own.

References

Anderson, B. (1983). *Imagined communities*. London: Verso.

Barber, L., Sevastopulo, D., & Tett, G. (2017). Donald Trump: Without Twitter, I would not be here. *Financial Times*. Retrieved from https://www.ft.com/content/943e322a-178a-11e7-9c35-0dd2cb31823a

Griffith, M. (2001). Twain: A man of 50,000 letters. *Los Angeles Times*. Retrieved from http://articles.latimes.com/2001/dec/02/local/me-10527

Nesbit, J. (2017). Trump's Twitter reach is actually small. *U.S. News & World Report*. Retrieved from https://www.usnews.com/news/at-the-edge/articles/2017-04-04/donald-trumps-twitter-reach-isnt-as-big-as-he-thinks-it-is

Smith, H. E. (2010a). [Introduction: Paragraph 77]. In H. E. Smith, B. Griffin, V. Fischer, M. B. Frank, S. K. Goetz, & L. D. Myrick (Eds.), *Autobiography of Mark Twain, Volume 1*. Berkeley, CA: University of California Press. Retrieved from http://www.marktwainproject.org/xtf/view?docId=works/MTDP10362.xml;style=work;brand=mtp;chunk.id=fr0011#d1e1426

Smith, H. E. (2010b). [Introduction: Paragraph 3]. In H. E. Smith, B. Griffin, V. Fischer, M. B. Frank, S. K. Goetz, & L. D. Myrick (Eds.), *Autobiography of Mark Twain, Volume 1*. Berkeley, CA: University of California Press. Retrieved from http://www.marktwainproject.org/xtf/view?docId=works/MTDP10362.xml;style=work;brand=mtp;chunk.id=fr0011#d1e715

Smith, H. E. (2010c). [Introduction: Paragraph 7]. In H. E. Smith, B. Griffin, V. Fischer, M. B. Frank, S. K. Goetz, & L. D. Myrick (Eds.), *Autobiography of Mark Twain, Volume 1*. Berkeley, CA: University of California Press. Retrieved from http://www.marktwainproject.org/xtf/view?docId=works/MTDP10362.xml;style=work;brand=mtp;chunk.id=fr0011#d1e747

Smith, H. E. (2010d). [Introduction: Paragraph 5]. In H. E. Smith, B. Griffin, V. Fischer, M. B. Frank, S. K. Goetz, & L. D. Myrick (Eds.), *Autobiography of Mark Twain, Volume 1*. Berkeley, CA: University of California Press. Retrieved from http://www.marktwainproject.org/xtf/view?docId=works/MTDP10362.xml;style=work;brand=mtp;chunk.id=fr0011#d1e726

Smith, H. E. (2010e). [Introduction: Paragraph 128]. In H. E. Smith, B. Griffin, V. Fischer, M. B. Frank, S. K. Goetz, & L. D. Myrick (Eds.), *Autobiography of Mark Twain, Volume 1*. Berkeley, CA: University of California Press. Retrieved from http://www.marktwainproject.org/xtf/view?docId=works/MTDP10362.xml;style=work;b rand=mtp;chunk.id=fr0011#d1e1761

Trump, D. [realDonaldTrump]. (2016, December 5). If the press would cover me accurately & honorably, I would have far less reason to "tweet." Sadly, I don't know if that will ever happen! [Tweet]. Retrieved from https://twitter.com/realdonaldtrump/status/805804034309427200

Twain, M. (2010a). The Tennessee Land. In H. E. Smith, B. Griffin, V. Fischer, M. B. Frank, S. K. Goetz, & L. D. Myrick (Eds.), *Autobiography of Mark Twain, Volume 1*. Berkeley, CA: University of California Press. Retrieved from http://www.marktwainproject.org/xtf/view?docId=works/MTDP10362.xml;style=work;b rand=mtp;chunk.id=d132371

Twain, M. (2010b). 26 March 1906 [Paragraph 27]. In H. E. Smith, B. Griffin, V. Fischer, M. B. Frank, S. K. Goetz, & L. D. Myrick (Eds.), *Autobiography of Mark Twain, Volume 1*. Berkeley, CA: University of California Press. Retrieved from http://www.marktwainproject.org/xtf/view?docId=works/MTDP10362.xml;style=work;b rand=mtp;chunk.id=d1e13987#d1e14054

Modeling Complexity in Human Built Systems: New Approaches, New Findings in Foster Care

Fred Wulczyn[A], John Halloran[B]

Abstract

Foster care is a widely utilized intervention to provide care to children who cannot live with their parents. The substantial portion of contemporary academic work on foster care either uses individual-level analyses, or focuses on linear analysis of aggregate annual statistics. This paper argues that a complexity-based perspective will deepen our understanding of how foster care policy operates. Wulczyn (1996) argued that a class of population growth models focused on latent resource limitations held great potential to elucidate behavior within aggregate foster care dynamics. Specifically, partial adjustment models propose that population growth rates and limits are dependent on prior resource and capacity states within a system. Recent advancements in data availability, methodology, and computation have made that theory testable. Analyzing data from the Foster Care Data Archive representing 81,142 child entries into out-of-home placements from 2000 through 2014, we apply Empirical Dynamic Modeling to identify nonlinearity and detect causal relationships in coupled time series (Sugihara et al., 2012). Findings indicate nonlinear casual relationships in the coupled foster care entry and exit time series data. These findings match expectations of strength and directionality. The empirical findings and the models that they support have the potential to frame child welfare systems differently, ultimately leading to different actionable policy conclusions.

Keywords: child maltreatment, child protection, human-built systems, systems theory

[A] Chapin Hall Center for Children at the University of Chicago

[B] Lewis University

doi: 10.18278/jpcs.3.2.7

摘要

对无法和父母一起生活的儿童而言，寄养是广泛使用的干预措施。寄养的当代学术工作实质上不是使用了个人层面分析，就是使用了年度统计集合（aggregate annual statistics）线性分析。 本文认为，基于复杂性的视角（complexity-based perspective）能加深对寄养政策如何实行的理解。 Wulczyn (1996)认为，聚焦于潜在资源限制的人口增长模型具备很大的潜力，能够解释总体寄养动态中的行为。具体而言，局部调整模型（partial adjustment models）认为人口增长率和人口增长限制都取决于之前的资源和系统内的容量状态。这一理论观点的检验需要用到近期在数据可用性、方法论和计算上取得的进展。寄养数据档案（Foster Care Data Archive）中包含了2000-2014年间81,142名儿童的寄养情况。为分析档案数据，本文应用了经验动态建模（Empirical Dynamic Modeling），以识别非线性，同时侦测耦合时间序列（coupled time series）的因果关系(Sugihara et al., 2012)。发现指出，非线性因果关系存在于寄养开始和寄养结束这两个时间序列之间。该发现符合有关寄养能力和方向性的期望。实证发现和其所支持的模型具备将各儿童福利系统进行区别对待的潜力，并最终促成不同的可执行政策结论。

关键词：儿童虐待，儿童保护，人造系统，系统理论

Resumen

El acogimiento de menores es una intervención generalmente utilizada para proveer cuidado a niños que no pueden vivir con sus padres. Una proporción sustancial del trabajo académico contemporáneo en el tema del acogimiento de menores usa, ya sea el análisis a nivel individual, o se enfoca en el análisis lineal de las estadísticas globales anuales. Este artículo sostiene que una perspectiva basada en la complejidad profundizará nuestro entendimiento de cómo el acogimiento de menores opera. Wulczyn (1996) argumentó que una clase de modelos de crecimiento poblacional que se enfocó en limitaciones de recursos latentes tuvo un gran potencial para dilucidar el comportamiento dentro de las dinámicas totales del acogimiento de menores. Específicamente, los ajustes parciales a los modelos sugieren que el grado de crecimiento poblacional y sus límites dependen de estados previos de recursos y capacidad dentro de un sistema. Avances recientes en disponibilidad de datos, metodología y computación han logrado que esa teoría se pueda verificar. Cuando analizamos datos del Foster Care Data Archive (Archivo de datos del acogimiento de menores) que representan 81.142 registros infantiles a colocaciones fuera del hogar de 2000 a 2014, aplicamos la modelización empírica dinámica para identificar la no linealidad y detectar relaciones en series cronológicas emparejadas (Sugihara et al., 2012). Los hallazgos indican que hay relaciones casuales no lineales en los registros emparejados de acogimiento de menores y las series de datos de salida. Estos hallazgos coinciden con las expectativas de fuerza y direccionalidad. Los hallazgos empíricos y los modelos que estos mismos apoyan tienen el potencial de enmarcar diferentemente a los sistemas de bienestar infantil, lo que resultaría en diferentes conclusiones de política accionable.

Palabras clave: maltrato infantil; protección infantil; sistemas construidos por humanos; teoría de sistemas.

Introduction

Foster care is an alternative living arrangement for children who are unable to live with their parents due, typically, to abuse or neglect. Foster care systems are found in every state in the United States and virtually every country around the world, although their size and complexity differ depending on context.

As an intervention foster care carries its own long-term consequences for children and families (Doyle, 2008, 2013; Lawrence, Carlson, & Egeland, 2006; Leve, Fisher, & Chamberlain, 2009). This risk of iatrogenic effects of foster care has led researchers to focus on the reasons why children are placed out of the home (Barth, Wildfire, & Green, 2006; Briar, 2008; Horwitz, Hurlburt, Cohen, Zhang, & Landsverk, 2011). This literature has been framed around two separate, but related, lines of inquiry. First, studies which parse the individual child or family risk factors for placement in foster care. Second, studies concentrating on the decision-making of child welfare professionals (Dettlaff, Graham, Holzman, Baumann, & Fluke, 2015). This research has been critical in producing interventions that prevent maltreatment, divert children from the child welfare system, and tools that assess risk and safety (Baird & Wagner, 2000).

These lines of research carry a tacit assumption: placements of children in out-of-home care are being determined by fact-specific analysis conducted of individual decision makers of individual cases. In this paradigm, the reasons for placement are viewed almost entirely through the lens of the single case. However, as we argue below, the dynamics of placement decisions almost certainly extend beyond the limits of the single case into the child welfare system as a whole (Bakwin, 1945; Emerson, 1982). Single cases do not occur in a vacuum and attention to broader system factors is likely needed before we fully understand how the foster care system operates and why children are placed into care (Wulczyn et al., 2010).

When we speak about system factors or system effects we are referring to classes of observable system features including nonlinearities, feedback loops, and contingencies that occur in the implementation of child welfare policy (Jervis, 1997; Lansing, 2003). These system effects are not visible in the analysis of a single case. Examples of potential system factors in child welfare might include relative risk where the risk level of a single case is assessed in the broader context of cases being evaluated at roughly the same time (Emerson, 1982, 1991), resource constraints such as total time allocated to decision-making as represented by the number of workers (Holmes et al., 2014), or the number of beds available relative to demand (Wulczyn, 1996).

We argue that applying systems analysis to foster care holds vast potential to yield insights that policy makers can use to deliver higher quality care. As metaphor, a complex systems perspective aligns well with foster care—there are well-defined structures, processes, resources, capacities, and functions linked to a set of goals. Once a state has made a policy decision that certain conditions require the removal of children from the home of their caregivers there is created a two-pronged policy problem: (1) where such children should be placed and (2) how to manage those placements. Out-of-home care systems are state-managed (from the selection of potential placement sites, to the actual process of placement) and are almost entirely state-financed. The empirical thrust of this paper, then, is establishing that systems science fits the subject matter in terms that are more concrete. To do this we explore developing opportunities in system science to deepen our understanding of human built systems though the unique fine-grained, reliable data that are generated as children move through the foster care system.

Foster care is a resource-constrained system and these resource constraints exist differently within different subsets of foster-care placements. More specifically the management of placements is necessarily different in traditional foster care than it is in congregate care.

Traditional foster care is a placement in the home of a state-licensed substitute caregiver who is often unknown to the child at the time of the placement. The capacity of traditional foster homes represents a resource stock that is managed by the state, requiring active recruitment, licensing, and resource management. Thus, the stock of traditional foster homes is more flexible than those beds available in congregate care. *Congregate care*, or group care, is placement in a group facility managed by the state or by a private entity contracted by the state. Though congregate care can take several different forms (e.g., an in-patient mental health center), a commonality is that these brick-and-mortar facilities are severely constrained in their ability to grow and shrink the available bed capacity. Thus, in comparison to traditional foster care, congregate care beds are the most inflexible, and, purely in terms of beds, the congregate care system is more resource constrained. If the resource constraints exist and impact flows then some evidence of causal coupling should be found in foster care time series analyzed through population growth models.

This resource-based analysis is described by Wulczyn (1996), who proposed a methodological framework for analyzing the aggregate level dynamics of foster care populations that sought to define problems related to system change in terms of models capable of directly answering change-based questions. Specifically,

$$dx(t)/dt = r[x^{\times}(t) - x(t-1)] \qquad (1)$$

referred to as the linear partial adjustment model, defines the change in population x between times $t-1$ and t as a variable relationship between the population of x at time t, the rate of population growth r, and "the level toward which causal forces are impelling [x]," defined as the $x^*(t)$ term (Tuma & Hannan, 1984, p. 440). This model starts with the concept that a system's state at time $t-1$ impacts that system at time t in a predictable way, and elaborate that to suggest that a portion of that impact and predictability is contained within a target value, $x^{\times}(t)$. Wulczyn (1996) hypothesizes that resource utilization serves as target value that both aggregate exits and entries dynamically adjust in accordance to. Recent advances in data availability, computation, and methodology provide an opportunity to test these hypothesized relationships.

Data

The data for this paper were drawn from the Multistate Foster Care Data Archive (FCDA), a longitudinal data warehouse maintained by the Center for State Child Welfare Data and Chapin Hall at the University of Chicago. The FCDA contains fine-grained data for individual children and their experiences within the child welfare system, including details about out-of-home care spells. A care spell represents the amount of time a child spends in out-of-home care, and is marked by a start date, or entry, and an end date, or exit. The entry/exit structure of this data set lends itself to the ecological metaphor where birth/death dynamics examine the growth and decline of biological populations (Wulczyn, 1996). In both ecology and the study of human-built systems, limited available data have influenced the ability to test dynamic models. The advantage of using the foster care population to study human built systems has to do with the quality of the data used to track the population. States are required to record when a child enters foster care and when a child leaves foster care. The time stamps used to track these comings and goings are recorded at the daily level, which means the time scale of the time series data can be adjusted easily and with very little measurement error. Importantly, a large number of state systems have been tracking the whereabouts of foster children for decades. Consequently, we have sufficient time points to study the dynamical properties of the time series data over extended periods. Due to the record keeping requirements of the out-of-home care system, the volume and granularity of the data provides a unique opportunity for theoretical and empirical advances in systems science.

The subset of FCDA data used in these analyses represents 81,142 child entries into out-of-home placements in the State of Tennessee from 2000 through 2014. These data are census data, representing all out-of-home care entries in the population during that period. These entry data are provided at the daily level, representing 5,479 days worth of entries during that time. For analysis, that data was summed at the weekly level, and provided 783 weeks of aggregated entries. The anonymized, aggregate state-level entry data were reported in two different care-type categories: congregate care (Entries, *Figure 1a*; Exits, *Figure 1b*); and traditional foster care (Entries, *Figure 2a*; Exits, *Figure 2b*).

Theory and Method

The analysis in this paper is rooted in empirical dynamical modeling (EDM), a method for time-series analysis developed by ecologist George Sugihara (Sugihara et al., 2012). EDM uses coupled population growth models from the same family as the coupled form of partial adjustment equations proposed in Wulczyn (1996). For non-chaotic systems that can be described as a point in high-dimensional space, EDM provides information on appropriate embedding dimensions through simplex projection, insights on nonlinearity in the system though S-mapping, and analysis of causal relationships between variables through convergent cross mapping (CCM).

Lag Determination

A threshold issue is selecting the appropriate lag to describe the temporal relationship between variables. The lag structure represents the flow of feedback information through the system from stimulus to response. Though out-of-home care time series do display seasonality and periodicity, there is no theoretical ground for the selection of lag length for a fine-grained aggregate out-of-home care time series. Thus, lag structure must be determined empirically. The lag structure of the data was produced using the general-to-specific methodology as described by Enders (2010), where a stepwise longitudinal linear regression is used to narrow down potential lag lengths and then the multivariate model is paired down using criterion testing (in this case, the Akaike Information Criterion). In the present analysis, we were seeking a univariate lag model for application in later analysis.

Simplex Projection and S-Mapping

As can be observed in *Figures 1* and *2*, these out-of-home care time series are highly variable from week to week and may appear to be noisy or chaotic. As a first step in EDM, we use simplex projection (Sugihara & May, 1990) to isolate whether the variation in the time series is indeed chaotic, or whether, when the points in the time series are projected into a higher embedding dimension patterns of relationships between near neighbor points emerge as the attractor is folded in phase space. As the time series is projected into higher embedding dimensions, higher coefficients (representing more predictable near neighbors) indicate that better predictions can be made about the shape of the attractor manifold. Thus, simplex projection assists in assessing not only whether the time series displays nonlinearities, but also whether the variability of a time series is chaotic or predictable.

Further refining the analysis of whether an observed time series displays predictable nonlinear behavior, S-mapping (Sugihara, 1994) uses a nonlinear tuning parameter to evaluate the nonlinearity of a time series in a given embedding dimension. The test predicts nonlinearity where values of a weighted tuning parameter, theta, in nonlinear space exceed those values in linear space. Specifically, theta amplifies the difference between observed distance and average distance between nearest neighbors. In a linear system, then, a theta of zero gives the best predictions. In a nonlinear system, however, increasing the weight (driven by an increase of theta) increases the locality of the prediction in phase space, indicating that nearer neighbors are more predictive than global average distances. Thus, an S-map indicates nonlinear dynamics in a time

series for a given embedding dimension determined by simplex projection.

Convergent Cross Mapping

Sugihara et al. (2012) describes CCM as a method to detect casual relationships between weakly coupled variables in dynamic ecological populations. Theoretical ecology provides an opportunity for the study of populations because of the rich theoretical traditions coupled with mathematical analysis designed to leverage often-incomplete highly variable data sets.

CCM is rooted in Generalized Takens' Theorem, which states that an attractor manifold can be reconstructed from lags of different time series so long as those time series are part of the same system. Leveraging this, CCM predicts whether variables share a causal relationship governed by an attractor manifold by cross-predicting variables that are observed from the same dynamic system (Ye, Deyle, Gilarranz, & Sugihara, 2015). In this case, utilizing Tennessee out-of-home care entries and lagged entries provide some assurance that the variables are coming from the same dynamic system. CCM, then, provides an analytic window into whether a predictable nonlinear attractor manifold exists within that system. Grounded in a latent, unobserved manifold that could be produced form a set of coupled difference equations representing a population growth mode similar to Equation *1*[1]:

[1]Comparing Equation (2) to Equation (1), we see that Equation (2) elaborates the definition of the x×(t) term in Equation (1), which includes not only x from earlier in time but also the influence of a y term. In Equation (2), the entry and exit populations are coupled within the expansion of the x×(t) term from Equation (1). Though these models are not directly interchangeable, they represent the same general class of population growth model.

$$X(t+1) = X(t)[r_x - rxX(t) - \beta_{x,y}(t)] \qquad (2)$$
$$X(t+1) = \mathrm{Y}(t)[\mathrm{r}y - \mathrm{r}yX(t) - \beta_{y,x}(t)]$$

CCM dispenses with the need to resolve that model by assuming that if the predictor variables are coupled and share an attractor manifold, that variable values should be predictably nearby on that manifold structure.

CCM tests for the degree to which entries are functionally coupled with lagged entries. This is accomplished by building an increasingly large library of randomly drawn lagged pair series—starting with a handful of pairs—a short library length—and building to a large number of pairs—a large library length. In the present analysis, library lengths increased from 10 pairs to 400 pairs, at 10-pair steps. Mean correlations are taken across these library lengths, across a number of bootstrapped random draws. In the present analysis, each library length was drawn 100 times. If the variables are coupled around an attractor manifold then the cross-mapped estimates are predicted to improve as the library length increases.

An ad hoc hypothesis test is performed to determine the significance of the CCM findings. This test uses random distribution of the time series data in two forms: first, overall random shuffles of the data set, and second, a semi-annual seasonal shuffle of the data set. These random distributions are analyzed using the CCM method described above, and then compared to the observed distributions to see the percentage of matched library pairs where the CCM coefficient of the random data set exceeds that of the observed data set. This value approximates a p-value.

Results and Analysis

Lag Determination

A pattern emerged using the general-to-specific lag determination methodology that a lag value of one was either the most explanatory, or among the two or three most explanatory, lag values both in terms of overall effect size, r-squared value, and information loss from multivariate lag models to a univariate lag model. Thus, a lag of one was selected for all analyses in this data set. Table 1 describes the coefficient, confidence interval, and r-squared for the series of univariate regression analyses in this study.

Table 1. Lag determination coefficients (Lag=1)

Outcome	Predictor	Beta (95% confidence)	R-squared
CC entries	CC entries	0.70 (0.65–0.75)	0.4951
CC entries	CC exits	0.55 (0.50–0.59)	0.4029
CC exits	CC exits	0.72 (0.68–0.77)	0.5253
CC exits	CC entries	0.80 (0.74–0.86)	0.4772
FC entries	FC entries	0.39 (0.32–0.45)	0.1483
FC entries	FC exits	0.00 (−0.05–0.06)	0.0001
FC exits	FC exits	0.27 (0.20–0.34)	0.0743
FC exits	FC entries	0.44 (0.35–0.52)	0.1112

Table 2. Comparison of convergent cross mapping coefficients

Outcome	Predictor	Regression beta	Correlation coefficient	CCM coefficient
CC entries	CC exits	0.55	0.63	0.80
CC exits	CC entries	0.80	0.69	0.77
FC entries	FC exits	0.00	0.01	0.38
FC exits	FC entries	0.44	0.33	0.49

Simplex Projection and S-Mapping

Simplex projection suggested embedding dimensions of E=7 for the foster series, and E=6 for the congregate care. Figure 3 contains a representative Simplex plot for congregate care entries. Each time series displayed some degree of non-linearity, though each series also contained some linear component. This suggests that both linear and nonlinear analysis techniques have information to contribute to the analysis of out-of-home care time series. Figure 4 contains a representative S-map plot for congregate care entries.

Convergent Cross Mapping

The cross map for each time series indicated some convergence of couple correlations accompanying increases in library size. This is indicative of causal coupled relationships between the two variables represented in each line. The strongest relationship was shown, as expected, in the congregate care time series (see Figure 5). The foster care time series showed the weakest relationship (see Figure 6). However, the foster care time series did show the largest marginal increase in predictive value between the linear models and the CCM model in the lagged foster care exits causing foster care entries series. These values are listed in Table 2.

The ad hoc significance test showed p-values exceeding $p<0.001$ for all cross-mapped data sets in congregate and foster care time series.

Analysis, Implications, and Limitations

Out-of-home care time series display predictable nonlinear causal dynamics at higher embedding dimensions. The complex shapes of the linear time series for congregate care and foster care are not generated through random linear processes, but through the contribution of nonlinear coupling. Moreover, these dynamics can be explored using only one lagged predictive variable. The short lag time (tau) of 1 week employed in this study suggests that out-of-home care time series are proximately sensitive to fluctuations in population sizes. This insight has implications for policy conversations that are frequently carried out on time scales that are quite lengthy in comparison. Understanding that there are nonlinear relationships continued within these time series suggests that some variability attributed to error in linear models may be informative. This finding has implications across the field for the strength of prevailing linear models for child welfare time series.

Overall, the CCM findings indicate that entries into out-of-home care cause exits, and that exits from out-of-home care cause entries. This suggests the existence of a shared manifold structure indicating the presence of an attractor in these coupled time series. The strength of these relationships differs across care types. These findings suggest that there may be some structural component to the out-of-home care system that creates the need of a certain equilibrium between entries and exits. These attractor structures represent resource constraints within the foster care system, and suggest that the dynamics of out-of-home care populations are driven, in part, by system feedback loops produced by the management of these resources.

Findings further indicate that the conceptualizations of population problems in child welfare systems are suitable for the application of population ecology theory and methods in some cases. Specifically those entry/exit regimes in child welfare do display similar behaviors to birth/death regimes in the study of biological populations. Thus, out-of-home care time series do contain variation that can be explained by population biology models. This opens up a robust literature where system behavior has been studied with incomplete information for centuries. Analogizing ways in which ecologists have solved these problems could create a leap in child welfare methodology.

The primarily limitation of this study is that the method provides no insights into what factors might comprise the attractor around which the time series manifold orbits. The presence of nonlinear causal relationships is of interest in and of themselves, but further study is necessary to understand what factors might influence those relationships and how those factors might operate in nonlinear space distinctly from how they operate in linear space. Additionally, this study only reviews state-level aggregate time series from a single state. Further investigations will be required to see if similar nonlinear behavior is found in the out-of-home care systems of other states, and at different levels of spatial aggregation.

Conclusion

Analytic methods focused on different geometries produce different stories. Andersen (1980) recognized that different analytic paradigms applied to the same data set could lead to different policy conclusions. We argue that this is due, in part, to the policy narratives that we attach to the geometries of the data. The results of this study offer alternative policy narratives, which can grow out of nonlinear causal models.

At the most basic level, the prediction of resource constraints acting as a causal influence contributing to the dynamics of the out-of-home care system as portrayed in entry/exit time series data predicted in Wulczyn (1996) receives strong empirical support in this analysis. As is the case with other systems operating in a rich ecological context, though we do not directly observe the decision-making processes, the entry/exit signal the system gives off reveals a great deal about the nature of causality within the system. This analysis suggests the influence of resource constraints as operate at a fine-grained time scale: Each day, each week, each month, caseworkers are tasked with making admission and discharge decisions. The cumulative impact of those decisions balances out over the period of interest to determine whether the caseload grows, shrinks, or stays the same. The availability of aggregate beds in the moment of aggregate decisions likely influences how the dynamic unfolds—as does the amount of aggregate worker time available to make decisions. These underlying aggregate-level time series dynamics, then, reveal rich, and previously unobservable, information about the operation of the child welfare system.

In the longer run, knowing the linear versus non-linear structure in the time series data will advance how we operate the out-of-home care system. In the presence of non-linearity, efforts to control the system can have counter intuitive effects. Knowing that non-linear behavior is present offers the possibility of more precise policymaking. However, before going down this path we have to understand what the hidden structure within the time series data now.

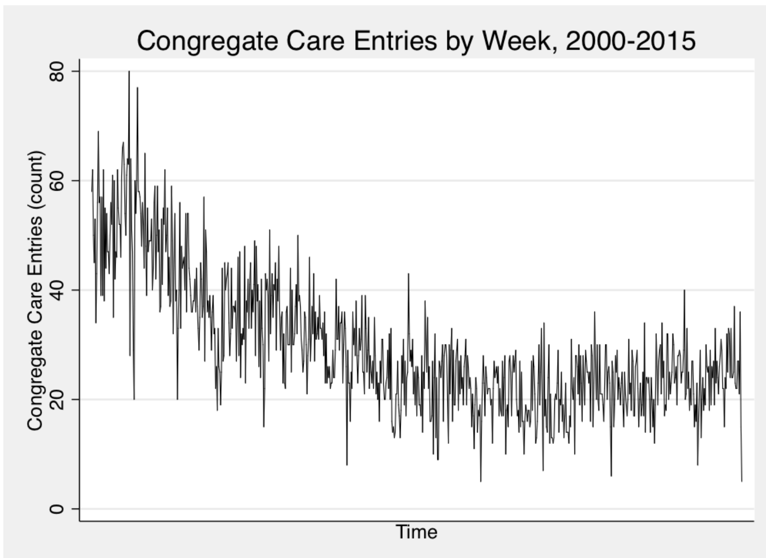


Figure 1a. Total congregate care entries 2000–2015.

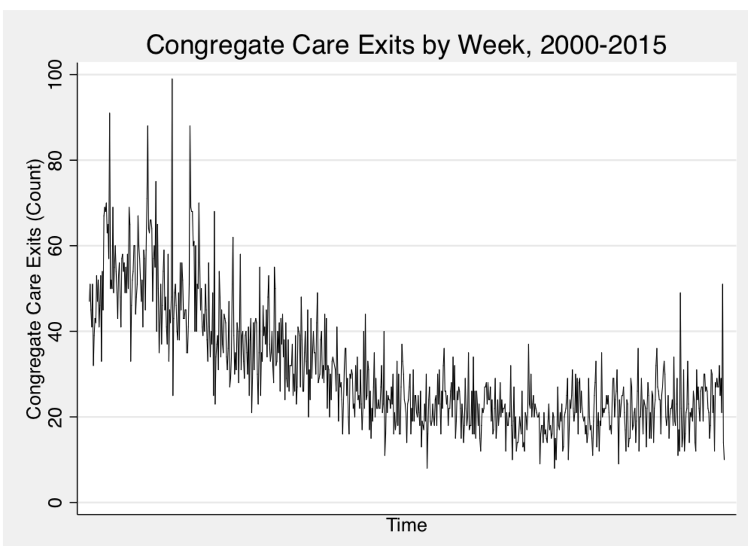


Figure 1b. Total congregate care exits 2000–2015.

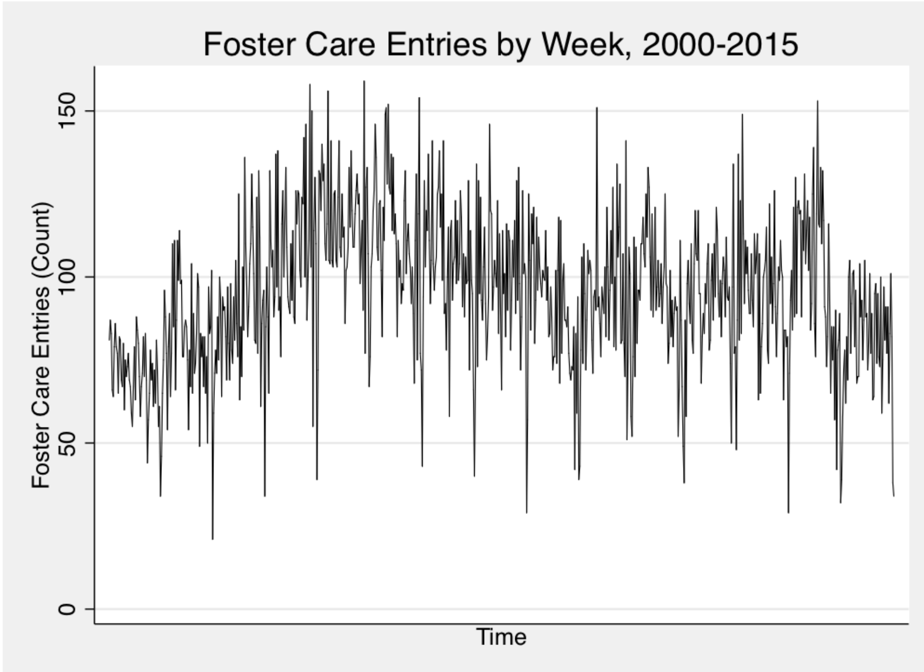


Figure 2a. Total foster care entries by week 2000–2015.

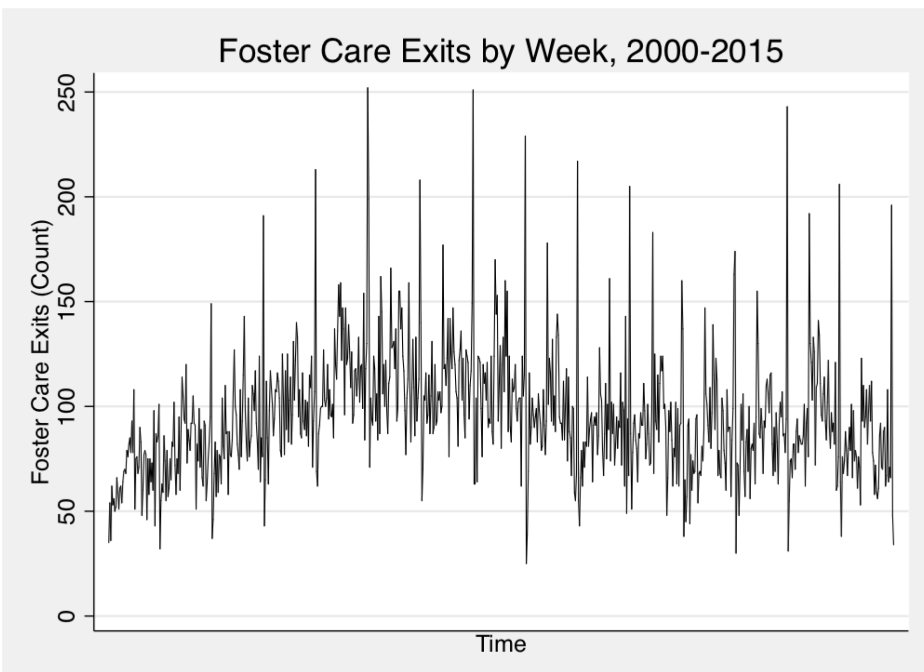


Figure 2b. Total foster care exits 2000–2015.

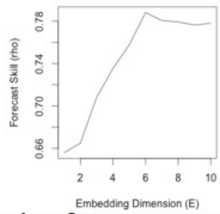


Figure 3. Simplex projection for congregate care entries 2000–2015.

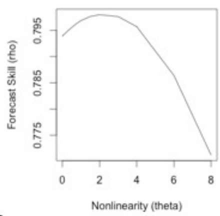


Figure 4. S-map for congregate care entries 2000–2015.

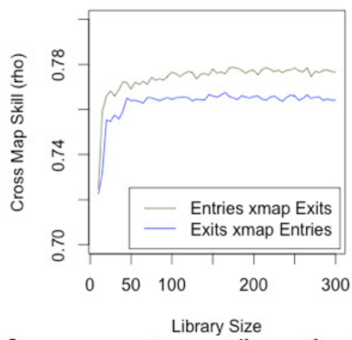


Figure 5. Convergent cross map for congregate care (lagged entries cause exits; lagged exits. cause entries; *E*=6, lag=1)

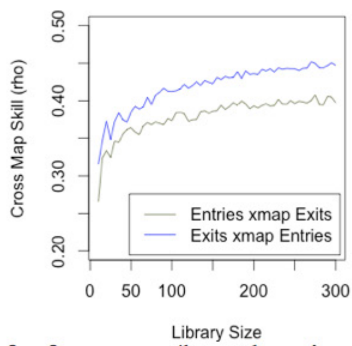


Figure 6. Convergent cross map for foster care (lagged entries cause exits; lagged exits cause entries; *E*=7, lag=1).

References

Andersen, D. (1980). How differences in analytic paradigms can lead to differences in policy conclusions. In J. Randers (Ed.), *Elements of the system dynamics method* (pp. 61–75). Cambridge, MA: MIT Press.

Baird, C., & Wagner, D. (2000). The relative validity of actuarial-and consensus-based risk assessment systems. *Children and Youth Services Review*, 22, 839–871.

Bakwin, H. (1945). Pseudodoxia pediatrica. *New England Journal of Medicine*, 232, 691–697.

Barth, R., Wildfire, J., & Green, R. (2006). Placement into foster care and the interplay of urbanicity, child behavior problems, and poverty. *American Journal of Orthopsychiatry*, 76, 358–366.

Briar, S. (2008). Clinical judgment in foster care placement. R*esidential Group Care & Treatment*, 1, 61–75.

Dettlaff, A. J., Graham, J. C., Holzman, J., Baumann, D. J., & Fluke, J. D. (2015). Development of an instrument to understand the child protective services decision-making process, with a focus on placement decisions. *Child Abuse & Neglect*, 49, 24–34.

Doyle, J. J. (2008). Child protection and adult crime: Using investigator assignment to estimate causal effects of foster care. *The Journal of Political Economy*, 116, 746–770.

Doyle, J. J. (2013). Causal effects of foster care: An instrumental-variables approach. *Children and Youth Services Review*, 35, 1143–1151.

Emerson, R. (1982). Holistic effects in social control decision-making. *Law & Society Review*, 17, 425–456.

Emerson, R. (1991). Case processing and interorganizational knowledge: Detecting the "real reasons" for referrals. *Social Problems*, 38, 198–212.

Enders, W. (2010). *Applied econometric time series* (3rd ed.). Hoboken, NJ: John Wiley & Sons, Ltd.

Holmes, L., Landsverk, J., Ward, H., Rolls-Reutz, J., Saldana, L., Wulczyn, F., & Chamberlain, P. (2014). Cost calculator methods for estimating casework time in child welfare services: A promising approach for use in implementation of evidence-based practices and other service innovations. *Children and Youth Services Review*, 39, 169–176.

Horwitz, S. M., Hurlburt, M. S., Cohen, S. D., Zhang, J., & Landsverk, J. (2011). Predictors of placement for children who initially remained in their homes after an investigation for abuse or neglect. *Child Abuse & Neglect*, 35, 188–198.

Jervis, R. (1997). Complexity and the analysis of political and social life. Political Science Quarterly, 12, 569–593.

Lansing, J. S. (2003). Complex adaptive systems. *Annual Review of Anthropology*, 32, 183–204.

Lawrence, C. R., Carlson, E. A., & Egeland, B. (2006). The impact of foster care on development. *Development and Psychopathology*, 18, 57–76.

Leve, L. D., Fisher, P. A., & Chamberlain, P. (2009). Multidimensional treatment foster care as a preventive intervention to promote resiliency among youth in the child welfare system. *Journal of Personality*, 77, 1869–1902.

Sugihara, G. (1994). Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature*, 334, 734–741.

Sugihara, G., & May, R. (1990). Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature*, 344, 477–495.

Sugihara, G., May, R., Hao, Y., Hsieh, C., Deyle, E., Forgarty, M., & Munch, S. (2012). Detecting causality in complex ecosystems. *Science*, 338(6106), 496–500.
Tuma, N., & Hannan, M. (1984). Social dynamics: Models and methods. New York, NY: Academic Press.

Wulczyn, F. H. (1996). A statistical and methodological framework for analyzing the foster care experiences of children. *Social Service Review*, 70(2), 318–329.
Wulczyn, F. H., Daro, D., Fluke, J., Feldman, S., Glodek, C., & Lifanda, K. (2010). Adapting a systems approach to child protection. New York, NY: UNICEF.

Ye, H., Deyle, E., Gilarranz, L., & Sugihara, G. (2015). Distinguishing time-delayed causal interactions using convergent cross mapping. *Nature Scientific Reports*, 5, 14750.

A Novel Evolutionary Algorithm for Mining Noisy Survey Datasets with an Application Toward Combating Chagas Disease

John P. Hanley, Donna M. Rizzo[A]

Abstract

Chagas disease is a deadly, neglected tropical disease endemic to every country in Central and South America. The principal vector of Chagas disease in Central America is the insect Triatoma dimidiata. The best methods of preventing household infestation with T. dimidiata (including Ecohealth interventions) involve mining large amounts of socioeconomic and entomologic survey data (comprised of nominal and ordinal data types) for numerous potential risk factors. The number of risk factors suggested by experts is too large for exhaustive search; and the use of traditional statistics can exclude risk factors that are purely epistatic. Therefore, we apply a novel evolutionary algorithm, the conjunctive clause evolutionary algorithm (CCEA), to mine these "Big Datasets" for the most important risk factors associated with T. dimidiata infestation using georeferenced survey data from two villages in Guatemala as examples. The CCEA identified socioeconomic risk factors to be important that are not significant using traditional statistics.

Keywords: Chagas disease, evolutionary algorithm, big data, data mining, Ecohealth

[A] University of Vermont

doi: 10.18278/jpcs.3.2.8

摘要

查加斯病是一种被忽视的致命性热带流行病，它在中美洲和南美洲的每个国家都有发生。中美洲查加斯病的主要载体是一种名为猎蝽（Triatoma dimidiata）的昆虫。为防止家庭感染，最好的方法（包括生态健康干预）是挖掘大量的社会经济调查数据和昆虫调查数据（由名义型数据和顺序型数据组成），找到诸多潜在的风险因素。由于专家建议的风险因素数量太大，无法使用穷举搜索（exhaustive search），并且使用传统统计学方法会排除具有完全上位效应（epistatic）的风险因素，因此，本文以危地马拉的两个村庄为例，采用一项创新性进化算法—the conjunctive clause 进化算法（conjunctive clause evolutionary algorithm，简称CCEA）—挖掘当地的“大数据”（调查数据），寻找与猎蝽有关的最重要风险因素。CCEA识别了社会经济风险因素的重要性，而传统统计学方法却并未做到这一点。

关键词：查加斯病，进化算法，大数据，数据挖掘，生态健康

Resumen

La enfermedad de Chagas es un padecimiento tropical, mortal e ignorado, endémico de todos los países de Centro y Sudamérica. El portador principal de la enfermedad de Chagas en Centroamérica es el insecto Triatoma dimidiata. Los mejores métodos para prevenir una infestación doméstica de T. dimidiata (Incluyendo intervenciones de Ecohealth) involucran la minería de grandes cantidades de datos de encuestas socioeconómicas y entomológicas (compuestos de tipos de datos nominales y ordinales) para un gran número de factores potenciales de riesgo. El número de factores de riesgo sugerido por expertos es muy grande para una búsqueda exhaustiva; y el uso de estadísticas tradicionales puede excluir factores de riesgo que son únicamente epistáticos. Por lo tanto, aplicamos un nuevo algoritmo evolutivo, el conjunctive clause evolutionary algorithm (CCEA), para minar estos sets de "macrodatos" para encontrar los factores de riesgo más importantes asociados con la infestación de T. dimidiata al utilizar datos de encuestas georreferenciados provenientes de dos aldeas en Guatemala como ejemplos. El CCEA identificó factores de riesgo socioeconómicos como importantes, que no serían significativos si se utilizaran métodos estadísticos tradicionales.

Palabras clave: enfermedad de Chagas, algoritmo evolutivo, macrodatos, minería de datos, Ecohealth

Introduction

The protozoan parasite, *Trypanosoma cruzi*, causes Chagas disease and is primarily spread via blood feeding insects in the order Hemiptera, family Reduviidae, and subfamily Triatominae (WHO, 2002). While the primary vector food source is vertebrates, *Trypanosoma cruzi* only infects mammals (Rassi, Rassi, & Marin-Neto, 2010). Human impacts, such as deforestation for agrarian land use, have caused triatomines to adapt (Coura, 2008); and one of the main vectors of Chagas disease, *Triatoma dimidiata*, has adapted to human domestic and peridomestic environments (Waleckx, Gourbière, & Dumonteil, 2015). This vector is endemic to Guatemala and researchers find it in Mexico all the way south to parts of Peru, Ecuador, Colombia, and Venezuela (WHO, 2002). People with Chagas disease often live in remote areas with poor sanitation, low socioeconomic status, and work manual labor jobs (Briceño-León & Méndez Galván, 2007; Prata, 2001). Approximately 70 million people in Latin America are at risk of infection with T*rypanosoma cruzi* and ~5.7 million people are infected (Chagas, 2015). In Central America, Guatemala has the largest number of vector transmitted cases (e.g., ~1,275 in 2010) (Chagas, 2015). Furthermore, Guatemala, El Salvador, and Honduras combined account for 85% of the new cases in Central America (Chagas, 2015). Chagas disease has an estimated disability-adjusted life years (DALY) of 546,000 (271,000–1,054,000) and is the second largest proportion of DALY for neglected tropical diseases in central Latin America, after hookworm disease (Murray et al., 2012). The estimated annual healthcare cost per Chagas patient in Latin America is ~$383 (range $207–$636); and the total annual cost to society (i.e., healthcare plus productivity losses) per chronic Chagas disease patient in Latin America is ~$4,059 (range $3,569–$4,434) (Lee, Bacon, Bottazzi, & Hotez, 2013). Thus, the disease burden of Chagas disease exceeds other infectious diseases such as cholera and rotavirus (Lee et al., 2013).

Humans infected with *Trypanosoma cruzi* can acquire Chagas disease by the transmission of the *Trypanosoma cruzi*-infected feces into the bite or open wound, or through the mucosa of the eye, nose, or mouth (Prata, 2001; Rassi et al., 2010; Teixeira, Nitz, Guimaro, Gomes, & Santos-Buch, 2006). Another possible source is via consumption of the infected feces in food items such as juice, vegetables, and possibly wild meat (Rueda, Trujillo, Carranza, & Vallejo, 2014). Oral transmission is believed to be the primary source of infection for wild animals (Coura, 2015); and the odoriferous glands of a marsupial infected with *Trypanosoma cruzi* can directly transmit the parasite to humans (Coura, 2015).

Chagas disease has three distinct phases. The first is the acute phase, which may last 1–4 months after infection with *Trypanosoma cruzi* (Prata, 2001; Stanaway & Roth, 2015). The characteristics of this phase are an increase in heart size, heart cell destruction, and depopulation of neurons (Teixeira et al., 2006). The acute phase is asymptomatic in 95% of the cases (Stanaway & Roth, 2015; Teixeira et al., 2006). However, for the remaining 5%, symptoms may include malaise, fever, jaundice, skin hemorrhages, enlargement of the liver, and muscle and joint pain (Prata, 2001; Rassi et al., 2010; Stanaway & Roth, 2015; Teixeira et al., 2006) and less than 1 in 2,500 infections result in death. Physicians usually attribute the latter to encephalomyelitis or severe cardiac failure (Prata, 2001; Stanaway & Roth, 2015; Teixeira et al., 2006). The *indeterminate* phase is asymptomatic, usually follows the acute phase, and can last 10–30 years or throughout a lifetime (Prata, 2001; Stanaway & Roth, 2015). Finally, the chronic phase, which usually follows the indeterminate phase, has symptoms that include heart disease (i.e., destruction of target heart cells), megaesophagus, megacolon, nervous system lesions, and sudden death (Prata, 2001; Rassi et al., 2010; Stanaway & Roth, 2015; Teixeira et al., 2006). Heart disease is one of the most common and deadly symptoms; however, there appears to be heterogeneity in Chagas-related heart disease with three distinct groups (Rassi et al., 2006). The 10-year mortality rate across all three groups is ~27%, but ranges from ~10% to ~84% (Rassi et al., 2006). In general, the relative risk ratio of mortality is ~1.74 for individuals with Chagas disease compared with similar individuals without the disease (Cucunubá, Okuwoga, Basáñez, & Nouvellet, 2016). In addition to the Chagas-related health effects, there is some evidence that Chagas is a risk factor for high blood pressure (Vicco, Rodeles, Yódice, & Marcipar, 2014), cognitive impairment in older adults (Lima-Costa et al., 2008), and ischemic stroke (Lima-Costa et al., 2008).

Currently, there is no preventive medicine for Chagas disease. Nonetheless, there are two anti-trypanosome drugs, nifurtimox, and benznidazole, used to treat *Trypanosoma cruzi* infections (González-Ramos et al., 2016; Jannin & Villa, 2007; Rassi et al., 2010; Teixeira et al., 2006). Both drugs have a common occurrence of adverse reactions that can prevent infected individuals from completing treatment (da Silva et al., 2014; Hasslocher-Moreno et al., 2012; Molina et al., 2015; Olivera, Cucunuba, Alvarez, & Nicholls, 2015). Studies found that between 13% and 31% cannot complete drug treatment (da Silva et al., 2014; Hasslocher-Moreno et al., 2012; Molina et al., 2015; Olivera et al., 2015).

Benznidazole is the preferred and most effective treatment of *Trypanosoma cruzi* infections (González-Ramos et al., 2016; Prata, 2001; Rassi et al., 2010). However, adverse reactions in adults include epigastric pain, skin disorders, nausea, abdominal bloating, sleep disturbance, temporary memory loss, headache, loss of appetite, myalgia, eosinophilia, and central and peripheral nervous system disorders (da Silva et al., 2014; Hasslocher-Moreno et al., 2012; Molina et al., 2015; Olivera et al., 2015). The percentage of adults with at least one of these reactions ranges from 49% to 80% (da Silva et al., 2014; Hasslocher-Moreno et al., 2012; Molina et al., 2015; Olivera et al., 2015). Extreme cases have resulted in intensive care unit treatment with symptoms that included tonic–clonic seizures and in one case, decreased liver function and multiple general organ failure accompanied by 30% skin detachment in another case (González-Ramos et al., 2016). Currently, drug treatment is optional for people over 50, because no proven benefits exist for people in this age cohort (Rassi et al., 2010). For other age cohorts, treatment efficacy for people with acute Chagas is between 30% and 83.5% (Jannin & Villa, 2007; Prata, 2001; Teixeira et al., 2006); and for chronic Chagas, the efficacy is much lower (5%–30%) (Jannin & Villa, 2007; Teixeira et al., 2006).

Thus, given the lack of preventative medicine, coupled with the adverse drug reactions and low efficacy of drug treatment, the preferred method of combating Chagas disease is by minimizing human contact with the vector. One of the most common tactics for controlling household infestation of *T. dimidiata* is the use of pyrethroid insecticide (Acevedo, Godoy, & Schofield, 2000; Dumonteil et al., 2004; Hashimoto, Cordon-Rosales, Trampe, & Kawabata, 2006; Manne et al., 2012; Nakagawa, Cordón-Rosales, Juárez, Itzep, & Nonami, 2003; Nakagawa, Hashimoto, et al., 2003; Quinde-Calderón, Rios-Quituizaca, Solorzano, & Dumonteil, 2016; Tabaru, Monroy, Rodas, Mejia, & Rosales, 1998; Yoshioka et al., 2015). While pyrethroid insecticides have successfully reduced infestation rates, rarely is the infestation rate reduced to 0 (Acevedo et al., 2000; Quinde-Calderón et al., 2016). Nonetheless, the residual effects of pyrethroid spraying appear to last only 4 months before adult *T. dimidiata* re-infest a house and 9 months before nymphs colonize the house (Dumonteil et al., 2004). Thus, while residual pyrethroid spraying prevents infestation by *Rhodnius prolixus* and *Triatoma infestans*, in most cases the prevention of *T. dimidiata* does not occur (Waleckx, Gourbière, et al., 2015). Researchers observed that the rebounding of infestation to original levels occurred almost 3 years after a single round of pyrethroid spraying in Jutiapa, Guatemala (Hashimoto et al., 2006). In addition to spraying houses with pyrethroid insecticides, recent work shows the potential of the fungi *Beauveria bassiana* and *Gliocladium virens* to control *T. dimidiata* (Vázquez-Martínez, Cirerol-Cruz, Torres-Estrada, & López, 2014). While researchers show that both fungi successfully kill T. dimidiata in a laboratory setting, Vázquez-Martínez et al. (2014) also show that T. dimidiata can transmit the deadly fungi to other T. dimidiata. However, short of extirpation of T. dimidiata, the vector will always pose the risk of infestation where it is endemic.

The only proven long-term control of *T. dimidiata* infestation is the implementation of home improvements often accompanied by educational outreach on Chagas disease and the vector (De Urioste-Stone et al., 2015; Ferral et al., 2010; Monroy et al., 2009). Home improvements that minimize the risk of household infestation with *T. dimidiata* run the gamut of cleaning and organizing the peridomestic environment immediately surrounding a house (Ferral et al., 2010; Zeledón & Rojas, 2006; Zeledón et al., 2008), plastering walls (Lucero et al., 2013; Monroy et al., 2009; Monroy, Rodas, Mejia, & Tabaru, 1998; Pellecer, Dorn, Bustamante, Rodas, & Monroy, 2013), replacing dirt floors with cement-like floors (Lucero et al., 2013; Pellecer et al., 2013), installing window screens (Ferral et al., 2010; Waleckx, Camara-Mejia, et al., 2015), impregnating curtains with insecticide (Ferral et al., 2010), and domestic rodent control (De Urioste-Stone et al., 2015). These home improvements have led to a reduction in household infestation that often lasts longer than spraying insecticide; however, none has eliminated infestation. Some researchers consider some of the aforementioned interventions Ecohealth interventions because they use sustainable methods and locally sourced materials (Lucero et al., 2013; Monroy et al., 2009; Pellecer et al., 2013). Window screens were very effective in nearly eliminating household infestation in pilot villages in the Yucatan, Mexico; however, in the Yucatan, *T. dimidiata* has been shown to invade homes seasonally, and thus barriers to entry are more effective (Ferral et al., 2010; Waleckx, Camara-Mejia, et al., 2015). Therefore, identifying unique combinations of factors or patterns that target cost-effective, mitigation strategies may help successfully eliminate infestation.

To determine the complex interactions associated with Chagas disease, a number of studies will perform univariate statistical analysis as a feature selection tool. Features below a designated p-value are then selected for multivariate analysis (da Silva et al., 2014; Kaplinski et al., 2015; King, Cordon-Rosales, Cox, Davies, & Kitron, 2011; Molina et al., 2015; Olivera et al., 2015; Rassi et al., 2006; Weeks et al., 2013). By performing feature selection using univariate statistical analysis, features that have epistatic interactions will most likely not be selected for use in subsequent multivariate models. This should be of particular concern with respect to the triatomine vectors of Chagas disease because infestation is inherently associated with epistatic interactions when the study views the system as ecological niche modeling. For triatomine vectors to survive in and infest a house, the vector requires, at a minimum, a source of shelter and a readily available food source. Other system features may be of importance (e.g., house attractiveness by the vector such as initial entry or passive modes of transportation into the house). Therefore, household infestation of triatomine vectors is a complex nonlinear system with epistatic relationships between features that are often included as potential risk factors in modeling triatomine infestation. Given the large number of potential risk factors associated with triatomine infestation, it is natural for studies to reduce the number of model features *a priori* because inclusion of large numbers of features makes exhaustive search of all possible models prohibitively expensive and/or impossible. In addition, *a priori* feature reduction may help remove noisy features that may lead to overfitting in the multivariate models. As a result, Bustamante et al. (2015) held a workshop in order to pre-select 25 features for multivariate modeling of *T. dimidiata*. Features were selected because previous studies showed they increase the odds of infestation; thus, the features selected are inherently biased toward previous univariate model selection (i.e., not able to search the datasets for epistatic interactions). In addition, missing data challenge many statistical methods. Thus, the removal of an entire observation (e.g., house) because it contains one or two missing features often results in the loss of many other features that may contain important information and change the individual feature distributions (Bustamante et al., 2015). Finally, another challenge associated with the statistical modeling of with triatomine vector infestation is the use of mixed data types. The inclusion of real, ordinal, and nominal input data is not possible for all statistical methods; and as a result, real-valued features are often converted to bins that represent a range. For instance, the number of chickens might get binned into categories (e.g., 0, 1–3, 4–10, etc.), when in reality, the number of features needed to represent all possible ranges for the number of chickens could be very large. As a result, Bustamante et al. (2015) used expert knowledge to create four binned classes to represent the ranges of chickens in the house. While there is nothing wrong with employing expert knowledge, especially when there is no other reasonable alternative, the posterior tinkering of features can result in the reinforcement of preconceived researcher hypotheses.

Finally, there is evidence of heterogeneity in modeling infestation with triatomine vectors of Chagas disease since both Bustamante, De Urioste-Stone, Juárez, and Pennington (2014) and Bustamante et al. (2015) found no statistical support for a single best model of infestation. Thus, any successful statistical modeling tool would need to consider model heterogeneity. The goals of this manuscript are two-fold. The first is to call attention to this neglected disease and the large number of risk factors that contribute to its transmission. The second is to show preliminary proof-of-concept of a recently developed evolutionary algorithm (Hanley, Eppstein, Buzas, & Rizzo, 2016; Hanley, Rizzo, Buzas, & Eppstein, 2017) designed to mine “Big” survey data for the most important risk factors associated with *T. dimidiata* infestation using georeferenced household data from two towns in Guatemala as examples.

Evolutionary Algorithms

Evolutionary algorithms are biologically inspired algorithms that have been used to solve a range of complex problems. While there are many different types of evolutionary algorithms, all evolutionary algorithms use populations of possible solutions where usually the fit solutions survive and/or produce offspring. Offspring are created every generation using crossover, mutation, or a combination of the two methods. Crossover involves the combination of parts of fit solutions to create a new solution, while mutation usually consists of random minor changes to a fit solution (Eiben & Smith, 2010). Using these core concepts of evolution, evolutionary algorithms are able to evolve a population of fit solutions for problems that have complex nonlinear search spaces.

In this article, we apply a novel evolutionary algorithm, the conjunctive clause evolutionary algorithm (CCEA) (Hanley et al., 2016, 2017), to explore the complex epistatic, heterogeneous, nonlinear interactions associated with the infestation of two towns in Guatemala with the Chagas vector *T. dimidiata*.

Methods and Material

Study Sites

Our study sites are the small rural towns of El Chaperno and El Carrizal located in the dry highlands of Jutiapa, Guatemala (red and yellow dots of Figure 1). Jutiapa, Guatemala (highlighted in red, Panel A) borders El Salvador with the study site locations shown as a yellow star. El Carrizal (Panel B) has spur roads radiating from the main road making the town less linear in shape. El Chaperno (Panel C) is linear in shape since most of the houses are adjacent to the principal road running through the town. In addition, El Chaperno is more heavily forested than El Carrizal due to its forest conservation efforts.

From October 1–3, 2012 in El Chaperno and February 4–5, 2013 in El Carrizal, teams comprised of personnel from the Escuela de Biología, La Universidad de San Carlos de Guatemala, and the Guatemalan Ministry of Health Office of Vector-Borne Diseases conducted entomological and socioeconomic surveys of 182 and 129 houses, respectively. Informed consent was obtained from all adult participants and from parents or legal guardians of minors. This project received ethical clearance from the Ministry of Health in Guatemala, La Universidad de San Carlos bioethics committee, and the Panamerican Health Organization.

The El Chaperno and El Carrizal house surveys contained 64 features thought to be potential risk factors for infestation with *T. dimidiata*. The dataset of each community was analyzed separately, and then combined and re-analyzed to test for larger-scale regional patterns. Given the difficulty in finding live *T. dimidiata* (Monroy et al., 1998), and because we are interested in identifying features associated with the risk of house infestation that help further the development of intervention strategies, we define infestation as any sign of T. dimidiata presence in the house (i.e., live or dead vectors, eggs, exuviae, or feces) as we believe these signs of T. dimidiata are indicative that the house is either currently infested or has been infested in the recent past. El Carrizal has a higher percentage of infested houses than El Chaperno does; however, both data-

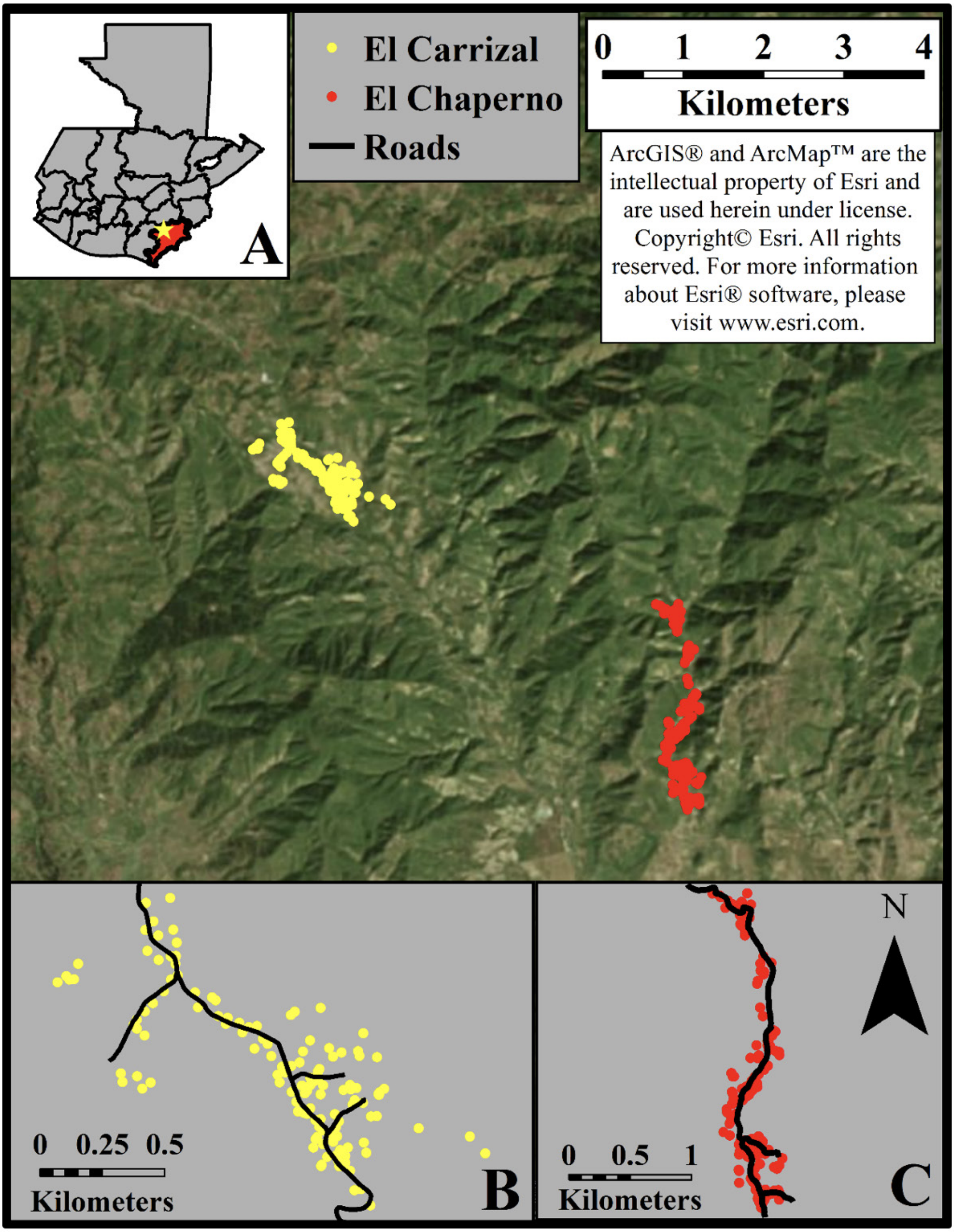


Figure 1. Satellite image of the study sites with the houses in El Chaperno and El Carrizal represented as red and yellow dots, respectively. Panel A is a map of the departments of Guatemala with the department of Jutiapa highlighted in red and the location of the study sites represented as a yellow star. Panels B and C show the locations of the houses and roads in El Carrizal and El Chaperno, respectively.

Table 1. Summary of the dataset characteristics for El Chaperno, El Carrizal, and the two towns combined

Dataset	# Houses	# Infested houses	# Ordinal, nominal, binary features	% Missing data	Median % missing data per feature	[Min, Max] % missing data per feature
El Chaperno	182	49 (26.9%)	[12, 8, 44]	28.9	15.7	[0.5, 86.8]
El Carrizal	129	51 (39.5%)	[14, 8, 42]	22.3	3.9	[0.8, 77.5]
Combined	311	100 (32.2%)	[14, 8, 42]	26.1	10.3	[1.2, 78.5]

sets have imbalanced outputs in the sense that the percent of infested houses is in the minority (Table 1, column 3).

Combinations in Datasets

While the number of households in a dataset may be small (e.g., 129 – 311) by some standards, the total number of features in the dataset makes these datasets too big to examine every combination of multivariate interactions using exhaustive search. For instance, assuming a dataset contains features with the same number of categorical responses per feature, then the number of potential feature value combinations for a given order of interaction can be calculated with $v^{O}\binom{L}{O}$ where v represents the number of values per feature, O represents the order of interaction, and L represents the number of features in the dataset. Where the first half of the equation, and v^{O} represents the total combination of feature values for the selected features. The term, $\binom{L}{O}$, represents the total number of feature combinations for the given order of interaction O. If we take a hypothetical dataset with 50 nominal features, each with five categorical values and each model is limited to one category per feature, then the number of models with two features is 3.06×10^4, three features is 2.45×10^6, and four features is 1.44×10^8. It should be noted that models that do not allow for a range of values for ordinal features could bias models against ordinal features. Therefore, when we test models that have

Table 2. The number of possible models for second- to fifth-order feature interactions for the El Chaperno, El Carrizal, and the combined datasets

Dataset	Number of combinations per order of feature interaction			
	2	3	4	5
El Chaperno	$3.91{\times}10^{5}$	$8.75{\times}10^{7}$	$1.25{\times}10^{10}$	$1.25{\times}10^{12}$
El Carrizal	$3.41{\times}10^{5}$	$7.07{\times}10^{7}$	$9.45{\times}10^{9}$	$8.90{\times}10^{11}$
Combined	$6.00{\times}10^{5}$	$1.60{\times}10^{8}$	$2.70{\times}10^{10}$	$3.13{\times}10^{12}$

ranges of values for both ordinal and nominal features, the number of two-way interactions (i.e., bivariate models) is on the order of hundreds of thousands (Table 2). For 5thfifth-order interactions, there are over one trillion possible models for two of the three datasets.

Conjunctive Clause Evolutionary Algorithm

We designed the CCEA to search the survey data for multivariate interactions across multiple data types (i.e., nominal, ordinal, and continuous) (Hanley et al., 2016, 2017). CCEA is a nonparametric statistical tool that searches across the entire range of multivariate interactions for each feature, where each feature can comprise features sets (ranges of values) that vary in size. The only assumption inherent in the models evolved by the CCEA is that ordinal and continuous features must be monotonic or unimodal. The CCEA evolves feature sets and the range of feature values using conjunctive clauses in the following form:

$$CC_k \coloneqq F_i \in a_i \wedge F_j \in a_j \ldots \quad (1)$$

where the term, :=, means "is defined as," F_i represents a feature i that may be nominal, ordinal, or continuous, and whose value lies in a_i, $\wedge$ and represents conjunction (i.e., logical AND). Note that a_i is a specified range or set of values that is a proper non-empty subset of a pre-specified universal set or a maximum range of each feature. The meaning of such a clause is interpreted as "if CC_k is true for a given input feature vector, then the class outcome is predicted to be associated with k." Association in this case means that the clause is more often associated with k than one would expect given the global distribution of k. Each one of the clauses, or groups of clauses, could be used as a classifier by stating that if CC_k matches an input feature vector, then classify it as k, else classify it as $\neg k$.

The CCEA is implemented using a customized version of an Age-Layered Population Structure (ALPS) (Hornby, 2006), with five linearly spaced age-layers and an age gap of 5 (Figure 2). In this study, we restrict each CCEA layer to a population size of L (where L is 64, the total number of features in the input vectors). In the CCEA, an additional sixth layer is used as an archive of probabilistically significant clauses. The CCEA was run for 200 generations and five repetitions.

At the start of the first generation (and every 5 generations there after), a novel population of clauses, each with age 1, is introduced into the first age layer. During each generation, all of the individuals in layers 1-–5, plus up to $L \times 5$ of the youngest individuals from the archived layer 6 (or fewer, if the archive does not yet hold this many individuals) are selected to reproduce with variation. The ages of these selected parents are incremented by 1; and they remain in the population. Variation is introduced either through crossover (with probability $P_x = 0.5$) or through mutation. If selected for crossover, a second parent is selected from the same or preceding (if one exists) age layer, using tournament selection with replacement (tournament size of 3); the age of the second parent is not incremented.

If selected for mutation, each feature from the parent is selected with probability $1/L$ (if zero features were initially selected, we select one at random). For each feature i that was selected, if the selected feature is not present in the clause, then the feature is added to the clause; and a_i is randomly initialized to a nonempty set or a range of allowable values that does not include the entire allowable subset or a range of values. However, if the feature is present in the

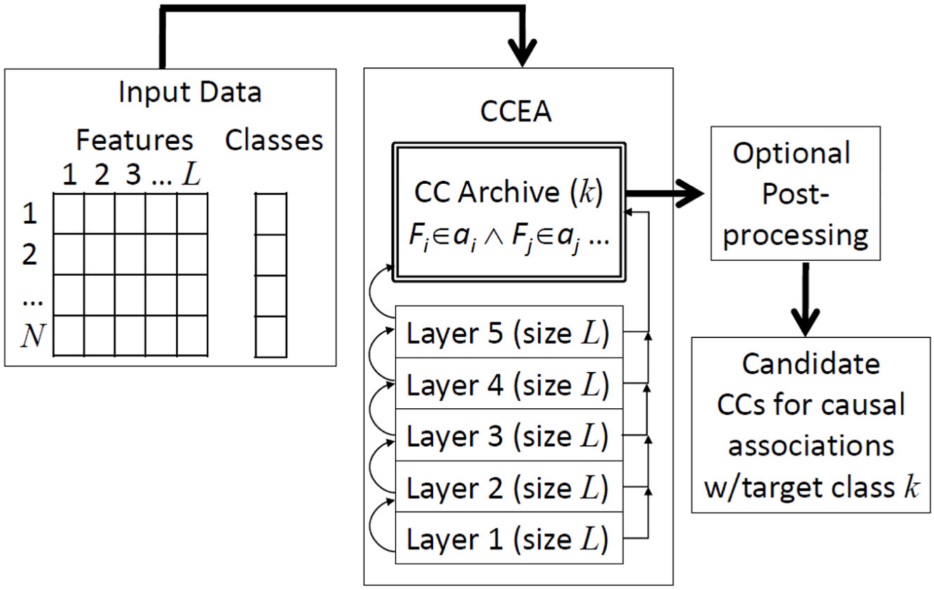


Figure 2. Flowchart for the ALPS-based CCEA. For a given target class k, we use the CCEA to evolve an archive of conjunctive clauses (CCs) that have a statistically significant probability of being associated with outcome class k; the CCs can be of arbitrary order, thus representing feature interactions. After running the CCEA, there is an optional post-processing step that can be used to tease out feature associations with target class k that may aid in the selection of causal CCs.

conjunctive clause, then with probability P_{wc}, the feature is removed. For this work, we selected a high probability (P_{wc} = 0.75), so that mutation favors order reduction and thus, aids in evolving parsimonious clauses that contain as few features as possible. If the feature Fi is not removed, then the corresponding ai is mutated as follows. If F_i is nominal, we randomly change, add, or delete a categorical value to ai, ensuring that the set remains nonempty and less than the allowable universal set of values. If F_i is ordinal or continuous, we randomly change the lower or upper bound of a_i, ensuring that the range remains nonempty and less than the maximum allowable range.

Every fifth generation, individuals in layers 1–5 age out of their layers into the next higher age layer and a new random population is created for layer 1. Those aging out of layer 5 are discarded from the population.

Fitness Function

To determine whether a conjunctive clause is probabilistically significant, the CCEA uses the hypergeometric probability mass function (PMF) (Kendall, 1952) as the fitness function. For conjunctive clauses evolved in the CCEA, the hypergeometric PMF is defined as follows:

$$Hypergeometric\ PMF = \frac{\binom{X_{tot}}{x_{match}}\binom{N_{tot}-X_{tot}}{n_{match}-x_{match}}}{\binom{N_{tot}}{n_{match}}} \quad (2)$$

where N_{tot} is the total number of observations that have non-missing values for the feature combination, X_{tot} is the total number of observations in the target class, *k*, that have non-missing values for the feature combination, n_{match} is the number of samples whose features match a given clause, and x_{match} is the number of observations that match the clause and are in target class *k*.

Equation. (2) quantifies the likelihood that the observed association between the clause and the target class is due to chance; thus, lower values of this fitness function are indicative of potential association between a clause and a target class. A conjunctive clause is considered to be probabilistically significant and thus worthy of being archived if its hypergeometric PMF is less than or equal to a threshold. In addition, the definitions of N_{tot} and X_{tot} help ensure that features with more missing data are not penalized for the missing data. Even though features with large percentages of missing data are less likely to form probabilistically significant multivariate conjunctive clauses, by defining N_{tot} and X_{tot} in this manner, as the total number of non-missing values, the hypergeometric PMF value, will be lower and thus more likely archived.

The CCEA can have a static threshold (i.e., the threshold will not heuristically decrease), or the threshold can deterministically evolve based on the number of archived conjunctive clauses for a given conjunctive clause order. In this work, we use a static threshold. Specifically, we archive conjunctive clauses that cover at least 10% or more of the houses infested with *T. dimidiata* by setting the hypergeometric fitness threshold to the fitness of a conjunctive clause that has 100% accuracy and 10% coverage of infested houses. Accuracy is defined as x_{match}/n_{match} and is analogous to the true positive rate of the conjunctive clause. Infested house coverage is the number of times a sampled conjunctive clause is associated with a target outcome over the total number of target outcomes in the dataset, x_{match}/X_{tot} . Note: While the accuracy and infested house coverage are related to the hypergeometric PMF, both are used as descriptive terms to show the expected true positive rate and generality of a conjunctive clause, respectively (this is akin to a more descriptive odds ratio that is often produced when performing logistic regression). If only a few conjunctive clauses are archived, we risk that the archived signals contain large amounts of noise and are subject to overfitting. As mentioned above, the CCEA used a static threshold to maintain a large population of archived conjunctive clauses. This is consistent with the concept in "Big Data" that more data can be used to find patterns of correlations with a desired output (true signal) (Mayer-Schönberger & Cukier, 2014).

The CCEA was run for five repetitions with 200 generations for each repetition for the El Chaperno, El Carrizal, and the combined datasets. The repetitions are seeded with a random number generator and provide another safeguard against the algorithm becoming trapped in one population of optima. For each dataset, we calculated the accuracy and infested house coverage of every archived conjunctive clause.

In addition to calculating the accuracy and infested house coverage, the population of archived conjunctive clauses was mined for patterns. For each repetition, the archived conjunctive clauses were analyzed on a house-by-house basis. The number of times a feature was present in a conjunctive clause that matched an infested house was calculated for all the features. These sums were then normalized between zero 0 and one1 for all the features. Thus, for every repetition and every dataset, there is a heat map matrix of values $\in$ [0,1] for every infested house and every feature. If an infested house is missing data for a given feature, then the corresponding cell in the matrix is not assigned a value. For each dataset, the maximum value across all five repetitions for each infested house and feature wereas then used to create new matrices (one per dataset) with values represented as heat maps (e.g., panels A-–C of Figure 4). Finally, for each dataset, "important" features were defined by the *majority* of infested houses had a value greater than 0.5. For heterogeneous datasets, one would not expect that one model and possibly one feature would cover all the infested houses; thus, the majority of houses wasere used as a cutoff to allow for heterogeneity of statistical models. However, the user may select any percentage of infested houses for this threshold. In this case, 0.5 means that the feature was present in at least half of the archived conjunctive clauses that matched a given infested house. Taken together, the cutoffs employed in this manuscript indicate that the important features were in the majority of archived conjunctive clauses that matched the majority of infested houses. Univariate statistical analysis was performed on these important features using JMP v11.0.

Results

The median number of CCEA archived conjunctive clauses per repetition is 52,615, 67,036, and 53,996 for El Chaperno, El Carrizal, and the combined datasets, respectively. Figure 3 contains the accuracy and infested house coverage of all the conjunctive clauses identified using the CCEA for the El Chaperno (Figure 3A), El Carrizal (Figure 3B), and the combined datasets (Figure 3C). Each dot represents a conjunctive clause and is color-coded based on the conjunctive clause. As expected, lower order conjunctive clauses tend to have a higher coverage of infested houses and higher order conjunctive clauses tend to have higher accuracy. The CCEA identifies 2ndsecond-order conjunctive clauses (represented as red dots) with 100% accuracy when it is run on the El Chaperno and El Carrizal datasets individually (Figure 3A and B); but this is not the case, for the combined dataset (Figure 3C).

The heat maps for each feature plotted against infested houses shows how often a feature was present in an archived conjunctive clause for a given infested house (Figure 4). Red denotes features that are important, while blue corresponds to those of less importance; white simply indicates missing data. Using the majority of the infested houses having a value greater than 0.5 as the threshold, the heat map indicates that 14 features that are important for El Chaperno, 23 are important for El Carrizal, and 15 for the combined datasets. Table 3 shows the name of the individual features for each dataset. The features embedded in conjunctive clauses that were identified as important across all three datasets are associated with: (1) the primary source of income is not business, (2) the primary source of income is not salary, (3) the household owns their home, (4) older homes, (5) longer periods of residency in the house, (6) accumulation of objects, (7) unhygienic beds, (8) adobe walls, (9) deteriorated bedroom walls, (10) deteriorated walls in the rest of the house, and (11) dirt floors. Features that were not previously identified as significant ($P>0.05$) using more traditional univariate methods at the combined town-level scale included features associated with primary source of income, home ownership, older homes, longer residency in the house, predominate wall material is adobe, and dirt floors.

Discussion

The CCEA was able to identify important sets of features across a range of conjunctive clause orders. While many of the archived conjunctive clauses may be considered noisy or overfitting, some interesting patterns emerge when the dataset is examined as a whole. Most importantly, the CCEA identifies socioeconomic features (i.e., source of income and home ownership) as important risk factors across all three datasets, while these same features are not found to be significant when using traditional statistics. While low socioeco-

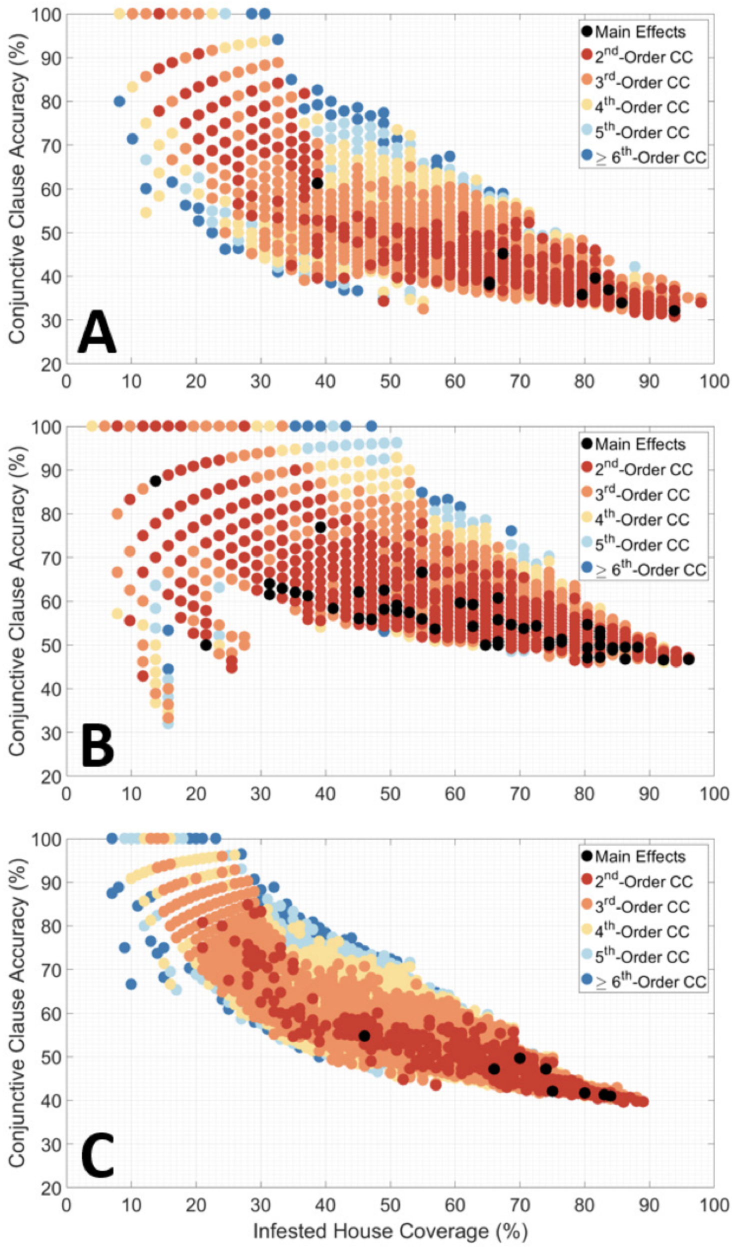


Figure 3. The accuracy and coverage of the conjunctive clauses identified using the CCEA for the (A) El Chaperno, (B) EL Carrizal, and (C) combined datasets. Each dot represents a conjunctive clause and they are color-coded based on the order of the conjunctive clause. Conjunctive clause accuracy is the number of time a conjunctive clause correctly classifies a house as infested represented as percent. Infested house coverage is the number of infested houses that the conjunctive clause matches represented in percent.

Table 3. Features identified as important by mining the CCEA archived conjunctive clauses for each of the three datasets

Feature	Survey question	Chaperno	Carrizal	Combined
2	Primary source of income of the head of household is day laborer		X	X
4	Primary source of income of the head of household is business	X	X	X
5	Primary source of income of the head of household is salary	X	X	X
6	Primary source of income of the head of household is other	X		
8	Do you own, rent, or are you borrowing the house?	X	X	X
9	How old is the house? (ordinal)	X	X	X
10	How long have you lived in the house? (ordinal)	X	X	X
18	How many dogs do you have?		X	
26	Are there signs of animals in the house?		X	X
27	Are there chicken nests in the house?		X	
29	Is there an accumulation of objects in the house?	X	X	X
34	Hygienic condition of the beds	X	X	X
35	Is the bed separated from the wall?		X	
36	Hygienic condition of the house		X	X
37	Are grains stored in the main bedroom?		X	
38	Is there an accumulation of firewood?		X	X
52	Is there an animal corral?		X	
53	What is the predominant material of the walls?	X	X	X
54	What is the condition of the walls in the bedroom?	X	X	X
55	What is the condition of the walls in the rest of the house?	X	X	X
57	What is the predominant material of the floor?	X	X	X
58	Is the bedroom dark (i.e., lacking natural light)?	X		
60	Does the bedroom have windows?	X		
61	Location of the kitchen?		X	
62	Does the house have plumbing?		X	
63	Does the house have electricity?		X	

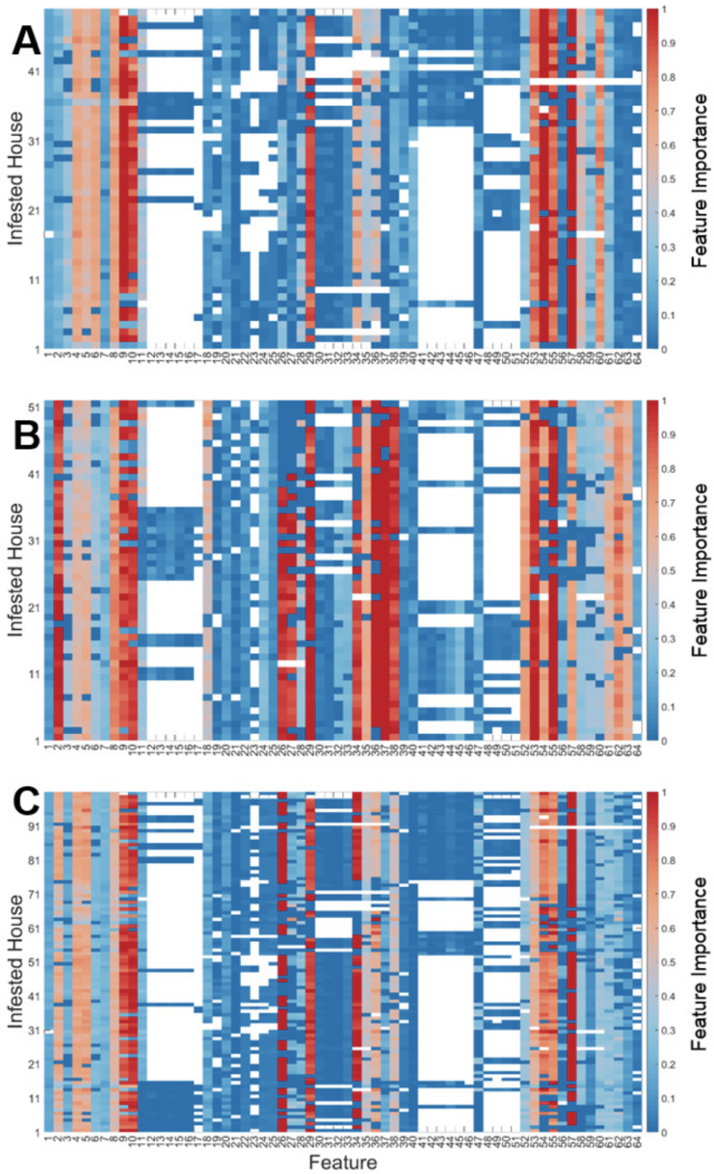


Figure 4. Heat map representing the importance of each feature for classifying a house as infested. Each column represents a feature presented to the CCEA; and each row represents an infested house in (A) El Chaperno, (B) El Carrizal, and (C) combined datasets. Red and blue denote features that are important to those that are not, repsectively; white indicates missing data.

nomic status is commonly linked to risk of infestation in the literature (Briceño-León & Méndez Galván, 2007; Prata, 2001), these factors are not identified as important features when using other statistical methodologies (Bustamante et al., 2014, 2015). Of the three socioeconomic risk factors identified by the CCEA as important, homeownership is associated with many of the archived conjunctive clauses across all three datasets. This feature may be particularly interesting for strategies designed to combat *T. dimidiata* infestation as the relationship between ownership and high infestation may be viewed several ways. First, it may be easier to persuade the large number of homeowners who live in infested homes to improve their homes (e.g., via the adoption of Ecohealth interventions) because they have the freedom to make improvements should resources be made available (compared, say, with those renting or borrowing the home). On the flipside, there may be a connection between economic status poverty and homeownership; and the reason that infested houses are associated with homeowners is that they cannot afford and/or lack the time to improve their house.

The CCEA is just one of many statistical tools that can be employed to analyze large, complex datasets. The goal of the CCEA is to identify probabilistically significant multivariate interactions that would either (1) be too costly (i.e., from a computational viewpoint) to search for higher-order, multivariate interactions or (2) not be detected using additive models. While we use the archived conjunctive clauses to identify important features in this article, these sets of features could be further analyzed in a number of ways. For example, future work may include analyzing infested houses associated with important features (sets of important features) for spatial patterns. Sensitivity analysis could be applied to higher-order feature sets to examine whether the model is more likely to be additive or epistatic to help improve mitigation strategies. For example, if both features in an archived second-order conjunctive clause (e.g., a large number of chickens and the proximity of the coop to the house) do not exhibit significant main effects, then the second-order clause exhibits epistatic interactions (both need to present to be associated with infestation). However, if both features exhibit significant main effects and when combined the set of features results in a more-fit second-order conjunctive clause, then the second-order conjunctive clause is additive. In addition, archived conjunctive clauses with lots of missing data could be isolated to see if those features show important trends; and if trends appear important, then more effort might be directed toward ensuring those feature values are collected in future surveys.

The goal of the CCEA algorithm is to provide an archive of conjunctive clauses with strong statistical signals that stakeholders can then use to fulfill their objectives. For instance, if researchers are interested in improving the floor material, they could mine the archived conjunctive clauses associated with floor material to see if there are other features (e.g., homeownership) that could be leveraged to help implement the replacement of dirt floors with cement-like floors. Cement-like floors not only reduce the risk of Chagas disease by eliminating a hiding place for *T. dimidiata* nymphs (Zeledón, Zúñiga, & Swartzwelder, 1969), they also help prevent infection with hookworm (Hotez, 2008), which is the only neglected tropical disease with a higher DALY than Chagas in central Latin America (Murray et al., 2012). The cost of replacing dirt floors with cement-like floors is ~$170 (Méndez, 2008). Guatemala has the highest prevalence of hookworm in Central

America (de Silva et al., 2003), with an estimated 3,000,000 people infected (Hotez et al., 2008). Given that the cost of hookworm infection in the Americas is ~$447 (range $40–$1,693) (Bartsch et al., 2016) and the annual total cost to society for each chronic Chagas disease patient is ~$4,059 (range $3,569–$4,434) (Lee et al., 2013), the cost of replacing dirt floors with cement-like floors is a fraction of the annual total cost to society. The hope is that researchers will use this new statistical tool as just one means of performing an unbiased statistical analysis of complex datasets often found in epidemiology.

While vectorial transmission of *Trypanosoma cruzi* is primarily limited to Latin America, Trypanosoma cruzi does exist in the United States. Researchers have identified 24 natural wildlife hosts of T*rypanosoma cruzi* in the southern portion of the United States ranging from Maryland to Florida and every state that borders Mexico (Bern, Kjos, Yabsley, & Montgomery, 2011). In addition to wildlife hosts, T*rypanosoma cruzi* has been identified in domestic dogs throughout the southern portion of the United States (Bern et al., 2011). Vectors infected with Trypanosoma cruzi endemic to the United States have been shown to feed on human blood, thus highlighting the risk of autochthonous human *Trypanosoma cruzi* infection (Waleckx, Suarez, Richards, & Dorn, 2014). In addition, there is evidence of limited autochthonous human *Trypanosoma cruzi* infection occurring in the southern portion of the United States (Bern et al., 2011; Cantey et al., 2012). In addition to the threat posed by triatomine transmission of *Trypanosoma cruzi*, recent laboratory research has shown that bed bugs (Cimex lectularius) can become infected with Trypanosoma cruzi after feeding on an infected host (i.e., mouse) and can serve as a competent vector of *Trypanosoma cruzi* (Salazar et al., 2015). Therefore, vectorial transmission of Trypanosoma cruzi in the United States has the potential to become a higher risk in the future; and if that were to occur, the CCEA could be used to mine Big Datasets for the most important risk factors.

Acknowledgements

We thank Laboratory for Applied Entomology and Parasitology (LENAP) at La Universidad de San Carlos Guatemala, especially Carlota Monroy, Antonieta Rodas, Lucia Orantes, Gabriela Rodas, Raquel Asuncion Lima, Elizabeth Solórzano, and Dulce Bustamante, for their help in collecting and managing the Chagas infestation dataset. We also thank the Guatemalan Ministry of Health for their help with collecting the data. This work was supported, in part, by Vermont EPSCoR with funds from the National Science Foundation Grant DBC-EID-1216193 (J.P.H.) and EPS-1101317 (D.M.R.).

References

Acevedo, F., Godoy, E., & Schofield, C. (2000). *Comparison of intervention strategies for control of Triatoma dimidiata* in Nicaragua. M*emórias Do Instituto Oswaldo Cruz*, *95*(6), 867–871.

Bartsch, S. M., Hotez, P. J., Asti, L., Zapf, K. M., Bottazzi, M. E., Diemert, D. J., & Lee, B. Y. (2016). The global economic and health burden of human hookworm infection. *PLoS Neglected Tropical Diseases*, *10*(9), e0004922.

Bern, C., Kjos, S., Yabsley, M. J., & Montgomery, S. P. (2011). Trypanosoma cruzi and Chagas' disease in the United States. *Clinical Microbiology Reviews*, *24*(4), 655–681.

Briceño-León, R., & Méndez Galván, J. (2007). The social determinants of Chagas disease and the transformations of Latin America. *Memórias Do Instituto Oswaldo Cruz*, 102 (Suppl 1), 109–112.

Bustamante, D. M., De Urioste-Stone, S. M., Juárez, J. G., & Pennington, P. M. (2014). Ecological, social and biological risk factors for continued Trypanosoma cruzi transmission by Triatoma dimidiata in Guatemala. *PLoS ONE*, *9*(8), e104599.

Bustamante, Z. D. M., Hernandez, M. M., Torres, N., Zuniga, C., Sosa, W., de Abrego, V., & Monroy Escobar, M. C. (2015). Information to act: Household characteristics are predictors of domestic infestation with the Chagas Vector Triatoma dimidiata in Central America. *American Journal of Tropical Medicine and Hygiene*, *93*(1), 97–107.

Cantey, P. T., Stramer, S. L., Townsend, R. L., Kamel, H., Ofafa, K., Todd, C. W., … Montgomery, S. P. (2012). The United States Trypanosoma cruzi Infection Study: Evidence for vector-borne transmission of the parasite that causes Chagas disease among United States blood donors: VECTOR-BORNE T. CRUZI IN US DONORS. *Transfusion*, *52*(9), 1922–1930.

Chagas disease in Latin America: an epidemiological update based on 2010 estimates. (2015). *Releve Epidemiologique Hebdomadaire*, *90*(6), 33–43.

Coura, J. R. (2015). The main sceneries of Chagas disease transmission. The vectors, blood and oral transmissions — A comprehensive review. *Memórias Do Instituto Oswaldo Cruz*, *110*(3), 277–282.

Cucunubá, Z. M., Okuwoga, O., Basáñez, M.-G., & Nouvellet, P. (2016). Increased mortality attributed to Chagas disease: A systematic review and meta-analysis. *Parasites and Vectors*, *9*(1), 42.

da Silva, G. M. S., Mediano, M. F. F., Alvarenga Americano do Brasil, P. E., da Costa Chambela, M., da Silva, J. A., de Sousa, A. S., … Hasslocher-Moreno, A. M. (2014). A clinical adverse drug reaction prediction model for patients with Chagas disease treated with Benznidazole. *Antimicrobial Agents and Chemotherapy*, *58*(11), 6371–6377.

de Silva, N. R., Brooker, S., Hotez, P. J., Montresor, A., Engels, D., & Savioli, L. (2003). Soil-transmitted helminth infections: Updating the global picture. *Trends in Parasitology*, *19*(12), 547–551.

De Urioste-Stone, S. M., Pennington, P. M., Pellecer, E., Aguilar, T. M., Samayoa, G., Perdomo, H. D., … Juarez, J. G. (2015). Development of a community-based intervention for the control of Chagas disease based on peridomestic animal management: An eco-bio-social perspective. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *109*(2), 159–167.

Dumonteil, E., Ruiz-Piña, H., Rodriguez-Félix, E., Barrera-Pérez, M., Ramirez-Sierra, M. J., Rabinovich, J. E., & Menu, F. (2004). Re-infestation of houses by Triatoma dimidiata after intra-domicile insecticide application in the Yucatán peninsula, Mexico. *Memórias Do Instituto Oswaldo Cruz*, *99*(3), 253–256.

Eiben, A. E., & Smith, J. E. (2010). *Introduction to evolutionary computing* (1. ed., 2. printing, softcover version of original hardcover ed. 2003). Berlin: Springer.

Ferral, J., Chavez-Nunez, L., Euan-Garcia, M., Ramirez-Sierra, M. J., Najera-Vazquez, M. R., & Dumonteil, E. (2010). Comparative field trial of alternative vector control strategies for non-domiciliated Triatoma dimidiata. *American Journal of Tropical Medicine and Hygiene*, *82*(1), 60–66.

González-Ramos, J., Noguera-Morel, L., Tong, H. Y., Ramírez, E., Ruiz-Bravo, E., Bellón, T., … Herranz-Pinto, P. (2016). Two cases of overlap severe cutaneous adverse reactions to benznidazole treatment for asymptomatic Chagas disease in a nonendemic country. *British Journal of Dermatology*, *175*(3), 604–607.

Hanley, J. P., Eppstein, M. J., Buzas, J. S., & Rizzo, D. M. (2016). Evolving probabilistically significant epistatic classification rules for heterogeneous big datasets. In *Proceedings of the 18th Annual Conference on Genetic and Evolutionary Computation* (pp. 445–452). ACM Press.

Hanley, J. P., Rizzo, D. M., Buzas, J. S., & Eppstein, M. J. (2017). A tandem evolutionary algorithm for identifying optimal association rules from complex data. Manuscript submitted for publication.

Hashimoto, K., Cordon-Rosales, C., Trampe, R., & Kawabata, M. (2006). Impact of single and multiple residual sprayings of pyrethroid insecticides against Triatoma dimidiata (Reduviiade; Triatominae), the principal vector of Chagas disease in Jutiapa, Guatemala. *The American Journal of Tropical Medicine and Hygiene*, *75*(2), 226–230.

Hasslocher-Moreno, A. M., do Brasil, P. E. A. A., de Sousa, A. S., Xavier, S. S., Chambela, M. C., & Sperandio da Silva, G. M. (2012). Safety of benznidazole use in the treatment of chronic Chagas' disease. *Journal of Antimicrobial Chemotherapy*, *67*(5), 1261–1266.

Hornby, G. S. (2006). ALPS: the age-layered population structure for reducing the problem of premature convergence. In *Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation* (p. 815). ACM Press.

Hotez, P. (2008). Hookworm and poverty. *Annals of the New York Academy of Sciences*, *1136*(1), 38–44.

Jannin, J., & Villa, L. (2007). An overview of Chagas disease treatment. *Memórias Do Instituto Oswaldo Cruz*, 102, 95–98.

Kaplinski, M., Jois, M., Galdos-Cardenas, G., Rendell, V. R., Shah, V., Do, R. Q., … Bern, C. (2015). Sustained domestic vector exposure is associated with increased Chagas cardiomyopathy risk but decreased Parasitemia and congenital transmission risk among young women in Bolivia. *Clinical Infectious Diseases*, *61*(6), 918–926.

Kendall, M. G. (1952). *The advanced theory of statistics* (5th ed., Vol. 1). New York, NY: Hafner Publishing Company.

King, R. J., Cordon-Rosales, C., Cox, J., Davies, C. R., & Kitron, U. D. (2011). Triatoma dimidiata infestation in Chagas disease endemic regions of Guatemala: Comparison of random and targeted cross-sectional surveys. *PLoS Neglected Tropical Diseases*, *5*(4), e1035.

Lee, B. Y., Bacon, K. M., Bottazzi, M. E., & Hotez, P. J. (2013). Global economic burden of Chagas disease: A computational simulation model. *The Lancet Infectious Diseases*, *13*(4), 342–348.

Lima-Costa, M. F., Castro-Costa, E., Uchôa, E., Firmo, J., Ribeiro, A. L. P., Ferri, C. P., & Prince, M. (2008). A population-based study of the association between Trypanosoma cruzi infection and cognitive impairment in old age (The Bambuí Study). *Neuroepidemiology*, *32*(2), 122–128.

Lucero, D. E., Morrissey, L. A., Rizzo, D. M., Rodas, A., Garnica, R., Stevens, L., … Monroy, M. C. (2013). Ecohealth interventions limit triatomine reinfestation following insecticide spraying in La Brea, Guatemala. *American Journal of Tropical Medicine and Hygiene*, *88*(4), 630–637.

Manne, J., Nakagawa, J., Yamagata, Y., Goehler, A., Brownstein, J. S., & Castro, M. C. (2012). Triatomine infestation in Guatemala: Spatial assessment after two rounds of vector control. *The American Journal of Tropical Medicine and Hygiene*, *86*(3), 446–454.

Mayer-Schönberger, V., & Cukier, K. (2014). *Big data: A revolution that will transform how we live, work, and think* (First Mariner Books edition). Boston, MA: Mariner Books, Houghton Mifflin Harcourt.

Méndez, M. (2008). Evaluación de la Relación Beneficio-Costo al comparar intervenciones tradicionales y ecosistémicas realizadas para el control de los vectores de la enfermedad de Chagas en la aldea La Brea, a lo largo de diez años. Primera parte, comprende el periodo del 2000 al 2004. Evaluación Ex ante. (pp. 1–37).

Molina, I., Salvador, F., Sánchez-Montalvá, A., Treviño, B., Serre, N., Sao Avilés, A., & Almirante, B. (2015). Toxic profile of benznidazole in patients with chronic Chagas Disease: Risk factors and comparison of the product from two different manufacturers. *Antimicrobial Agents and Chemotherapy*, *59*(10), 6125–6131.

Monroy, C., Bustamante, D. M., Pineda, S., Rodas, A., Castro, X., Ayala, V., … Moguel, B. (2009). House improvements and community participation in the control of Triatoma dimidiata re-infestation in Jutiapa, Guatemala. *Cadernos De Saude Publica, 25 (Suppl 1)*, S168–S178.

Monroy, C., Rodas, A., Mejia, M., & Tabaru, Y. (1998). Wall plastering and paints as methods to control vectors of Chagas disease in Guatemala. *Medical Entomology and Zoology*, *49*(3), 187–193.

Murray, C. J. L., Vos, T., Lozano, R., Naghavi, M., Flaxman, A. D., Michaud, C., … Lopez, A. D. (2012). Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, 380(9859), 2197–2223.

Nakagawa, J., Cordón-Rosales, C., Juárez, J., Itzep, C., & Nonami, T. (2003). Impact of residual spraying on Rhodnius prolixus and Triatoma dimidiata in the department of Zacapa in Guatemala. *Memórias Do Instituto Oswaldo Cruz*, *98*(2), 277–282.

Nakagawa, J., Hashimoto, K., Cordón-Rosales, C., Abraham Juárez, J., Trampe, R., & Marroquín, M. L. (2003). The impact of vector control on Triatoma dimidiata in the Guatemalan department of Jutiapa. *Annals of Tropical Medicine and Parasitology*, *97*(3), 288–297.

Olivera, M. J., Cucunuba, Z. M., Alvarez, C. A., & Nicholls, R. S. (2015). Safety profile of Nifurtimox and treatment interruption for chronic Chagas disease in Colombian adults. *American Journal of Tropical Medicine and Hygiene*, *93*(6), 1224–1230.

Pellecer, M. J., Dorn, P. L., Bustamante, D. M., Rodas, A., & Monroy, M. C. (2013). Vector blood meals are an early indicator of the effectiveness of the Ecohealth approach in halting Chagas transmission in Guatemala. *American Journal of Tropical Medicine and Hygiene*, *88*(4), 638–644.

Prata, A. (2001). Clinical and epidemiological aspects of Chagas disease. *The Lancet Infectious Diseases*, *1*(2), 92–100.

Quinde-Calderón, L., Rios-Quituizaca, P., Solorzano, L., & Dumonteil, E. (2016). Ten years (2004–2014) of Chagas disease surveillance and vector control in Ecuador: Successes and challenges. *Tropical Medicine and International Health*, *21*(1), 84–92.

Rassi, A., Rassi, A., Little, W. C., Xavier, S. S., Rassi, S. G., Rassi, A. G., … Scanavacca, M. I. (2006). Development and validation of a risk score for predicting death in Chagas' heart disease. *The New England Journal of Medicine*, *355*(8), 799–808.

Rassi, A., Rassi, A., & Marin-Neto, J. A. (2010). Chagas disease. *Lancet (London, England), 375*(9723), 1388–1402.

Rueda, K., Trujillo, J. E., Carranza, J. C., & Vallejo, G. A. (2014). Transmisión oral de Trypanosoma cruzi: un nuevo escenario epidemiológico de la enfermedad de Chagas en Colombia y otros países suramericanos. *Biomédica*, *34*(4), 631–641.

Salazar, R., Castillo-Neyra, R., Tustin, A. W., Borrini-Mayori, K., Naquira, C., & Levy, M. Z. (2015). Bed bugs (Cimex lectularius) as vectors of Trypanosoma cruzi. *American Journal of Tropical Medicine and Hygiene*, *92*(2), 331–335.

Stanaway, J. D., & Roth, G. (2015). The burden of Chagas disease. *Global Heart*, *10*(3), 139–144.

Tabaru, Y., Monroy, C., Rodas, A., Mejia, M., & Rosales, R. (1998). Chemical control of Triatoma dimidiata and Rhodnius prolixus (Reduviidae : Triatominae), the principal vectors of Chagas' disease in Guatemala. *Medical Entomology and Zoology*, *49*(2), 87–92.

Teixeira, A. R. L., Nitz, N., Guimaro, M. C., Gomes, C., & Santos-Buch, C. A. (2006). Chagas disease. *Postgraduate Medical Journal*, *82*(974), 788–798.

Vázquez-Martínez, M. G., Cirerol-Cruz, B. E., Torres-Estrada, J. L., & López, M. H. R. (2014). Potential for entomopathogenic fungi to control Triatoma dimidiata (Hemiptera: Reduviidae), a vector of Chagas disease in Mexico. *Revista Da Sociedade Brasileira de Medicina Tropical*, *47*(6), 716–722.

Vicco, M. H., Rodeles, L., Yódice, A., & Marcipar, I. (2014). Chagas disease, a risk factor for high blood pressure. *Blood Pressure*, *23*(6), 345–348.

Waleckx, E., Camara-Mejia, J., Ramirez-Sierra, M. J., Cruz-Chan, V., Rosado-Vallado, M., Vazquez-Narvaez, S., … Dumonteil, E. (2015). An innovative ecohealth intervention for Chagas disease vector control in Yucatan, Mexico. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *109*(2), 143–149.

Waleckx, E., Gourbière, S., & Dumonteil, E. (2015). Intrusive versus domiciliated triatomines and the challenge of adapting vector control practices against Chagas disease. *Memórias Do Instituto Oswaldo Cruz*, *110*(3), 324–338.

Waleckx, E., Suarez, J., Richards, B., & Dorn, P. L. (2014). Triatoma sanguisuga blood meals and potential for Chagas disease, Louisiana, USA. *Emerging Infectious Diseases*, *20*(12), 2141–2143.

Weeks, E. N. I., Cordón-Rosales, C., Davies, C., Gezan, S., Yeo, M., & Cameron, M. M. (2013). Risk factors for domestic infestation by the Chagas disease vector, Triatoma dimidiata in Chiquimula, Guatemala. *Bulletin of Entomological Research*, *103*(06), 634–643.

World Health Organization (Ed.). (2002). *Control of Chagas disease: Second report of the WHO Expert Committee*. Geneva: WHO.

Yoshioka, K., Nakamura, J., Perez, B., Tercero, D., Perez, L., & Tabaru, Y. (2015). Effectiveness of large-scale Chagas disease vector control program in Nicaragua by residual insecticide spraying against Triatoma dimidiata. *American Journal of Tropical Medicine and Hygiene, 93*(6), 1231–1239.

Zeledón, R., & Rojas, J. C. (2006). Environmental management for the control of Triatoma dimidiata (Latreille, 1811), (Hemiptera: Reduviidae) in Costa Rica: A pilot project. *Memórias Do Instituto Oswaldo Cruz*, *101*(4), 379–386.

Zeledón, R., Rojas, J. C., Urbina, A., Cordero, M., Gamboa, S. H., Lorosa, E. S., & Alfaro, S. (2008). Ecological control of Triatoma dimidiata (Latreille, 1811): Five years after a Costa Rican pilot project. *Memórias Do Instituto Oswaldo Cruz*, *103*(6), 619–621.

Zeledón, R., Zúñiga, A., & Swartzwelder, J. C. (1969). The camouflage of Triatoma dimidiata and the epidemiology of Chagas' disease in Costa Rica. *Boletin Chileno De Parasitologia*, *24*(1), 106–108.

Modeling a Victim-Centered Approach for Detection of Human Trafficking Victims Within Migration Flows

Brant M. Horio[A], Kyle M. Ballard[B]

Disclaimer: The views expressed in this paper are those of the authors and do not necessarily reflect the views or policies of any U.S. government agency.

Abstract

Human trafficking—the use of force, fraud, or coercion to compel individuals into sex trafficking or forced labor—is a global problem. It is challenging to determine the magnitude of this "hidden crime" because the detection of victims is largely dependent on establishing a sufficient level of trust between them and authorities that encourages victims to self-identify. The U.S. government promotes a victim-centered approach for the detection of human trafficking, the implementation of which is becoming widely accepted as the most effective means to detect trafficking victims, especially in scenarios of increased vulnerability such as the European migration crisis. Using the migration crisis and the networks among migrant populations as context, this paper presents human trafficking as a dynamic process whereby a victim's willingness to self-identify adapts over time both spatially (e.g., due to influence from nearby neighbors) and through networks (e.g., familial, country of origin, traveling parties). We employ an agent-based model for exploring the victim-centered approach and its effectiveness for detecting human trafficking victims in an abstract representation of migrant flows. Our sensitivity analysis over a range of initial cooperation levels among migrants finds a tipping point exists for when the victim-centered approach will achieve positive results. It also suggests minimum system conditions for the propagation and persistence of migrants' strategic intention to self-identify. We also discuss our findings with respect to tradeoff considerations for balancing immigration policy objectives, with the need for allowing sufficient time for positive cooperation among the migrant population to develop.

Keywords: agent-based model, human trafficking, migration, social network analysis, preferential attachment network, transnational organized crime

[A] LMI

[B] Office to Monitor and Combat Trafficking in Persons, U.S. Department of State

doi: 10.18278/jpcs.3.2.9

摘要

贩卖人口一使用武力、欺骗或强迫的方式让个体被迫参与性贩卖或强制劳动一是一个全球问题。确定这种“隐藏的犯罪”规模具有挑战性，因为受害者的侦测很大程度上取决于其和权威机构之间建立的信任是否充分，后者需要鼓励前者进行自我认同。美国政府鼓励使用一项以受害者为中心的方法侦测人口贩卖。该方法的实施十分有效，因此正在被广泛接受（例如欧洲移民危机时，人口贩卖的可能性增加，此法尤其在这类场合发挥了重要作用）。以移民危机和流动人口网络为背景，本文将人口贩卖作为一个动态过程进行呈现。在该动态过程中，受害者进行自我认同的意愿会随着时间的推移而产生空间变化（例如受到附近人的影响）和网络变化（例如家庭、出生国家和同行的人）。本文运用一项基于主体的建模（AGENT-BASED MODEL），探索了以受害者为中心的方法和该方法在人口迁移流动中侦测人口贩卖受害者的有效性。针对不同移民初次合作的一系列水平，本文运用敏感性分析发现，需要满足一定条件才能让以受害者为中心的方法实现积极的结果。该分析还建议，应创造最低的系统条件，让移民持续支持并宣传自我认同的战略意图。本文还讨论了有关平衡移民政策目标的取舍问题，并认为移民人口的积极合作需要充足的时间来发展。

关键词：基于主体建模，人口贩卖，移民，社交网络分析，择优连接网络，跨国有组织犯罪

El tráfico de personas— el uso de la fuerza, fraude o coerción para obligar a individuos a ser parte del tráfico sexual o trabajo forzado— es un problema global. Es un reto determinar la magnitud de este “crimen escondido” porque la detección de víctimas depende más que todo de establecer un nivel suficiente de confianza entre ellos y las autoridades que anime a las víctimas a auto identificarse. El gobierno de los EEUU promueve un acercamiento centrado en las víctimas para la detección del tráfico de personas, cuya implementación está siendo generalmente aceptada como una de las formas más efectivas para detectar a las víctimas de tráfico, especialmente en situaciones de alta vulnerabilidad como la crisis migratoria europea. Al utilizar la crisis migratoria y las redes dentro de las poblaciones de migrantes como contexto, este artículo presenta el tráfico de personas como un proceso dinámico en el que la disposición de auto identificarse de una víctima se adapta espacialmente con el tiempo (por ejemplo, debido a la influencia de los vecinos) y a través de las redes (por ejemplo, familiares, de país de origen, de grupos de viaje). Empleamos un modelo basado en agentes para explorar el acercamiento centrado en víctimas y su efectividad para detectar el tráfico de personas en una representación abstracta de flujos migratorios. Nuestro análisis de sensibilidad sobre un rango de niveles de cooperación inicial entre los migrantes encuentra que un punto de quiebre existe cuando el acercamiento centrado en víctimas tiene resultados positivos. También sugiere condiciones mínimas del sistema para la propagación y persistencia de la intención estratégica de los migrantes para auto identificarse. También discutimos nuestros hallazgos respecto a consideraciones de balancear los objetivos de la política de inmigración, con la necesidad de permitir suficiente tiempo para que se desarrolle la cooperación positiva en la población migrante.

Palabras clave: modelo basado en agentes, tráfico de personas, migración, análisis de redes sociales, red de conexiones preferenciales, crimen organizado transnacional

Introduction

Human trafficking is a global problem. The 2016 Trafficking in Persons Report published by the U.S. Department of State (U.S. DOS) documents victims of human trafficking in 188 countries (U.S. DOS, 2016), while the International Labor Organization estimates 21 million victims worldwide in an illicit trafficking economy worth $150 billion per year (International Labor Organization, 2016). Vulnerability to human trafficking increases significantly in situations of mass migration, such as the current European migration crisis, because migrants lack legal status, may be socially marginalized, and may be unaware of local languages or laws. This results in individuals becoming more vulnerable to exploitation by others, including smugglers, unscrupulous labor recruiters, and corrupt border officials, on whom they often rely (U.S. DOS, 2016). The European migration crisis is a large-scale problem; according to the International Organization for Migration (International Organization for Migration, 2015a), nearly 1 million migrants entered the European Union in search of safe haven and economic opportunity in 2015. The migration crisis carries a high risk for exploitation by human traffickers, and effective methods for detecting victims are critically needed. The U.S. government (U.S. DHS, 2016a) promotes a victim-centered approach as the most successful means of detecting human trafficking. Although this approach is viewed as an effective way forward, its implementation challenges are exacerbated in the European migration crisis because of complex interactions with strained government resources, anti-immigration sentiment, and the geopolitical significance of migrant flows.

Combating Human Trafficking

The international legal definition of "trafficking in persons" is the use of force, fraud, or coercion in the exploitation of individuals in commercial sex or forced labor (United Nations OCHCR 2016). Instances of human trafficking are composed of three elements: the acts (e.g., abduction, recruitment, transportation); the *means* (e.g., violence, threats, deception); and the *purpose* (e.g., domestic service, prostitution). Combating human trafficking is a major challenge, even for governments with the political will and adequate capacity. The U.S. Department of Homeland Security describes human trafficking as a hidden crime because victims are often unwilling to self-identify for reasons that include "language barriers, fear of the traffickers, and/or fear of law enforcement" (U.S. DHS 2016b). Instances of human trafficking may be further obscured by other crimes—such as drug and weapons trafficking, prostitution, and immigration violations—which contribute to the significant challenge of detecting trafficking victims. Consequently, successful victim identification requires authorities to investigate beyond initial crimes, such as illegal immigration and prostitution, and recognize indicators of human trafficking to determine whether the perpetrator is actually a victim.

Combating trafficking is further complicated by the fact that trafficking is often a transnational crime, and jurisdictions within the international system operate under different definitions. For example, the Cuban government defines human trafficking as sex trafficking and does not criminalize forced labor (U.S. DOS 2016). In addition, although the vast majority of governments are committed to combating some form of the crime, some governments perpetrate the crime directly or otherwise facilitate or incentivize trafficking (U.S.

DOS 2016). Other governments may be inhibited by a lack of capacity, such as the government of Lebanon, which has an underfunded and untrained law enforcement community that applies its anti-trafficking law "unevenly" (U.S. DOS 2015).

Limitations of Predominate Approaches

There are generally two predominate perspectives for understanding human trafficking (Salt 2000)—an economic perspective that emphasizes "trafficking as a business," and a law enforcement perspective that focuses on trafficking as criminal activity. Although both approaches are important for understanding human trafficking as a phenomenon, neither has proven effective for combating the practice. Both perspectives maintain an underlying assumption that human trafficking is a function of cohesive, coordinated transnational networks. This often leads to approaches that treat human trafficking like other forms of transnational crime, such as trafficking weapons, organs, and drugs (Bhabha, 2006). This is problematic because it effectively dismisses the agency of victims, and implies that strategies for combating the trafficking of inanimate objects are also effective for combating the trafficking of humans. Prosecutions of human trafficking cases often hinge on the cooperation of victims who are frequently the sole source of evidence. The interception of potential trafficking victims does not yield the same evidence as seizing drugs or weapons. Victims must be willing to cooperate to provide the necessary evidence, but are often reluctant to do so because they may be involved in criminal activity themselves (e.g., illegal immigration, prostitution), under duress due to being forced into the situation by their trafficker, fearful of harm or losing their jobs, or in situations of debt bondage in which they or their family face the loss of collateral if they fail to fulfill certain commitments (e.g., employment contracts, debts to labor recruiters).

Victim-Centered Approach to Human Trafficking

The U.S. Department of Health & Human Services (U.S. HHS) (2016) identified two primary reasons victims of human trafficking are not detected; victims do not identify themselves as victims, and others do not view them as such. The U.S. DHS and Department of State both follow and promote the victim-centered approach, which is based on the assumption that victims do not often self-identify and therefore require incentives to cooperate. This approach requires law enforcement knowing the dynamics of human exploitation and then leveraging that understanding to search for indicators of trafficking, such as physical abuse, the use of labor recruiters, and imbalanced labor contracts. Recognizing that victims may have been compelled to commit crimes such as prostitution, authorities using a victim-centered approach are encouraged not to penalize victims for crimes they commit as a result of being trafficked. This approach then makes accommodations for providing victims with needed services, such as shelter, psycho-social help, or repatriation assistance, which act as incentives for self-identification and cooperation with authorities. The overall premise is that, if an individual victim believes authorities are more interested in pursuing human traffickers and helping victims than persecuting her for her coerced crimes, she will be more willing to self-identify and cooperate (U.S. DHS, 2016a).

Human Trafficking in the Context of Migration

Defining Migrant Motivations

The motivations of migrants generally fall into two categories: those seeking *economic* opportunity to improve their quality of life, and those *fleeing* persecution and conflict (International Organization for Migration, 2015b). For this research, we refer to these motivations respectively as "seeking" and "fleeing." Figure 1 shows 2015 migration data that describe the top countries from which European Union asylum applicants originate. The top three origins—Syria, Afghanistan, and Iraq—are countries in active conflict (assumed to have a "fleeing" motivation), while the others suffer from significant poverty as well as political instability (assumed to have a "seeking" motivation). We recognize this is a broad simplification for the complex concept of migrant motivation, but we suggest it is a reasonable abstraction for this research given sparse availability of relevant data.

According to BBC News (2016), Germany received the highest number of 2015 asylum applicants in the European Union, around 36% of all applications. This does not include other forms immigration (e.g., student and work visas) or illegal immigration. Of this population, approximately half of these individuals come from conflict countries—27.2% from Syria, 13.2% from Afghanistan, and 9.5% from Iraq—with the remainder being migrants seeking economic opportunity (Singleton, Heiermann, & Kierans, 2016).

State Immigration Policies

Immigration policies are generally a balance between state economic interests and humanitarian objectives (Facchini & Mayda, 2009). With respect to the former, immigration policies often focus on maximizing the economic benefits of the receiving country. This can take several forms. For example, governments may look to fill gaps in the labor market, at all skill levels, or take on a protectionist posture by limiting

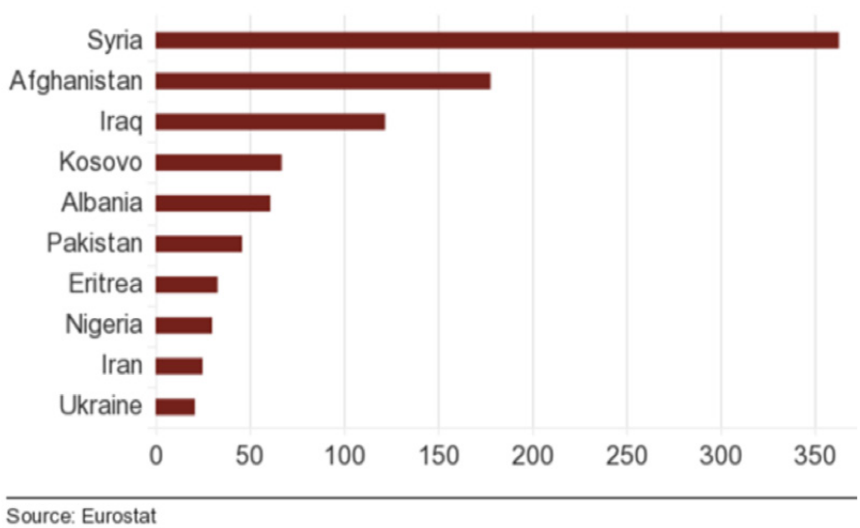


Figure 1. Top 10 origins of people applying for asylum in the European Union. Data show first-time applicants in 2015, in thousands. Source: Eurostat data, BBC News 2016.

migrants' access to labor markets. A 2009 UN Development Programme research paper found that over 80% of surveyed governments had policies to maintain or increase the number of skilled migrants accepted into their countries while five others reported having policies designed to reduce the arrivals of skilled workers (Facchini & Mayda, 2009). With respect to humanitarian objectives, governments may use immigration policy to provide safe haven to individuals in need. Unlike with economic interests, humanitarian programs most often target specific demographics largely determined by a combination of international events, domestic sensitivities, and geopolitics.

In reality, few immigration policies are purely one or the other. Germany, for example, has distinct economic and humanitarian components to its immigration policy. Its 2002 immigration law built a point system by which it considers migrants for admittance. The point system places an emphasis on demand-driven economic factors such that the demographics of admitted migrants are specified by the immediate needs of the labor market; the law overtly favors "skilled workers" and entrepreneurs (Oezcan, 2002). Germany's policy also has a humanitarian element. For example, the Federal Ministry of the Interior issued a reception order in 2014 providing special treatment to refugees fleeing the conflict in Syria (Federal Office for Migration and Refugees 2014). In response to the European migration crisis, while initially declaring Germany a quota-free country, domestic politics and the burden of increasing migrants led President Angela Merkel to shift policy (Wagstyl & Rachman, 2016). In accordance with the European Commission's quota system, Germany established a policy of accepting 27,000 migrants per year, although this number is much larger in practice when one accounts for the various other forms of immigration (BBC News, 2016).

State Immigration Policies

Immigration policies are generally a balance between state economic interests and humanitarian objectives (Facchini & Mayda, 2009). With respect to the former, immigration policies often focus on maximizing the economic benefits of the receiving country. This can take several forms. For example, governments may look to fill gaps in the labor market, at all skill levels, or take on a protectionist posture by limiting migrants' access to labor markets. A 2009 UN Development Programme research paper found that over 80% of surveyed governments had policies to maintain or increase the number of skilled migrants accepted into their countries while five others reported having policies designed to reduce the arrivals of skilled workers (Facchini & Mayda, 2009). With respect to humanitarian objectives, governments may use immigration policy to provide safe haven to individuals in need. Unlike with economic interests, humanitarian programs most often target specific demographics largely determined by a combination of international events, domestic sensitivities, and geopolitics.

In reality, few immigration policies are purely one or the other. Germany, for example, has distinct economic and humanitarian components to its immigration policy. Its 2002 immigration law built a point system by which it considers migrants for admittance. The point system places an emphasis on demand-driven economic factors such that the demographics of admitted migrants are specified by the immediate needs of the labor market; the law overtly favors "skilled workers" and entrepreneurs (Oezcan, 2002). Germany's policy also has

a humanitarian element. For example, the Federal Ministry of the Interior issued a reception order in 2014 providing special treatment to refugees fleeing the conflict in Syria (Federal Office for Migration and Refugees 2014). In response to the European migration crisis, while initially declaring Germany a quota-free country, domestic politics and the burden of increasing migrants led President Angela Merkel to shift policy (Wagstyl & Rachman, 2016). In accordance with the European Commission's quota system, Germany established a policy of accepting 27,000 migrants per year, although this number is much larger in practice when one accounts for the various other forms of immigration (BBC News, 2016).

Approach

Research Question

In the context of the migration crisis, the complex dynamics of stakeholder interactions—shaped by fear, mistrust, vulnerability, and political considerations—have significantly influenced the behavior of state authorities and migrants. This paper introduces an agent-based model that explores the conditions in which a victim-centered approach to human trafficking is implemented in the context of the current European migration crisis. It seeks to answer the following questions, given the scenario of migrants at a country border and assumptions for the communication channels across the migrant population:

- Does a victim-centered approach improve detection of trafficking victims?
- How does a state's immigration policy affect the willingness of human trafficking victims to self-identify (i.e., cooperate) and, in turn, the state's ability to identify those victims?
- Does a tipping point exist with respect to initial cooperation levels at which the victim-centered approach will be effective, given spatial and network influences?

Prior Modeling Approaches

Human trafficking has garnered some attention in the academic and research community, although few computational modeling applications to the problem exist. A notable exception is the extensive systems dynamics model of forced labor human trafficking (Parakh, 2016), although it does not address the vulnerability of migrant populations. A systems dynamics modeling approach also does not address the complexity of the system with respect to interactions between individuals (Amin, 2010). However, the authors have not seen an implementation of such models. Some like Gutierrez-Garcia, Orozco-Aguirre, and Landassuri-Moreno (2013) use agent-based models to explore the emergent qualities of crime by focusing on the conditions and interactions that lead individuals to perpetrate crimes. Pint, Crooks, and Gellar (2010) go a step further to explore the emergence of organized crime and criminal networks. Neither of these efforts addresses victims of crime, and both focus on crimes like gang activity or drug dealing and use that are fundamentally different from human trafficking for reasons noted previously. With respect to migration modeling, some research has been published, but the instances we found were not in the context of human trafficking (Bagherpour, Donaldson, & Scharpnick, 2016). Our research seeks to fill this gap through the modeling and computational exploration of human trafficking, with a particular focus on the vulnerable populations of migrant flows and the role of trust in detecting trafficking victims.

Model

Agent-based models are increasingly recognized as an effective tool for simulating systems that are characterized by complex interactions among their components. The success of the victim-centered approach is strongly dependent on the evolving dynamics of trust between migrants and authorities. We use an agent-based model that accounts for heterogeneous characteristics of migrants who interact spatially and over networks for scenarios of state immigration policies. Our model instantiates a migrant population and seeds the agent set with some number of interspersed trafficking victims. We measure the effectiveness of the victim-centered approach by its detection success of these seeded victims.

The model is a highly stylized and abstract representation of a migrant population. The model is developed on a grid, with each cell holding an agent object that represents migrants. The grid dimensions are 41×41; thus, all model runs are executed with a total migrant population of 1,681. The motivation for using a grid layer was to use the Moore neighborhood (Gilbert & Troitzsch, 2005) as a representation of spatial neighbor influences within the densely packed migrant population on the country borders of a state—an individual is assumed to have greater word-of-mouth communication opportunities with those who are directly adjacent to them. For this research, our model will be in context of the European migration crisis and will attempt to calibrate model parameters to represent a migrant population on the German border.

In the context of migration flows, networks play a key role in the problem of human trafficking. Of particular note are the local kin- and community-based migration networks which migrants rely on to identify potential destinations based on economic opportunity and/or greater safety and security (Salt, 2000). We also wanted to capture network-based communications for these migrant populations (e.g., familial networks, country of origin networks), and we approximate these relationships by linking all of our migrant agents in the population to each other on an underlying preferential attachment network. We assume that the clustering characteristics for this class of networks provide a sufficient proxy in this research to represent real-world migrant networks—some individuals have greater connectivity within the population than others do.

Figure 2 shows a screenshot of our model interface, which features user controls for adjusting population sizes and model parameters for the migrant agents and policies of the state. The underlying preferential attachment network is not visualized. Colors represent the cooperation strategy of each agent—to cooperate (yellow) or deceive (red). For example, from a trafficking victim agent's perspective, we perceive "cooperate" as representing an intention to self-identify, and "deceive" as the intention to not. Shapes indicate if they have interacted with the state yet, and the interaction outcome (accepted or rejected). Using colors and shapes, the researcher may observe a visual representation of the evolution in the system over time and how cooperation may build and/or erode over time. The example realization depicted in Figure 2 shows clustered communities of cooperators persisting among a larger system population of deceivers.

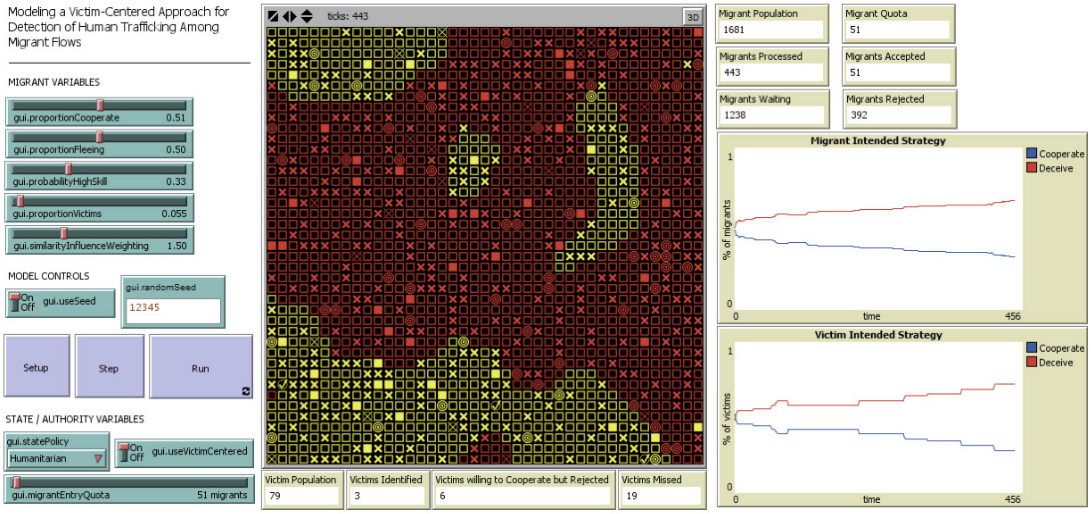


Figure 2. Model screenshot for a sample realization.

Agents

There are two agent types in the model—migrant agents and one state agent.

Migrant Agent

Migrant agents are the primary agent objects of the model and make up the population of individuals, whom we assume for the context of this research as a migrant population congregated awaiting to be processed by the border authorities of the state. We denote a migrant as $m_i \in M$, where M

is the agent set of all migrants in the population.

<u>*Migrant Agent Instance Variables*</u>

Each migrant m_i is heterogeneous and characterized by the instance variables shown in Table 1. Instance variables for s_i, k_i, and v_i were meant to be calibrated to real-world data if possible, with s_i representing whether the migrant is one of our seeded trafficking victims. This may be directly calibrated to existing data. For example, Germany reported that 5.5% of trafficking victims were self-identified in 2015; thus, we may seed a similar proportion of the migrant population as victims that need to be identified. The variable for skill level k_i may be "high" (represents the migrant having a high school degree) or "low" (as not having a degree), and v_i represents the country of origin and the circumstances of that country. Using Figure 1 as an example, we assume that migrants arriving from Syria, Afghanistan, and Iraq are fleeing, while migrants from all other countries of origin are seeking.

Cooperation intention y_i is randomly assigned to each migrant agent during instantiation as per a user-specified parameter for proportion of initial cooperators—agents are preassigned with an intention to cooperate or deceive.

Table 1. Instance variables for the migrant agents

Variable	Variable name	Description
s_i	Victim status	Boolean indicator if the migrant agent m_i is a "seeded" trafficking victim in the population.
k_i	Migrant skill level	The "marketable" skill/education level of migrant agent m_i that would be of interest to the state agent employing an "economic"-based immigration policy. Levels are "high" or "low."
v_i	Migrant motive	The motive of migrant agent m_i for being a migrant, assumed as either "fleeing" (escaping conflict) or "seeking" (seeking economic gain and stability).
y_i	Cooperation intention	The intended strategy of the migrant agent m_i *prior* to interacting with the state agent, to "cooperate" or "deceive."
x_i^S	Interaction status	Boolean indicator if the migrant agent m_i has or has not yet interacted with the state agent.
x_i^O	Interaction outcome	The outcome from the interaction of m_i with the state agent; accepted/granted entry, or rejected/denied entry.

Migrant Agent Behaviors

The migrant agents have two behaviors: (1) interact with the state agent, and (2) update their cooperation strategy to cooperate or deceive. For the interaction with the state, the migrant agent m_i simply engages with the state agent and updates their interaction status x_i^S and outcome x_i^O variables accordingly. The logic for whether m_i is allowed entry past the border is controlled by the state agent.

For the updating of the cooperation intention y_i, this behavior is executed at each tick of the model but handled differently for two different subsets of the migrant population M. Note that only migrants who have not yet interacted with the state agent will go through this process of updating their intention. The cooperation intention y_i is meant to reflect the strategy that the migrant will follow if and only if they are randomly selected to interact with the state agent. An agent will keep updating their cooperation intention for each time step of the model that they are not interacting with the authorities.

We denote the set of migrants waiting to be processed as M^X, where migrants $m_i^X \in M^X$ have a first-order connection to the migrant agent, m_0, who just completed their interaction with the state agent. This is, in effect, an abstraction of the individuals that an agent will relate their recent interaction experience to. We either define a first-order connection as being either a member of a Moore neighborhood or linked via the underlying preferential attachment network to the migrant m_0.

For each $m_i^X M^X$, the agent will count their own first-order connections and imitate the majority strategy of "cooperate" or "deceive." Each agent m_i^X will also compare themselves to m_0 and employ a user-defined similarity multiplier weight to have the cooperation strategy of m_0 count more if m_i^X has a similar attribute that could be perceived as a contributing factor to the m_0 being either accepted or rejected by the state. This was intended to capture an assumption that migrants are aware of outcomes with their connections and may alter their perceptions of cooperation based on an empathy with the recent outcome of m_0. This similarity comparison is based on the specific immigration policy being employed by the state agent (described previously). If the state is using an "economic" policy, migrant m_i counts cooperation strategy y of agent m_0 with a multiplier if m_i has the same skill level k as agent m_0, making m_0 more influential. Similarly, if the state is using a "humanitarian" policy, migrant m_i counts cooperation strategy y of agent m_0 with a multiplier if m_i has the same migrant motive v as agent m_0.

We denote a second set of migrants waiting to be processed—the remainder of the migrant population that has not yet been processed by the authorities—as $\bar{M}^X$, where migrants

$$\bar{m}_i^X \in \bar{M}^X$$

have ***no*** first-order connection to the migrant agent, m_0. Each agent $\bar{m}_i^X$ will count their first-order connections and imitate the majority cooperation strategy y. Note that agent $\bar{m}_i^X$ could include agent m_0 in $\bar{M}^X$, but, as they are not a first-order connection of m_0, no multiplier is used.

State Agent

The state agent interacts with the migrants and determines whether they will be granted entry. Therefore, in this scenario, the state agent may be described as any state authority that makes determinations on whether a migrant agent will be granted entry into the country (e.g., law enforcement, customs, or immigration officials). Our approach assumes a single instantiation of this agent type that interacts with individual migrant agents.

State Agent Instance Variables

The state agent uses the following instance variables shown in Table 2.

The migrant entry quota Q may proportionally represent real-world quota limits for migrant entries granted. For example, Germany agreed to grant up to 27,000 asylum requests, but from an overall population of more than 1 million migrants, which translates to approximately 3% of the migrant flow. In the context of our model, this is a quota limit of 51 migrants, or approximately 3% of the total migrant population.

Table 2. Instance variables for the state agent

Variable	Variable name	Description
Q	Migrant entry quota	Entry limit for how many migrants the state agent will accept, regardless of policy or migrant situation.
P	Immigration policy	"Economic" (maximize the economic value) or "humanitarian" (fill quota and favor fleeing agents and then seal the border).
VCA	Victim-centered approach	Boolean variable for denoting whether the state agent is implementing a victim-centered approach, meaning they will favor migrants with high vulnerability if they are willing to cooperate.

The immigration policy P reflects how the state treats the acceptance process. Immigration policy is a balance between economic and humanitarian elements. If P is an "economic"-focused policy, the state seeks to maximize its economic benefit to the country. This could imply a protectionist posture with a minimization of migrant access to labor markets or the admittance of migrants who fill a specific economic need. For this research, we assume that an economic policy implies the state will accept all migrants with skill level k_i="High" until the quota is met. If P is a "humanitarian"-focused policy, the state seeks to provide safe haven to those in need. For the purposes of this research, we assume that a humanitarian policy implies the state will accept all migrants with motive v_i="Fleeing" until the quota has been met.

The victim-centered approach variable, VCA, denotes whether the state agent will first give preference to migrants of high vulnerability—we assume that seeded victims are migrants of high vulnerability.

State Agent Behaviors

A state agent's only behavior rule relates to its interaction with migrant agents. This interaction is limited to one migrant at a time and is driven by the state's immigration policy, number of migrants allowed entry against the quota, and the individual attributes of the migrant agents. The interaction logic is shown in Figure 3.

Figure 3 also highlights the possible migrant acceptance outcomes. The key success metric we are tracking for our research is when a seeded victim is successfully identified under a victim-centered approach by the state agent. Another valuable metric is when victims are willing to cooperate but the quota has been filled, and thus they are rejected. We intend to use this metric as a proxy for estimating the persistence of cooperation through the migrant population over time.

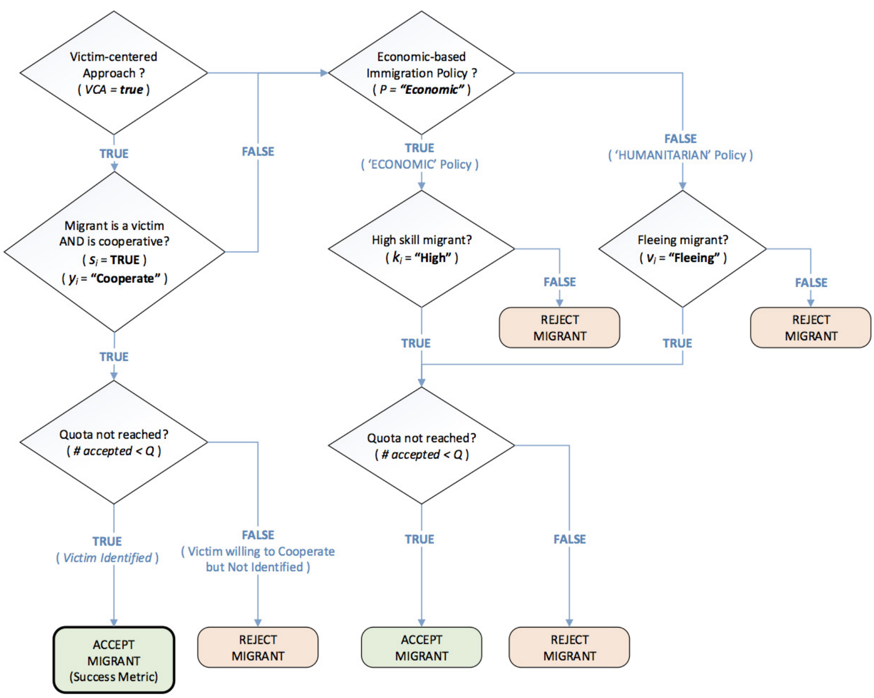


Figure 3. Behavior logic for how the state agent interacts with migrant agents. This logic is used by the state agent for determining whether the migrant will be accepted or rejected.

Process Flow and Scheduling

The process flow of our model is described with the pseudocode below.

```
FOR EACH migrant m_i ∈ M, in a random order
{
  IF the quota Q has not been reached yet
  {
  Interact with state agent and determine accept/reject outcome x_i^O
  Update the cooperation strategy y_i for the 1st order neighbors of m_i
  Update y_i for remaining agents m ∈ M with interaction status x_i^S = FALSE
  Update migrant entries and compare to the quota Q
  } ELSE
  {
  State rejects the migrant m_i
  }
}
```

The process flow will continue until all migrants have interacted with the state agent, and each migrant agent m_i will only interact once per simulation run.

Implementation

Implementation Description

We implemented the model in NetLogo (Wilensky, 1999), an open-source, agent-based modeling platform. All agent updates are executed by default in NetLogo by using asynchronous random activation order.

Verification and Validation

Model verification was a continuous process throughout model development. The primary way to verify that the model was running as expected was through observing how the visualization changed as we manipulated model parameter.

To help with this process, we implemented agent visual representations to clearly indicate where in the process they were. These representations are shown in Figure 4.

Figure 5 shows a panel plot for one sample realization of our model. Each plot shows the emergent pattern of migrant cooperation strategies for the given time tick t in the model. We observe that the migrant quota was reached at time *t*=51, but clustered communities of cooperation persist well into tick 500.

Following Axtell and Epstein (1994), we suggest we have achieved Level 1 validation—our model can represent stylized facts for how influence and social contagion might plausibly spread through a migrant population, but is not at a level of realism that we can achieve statistical validity. This level of validation, while not good

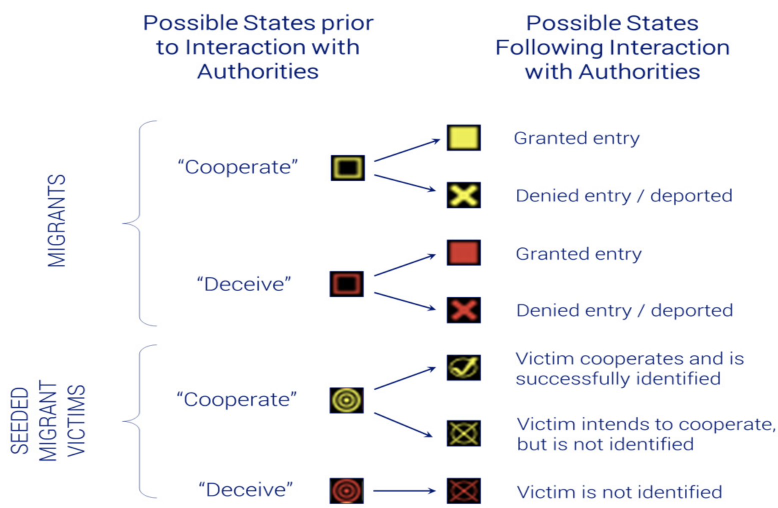


Figure 4. Visualization legend key for the model. For each agent type of a migrant or a seeded trafficking victim, the key shows the initial visualization for their intended strategy and the possible new state(s).

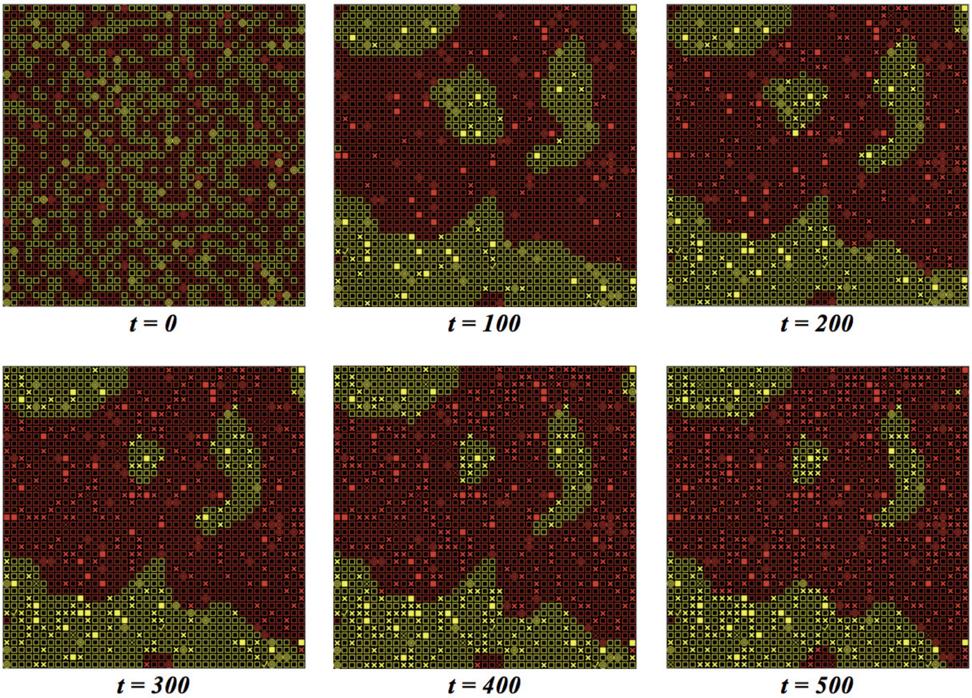


Figure 5. Sample model realization with system state changing over time. Shows model screenshots over time, with the variable *t* representing the tick count.

for point prediction estimates, is still useful to evaluate potential mechanistic explanations for the role of interaction dynamics between migrants and immigration policy on the potential high-level outcomes for detection of trafficking victims.

Experiment Results

We conducted a parameter sweep experiment aimed to determine the sensitivity of outcomes (i.e., successful detection of migrant victims) to the initial levels of cooperation in the population.

Assumptions

We made the following assumptions for the experiments:

- Rules around accepting migrants will always reflect self-interest and domestic politics; a victim-centered approach will be secondary to immediate interests.
- The number of migrants a country is willing to accept is bounded by a quota.
- Under an economic policy, acceptance is based on whether a migrant's specific skill level matching the state's stated need.
- Under a humanitarian policy, acceptance is based on migrant motive (i.e., fleeing or seeking).
- Under a victim-centered approach, all victims are accepted if they are willing to cooperate.
- A victim may be granted entry based on the immigration policy, but if she does not want to "cooperate," she is still counted as an undetected victim.

Experiment Initialization

To initialize the experiment runs, we used the following parameter ranges in Table 3 that reflect real-world representations (where possible) of the migration crisis on Germany's border.

We ran 100 replications for each parameter combination specified in Table 3, for a total of 1,100 runs, to complete the experiments.

Results

Figure 6 shows that the detection of trafficking victims first emerges when 48% of the agent population starts with a cooperative strategy. Detection rates increase with the percentage of agents using our abstraction of cooperative strategies, stabilizing around 52%. Figure 7 measures the number of seeded victims who were denied entry, despite being willing to cooperate. Because the willingness to cooperate is a function of interactions with other migrants, the results also provide an indication for how an overall culture of cooperation persists throughout the migrant population despite a country meeting its quota and closing its border. Results suggest that when initial cooperation intention of the population is >52%, nearly all seeded victims maintain a willingness to cooperate.

These results suggest initial treatment of trafficking victims has a significant influence on the success for victim detection, and that early propagation of a readiness to cooperate is important for sustaining victims' long-term willingness to cooperate. The more effectively the state can promote cooperation prior to interactions with state agents, the greater the contagion and persistence of cooperation in the population over time. These insights have implications for how states should proactively shape migrant perceptions and encourage self-identification. These findings can serve as a basis for further study of how state behavior influences individual strategies for interacting with authorities.

In addition to a computational exploration for the underlying mechanics of the victim-centered approach, model outcomes suggest for our abstraction of the problem, a victim-centered approach facilitates the identification of victims. Both the economic and humanitarian immigration policies performed better when paired with the victim-centered approach.

Table 3. Experiment parameter settings to represent migration into Germany in 2015

Model parameter	Experiment Value(s)
Population	1,681 (fixed)
Proportion of initial cooperators	Over range [0.45, 0.55], incrementing by 0.01 (initial explorations of the data suggested this was a parameter range of interest).
Proportion of fleeing migrants	0.50 (represents proportion from Syria, Afghanistan, and Iraq)
Proportion of high-skilled migrants	0.33 (represents proportion with a high school degree)
Proportion of trafficking victims	0.055 (represents documented self-identification rate)
State policy	Economic, humanitarian
Victim-centered approach	True, Ffalse
Migrant entry quota	51 Migrants (represents 3% of the total population)

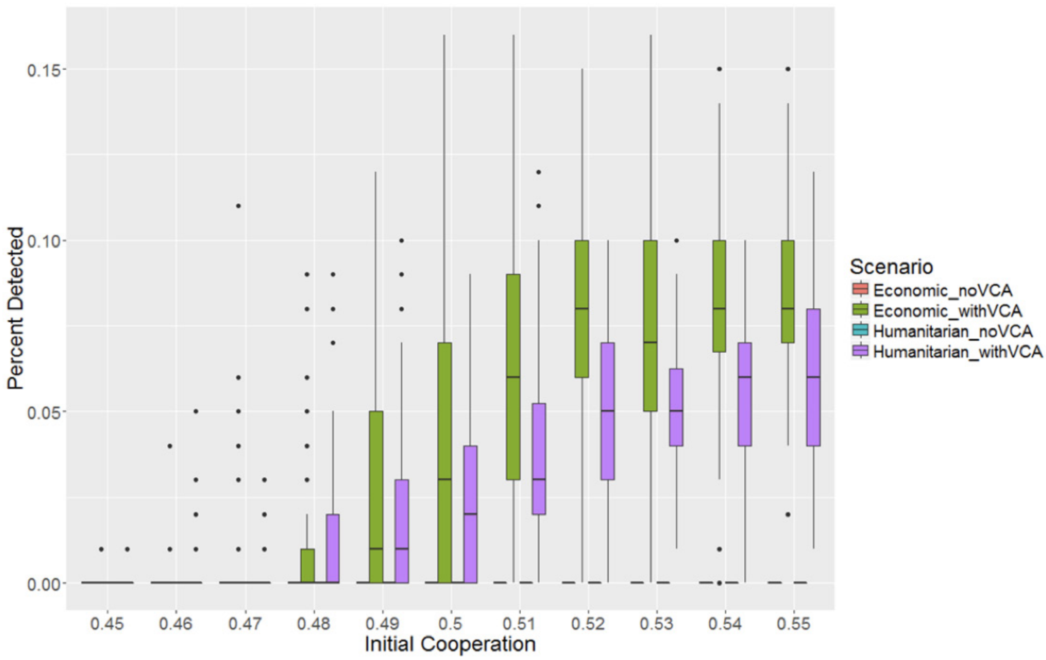


Figure 6. Percent of human trafficking victims successfully detected. The plot shows the percent of the migrant population that were willing to cooperate (and the quota was not yet full) when the state agent interacted with them.

The outputs suggest however that the economic immigration policy yields more cooperation among trafficking victims and a slightly improved rate of detecting those victims. A satisfactory explanation of this outcome is still being investigated. Previously cited studies on international migration offer some theoretical basis for validation. For example, Kugler, Boussalis, and Coan (2012) suggest that migrants prioritize the opportunity for economic success over access to public goods.

This may also suggest a tradeoff between more permissive immigration policies (represented in this model by the humanitarian policy) and a victim-centered approach to human trafficking. In other words, policies that admit migrants based on broad or commonly occurring attributes (such as humanitarian suffering) might lead states to more quickly fill their immigration quotas and deplete resources, while more selective immigration policies might afford state authorities more time and resources to successfully employ a victim-centered approach.

Conclusions

Our research aimed to provide insight into some of the tradeoffs governments must balance in the context of large migration flows. We developed and presented a proof-of-concept model to better understand of how a state's immigration policy might affect the willingness of human trafficking victims to self-identify (i.e., cooperate) and, in turn, the state's ability to identify victims among the flow of migrants. Ours is a parsimonious model that is focused on representing a population environment (i.e., migrants) where agents are sensitive to spatial and network-based communication, adaptation, and learning.

Our experiments identified a tipping point of initial cooperation levels at which

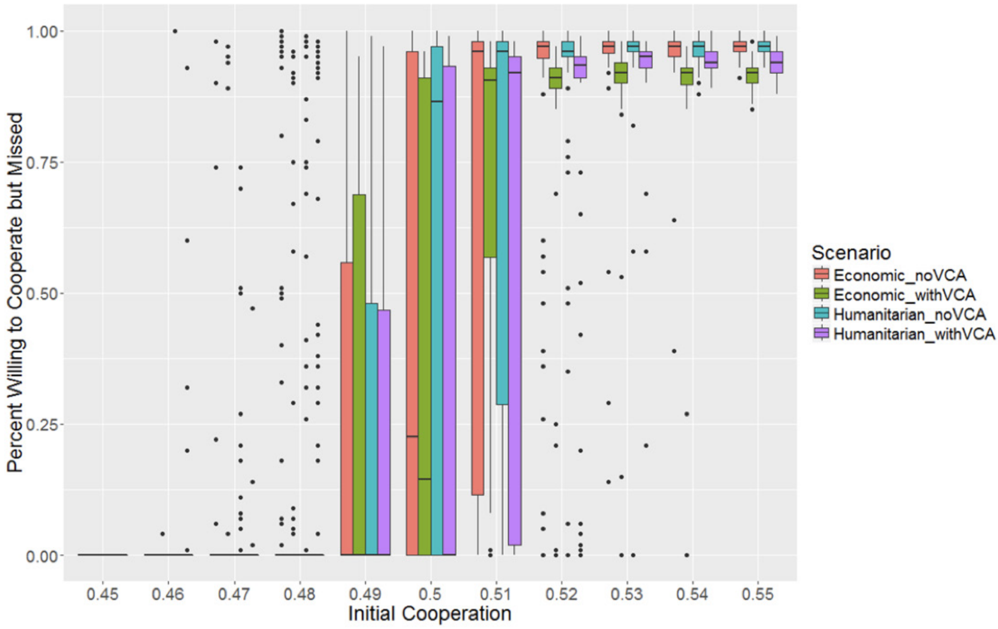


Figure 7. Percent of human trafficking victims that were willing to cooperate but were not identified under the victim-centered approach.

the victim-centered approach became effective, given spatial and network influences. The model, as in the real-world, is sensitive to the initial cooperation of the victim population. Our results imply that modest success in the model requires a very high level of initial cooperation, nearly half of the population. This, in turn, suggests victim experiences and perceptions before interaction with authorities are very important because those experiences and perceptions lead to persistent clusters of migrants willing to cooperate even after border closure. Results also suggest state authorities need early successes for cooperative strategies to propagate across the population.

Initial findings also suggest a victim-centered approach coupled with more selective immigration policies appear to yield the most success in detecting human trafficking victims. In this model, the economic immigration policy was the more selective of the two. More research is required to determine if more selective policies are indeed more conducive to the victim-centered approach. The results are not an indictment on humanitarian policies or a promotion of economic policies, per se, but rather it may be an indication that overarching goals of a state's immigration policy—and a state's subsequent allocation of resources and efforts—have a direct (and counterintuitive) impact on whether a victim-centered approach is successful. In other words, we suggest that less selective policies allow the migrant quota to be filled so quickly that trust and cooperation levels necessary for the victim-centered approach to be successful do not have time to propagate and influence victim beliefs. This tradeoff between policy goals and resources is exacerbated by mass migration flows, and emphasizes even more so, the critical role of initial cooperation levels and perceptions of the state among target populations.

Future Research

This agent-based model is an exploratory model that requires more development and validation to empirical data sets. To be more nuanced and more useful for policymakers, the parameters of the model could be expanded to account for more individual details among migrant populations and government policies. For example, more detail for describing migrant skill levels, more nuanced state policies (e.g., hybridization between economic and humanitarian policies), and a better definition and representation for migrant "vulnerability." These expanded data elements could facilitate a more robust consideration of migrant motivations and enhance the model representation for government efforts and the resulting ability to detect potential trafficking victims. Greater fidelity in migrant attributes is necessary for a more realistic representation of the victim-centered approach.

Furthermore, several parameters in future iterations will be calibrated to existing data. For example, the current model imposes government immigration quotas. Implementing more precise representations and estimates of Germany's migrant acceptance policies, rates, and limits may yield more representative insights. Similarly, the networks used to propagate perceptions across the agent population are highly stylized. These networks should be explored for a representation more realistic than assuming a preferential attachment network model.

Adding additional jurisdictions would also help us better understand the transnational elements of migration and human trafficking. Representing other governments would allow for a more complete exploration of the various factors that drive migrant perceptions (e.g., experiences in source and transit countries) and recreate the more realistic complexities actual migrants face as they choose where to settle. Geopolitical dynamics could also be more fruitfully explored, such as the effects of closed-border policies in countries like Hungary on more permissive policies, such as those in Germany.

Acknowledgments

The authors would like to express their appreciation for feedback on this research from Dr. Rob Axtell and Dr. Qing Tian from the Computational Social Science Program of the College of Science at George Mason University. We would also like to thank the anonymous reviewers for their constructive feedback.

References

Amin, S. (2010). *A step towards modeling and destabilizing human trafficking networks using machine learning methods*. Paper presentation, AAAI Spring Symposium Series, Palo Alto, CA, March 2010. Retrieved from http://www.aaai.org/ocs/index.php/SSS/SSS10/paper/view/1155/1341

Axtell, R. L., & Epstein, J. M. (1994). Agent-based modeling: Understanding our creations. *The Bulletin of the Santa Fe Institute*, 28–32.

Bagherpour, A., Donaldson, S., & Scharpnick, M. (2016). Europe at a crossroads: Pathways to mitigating the refugee crisis in 2016. Washington, DC: Global Impact Strategies. Retrieved from https://www.gistrat.com/wp-content/uploads/2016/02/giStrat_EU_Migration_Feb_2016.pdf

BBC News (2016, March 4). *Migrant crisis: Migration to Europe explained in seven charts.* Retrieved from http://www.bbc.com/news/world-europe-34131911

Bhabha, J. (2006). *Human smuggling, migration and human rights*. Working paper for the International Council on Human Rights Policy Review Meeting, "Migration: Human Rights Protection of Smuggled Persons," Geneva, Switzerland, July 2006.

Facchini, G., & Mayda, A. M. (2009). *The political economy of immigration policy* (Human Development Research Paper 2009/03). New York: United Nations Development Programme. Retrieved from http://hdr.undp.org/sites/default/files/hdrp_2009_03.pdf

Federal Office for Migration and Refugees. (2014). *The reception procedure for Syrian refugees.* Retrieved from http://www.bamf.de/EN/Migration/AufnahmeSyrien/aufnahmeverfahren-syrien-node.html

Gilbert, N., & Troitzsch, K. J. (2005). *Simulation for the social scientist* (2nd ed.). Maidenhead, UK: Open University Press.

Gutierrez-Garcia, J. O., Orozco-Aguirre, H., & Landassuri-Moreno, V (2013, November). A*gent-based simulation of crime*. Paper presentation, 12th Mexican International Conference on Artificial Intelligence, Mexico City, Mexico. Retrieved from http://ieeexplore.ieee.org/document/6714643

International Labor Organization. (2016). *Forced labour, human trafficking and slavery.* Retrieved from http://www.ilo.org/global/topics/forced-labour/lang--en/index.htm

International Organization for Migration. (2015a). *Irregular migrant, refugee arrivals in Europe top one million in 2015: IOM*. News release. Retrieved from https://www.iom.int/news/irregular-migrant-refugee-arrivals-europe-top-one-million-2015-iom

International Organization for Migration. (2015b). *World migration Report 2015—Migrants and cities: New partnerships to manage mobility.* Geneva: International Organization for Migration. Retrieved from http://publications.iom.int/system/files/wmr2015_en.pdf

Kugler, T., Boussalis, C., & Coan, T. G. (2012). Political performance as a moderator of international migration. In J. Kugler & R. L. Tammen (Eds.), *The performance of nations* (pp. 119–143). Lanham, MD: Rowman & Littlefield Publishers, Inc.

Oezcan, V. (2002). *German immigration law clears final hurdle*. Migration Policy Institute. Retrieved from http://www.migrationpolicy.org/article/german-immigration-law-clears-final-hurdle

Parakh, S. (2016). *Modeling trafficking of OFWs*. Presentation. Retrieved from https://saurabhparakh.kumu.io/modeling-trafficking-of-ofws?token=1JSNlfq4I7vxKjxQ

Pint, B., Crooks, A., & Geller, A. (2010). *Exploring the emergence of organized crime in Rio de Janeiro: An agent-based modeling approach.* Paper presentation, 2010 Second Brazilian Workshop on Social Simulation (BWSS), Sao Paulo, Brazil. Retrieved from http://ieeexplore.ieee.org/document/6030007

Salt, J. (2000). Trafficking and human smuggling: A European perspective. I*nternational Migration*, Special Issue 2000/1, 31–56.

Singleton, A., Heiermann, A. C., & Kierans, D. (2016). *Migration, asylum and refugees in Germany: Understanding the data.* Global Migration Data Analysis Center. Retrieved from https://www.iom.int/sites/default/files/country/docs/Germany/Germany-Data-Briefing-1Jan2016.pdf

United Nations OCHCR. (2016). *Protocol to prevent, suppress and punish trafficking in persons especially women and children, supplementing the United Nations Convention against Transnational Organized Crime.* Retrieved from http://www.ohchr.org/EN/ProfessionalInterest/Pages/ProtocolTraffickingInPersons.aspx

U.S. Department of Health and Human Services. (2016). *Identifying victims of human trafficking: Inherent challenges and promising strategies from the field.* Retrieved from https://aspe.hhs.gov/basic-report/identifying-victims-human-trafficking-inherent-challenges-and-promising-strategies-field

U.S. Department of Homeland Security. (2016a). *A victim-centered approach.* Retrieved from https://www.dhs.gov/blue-campaign/victim-centered-approach

U.S. Department of Homeland Security. (2016b). *What is human trafficking?* Retrieved from https://www.dhs.gov/blue-campaign/what-human-trafficking

U.S. Department of State. (2016). *Trafficking in persons report.* Washington, DC: U.S. Department of State. Retrieved from http://www.state.gov/documents/organization/258876.pdf

Wagstyl, S., & Rachman, G. (2016, January 20). German President Backs Refugee Quotas as Austria Sets First Cap. *Financial Times.* Retrieved from https://next.ft.com/content/14eab1c2-bf6d-11e5-846f-79b0e3d20eaf

Wilensky, U. (1999). *NetLogo.* Evanston, IL: Center for Connected Learning and Computer-Based Modeling, Northwestern University. Retrieved from http://ccl.northwestern.edu/netlogo/hubnet.html

Simulating Heterogeneous Farmer Behaviors under Different Policy Schemes: Integrating Economic Experiments and Agent-Based Modeling

Shang Wu,[A] Asim Zia,[B] Mengyuan Ren,[C] Kent Messer[D]

Abstract

In this paper, we develop an agent-based model that scales up results from economic experiments on technology diffusion and abatement of non-point source water pollution under the conditions of an actual watershed. The results from the economic experiments provide the foundation for assumptions used in the agent-based model. Data from geographic information systems and the US Census of Agriculture initialize and parameterize the model. This integrated model enables the exploration of the effects of several policy interventions on technology diffusion and agricultural production and, hence, on agricultural non-point source pollution. Simulation results demonstrate that information "nudges" based on social comparisons increase ambient-based policy performance as well as efficiency, especially individual-level tailored information on what others like them have done in past similar situations.

Keywords: economic experiments, agent-based model, non-point source pollution, policy evaluation, farmer behavior

[A] Shang Wu is a Ph.D. candidate in Economics and research assistant in Center for Experimental and Applied Economics at the University of Delaware.

[B] Asim Zia is an Associate Professor of Public Policy and Decision Analysis in the Department of Community Development and Applied Economics, with a secondary appointment in the Department of Computer Science, at the University of Vermont.

[C] Mengyuan Ren holds a M.S. in Computer Science from the University of Delaware.

[D] Kent Messer is Unidel Howard Cosgrove Chair for the Environment and Professor in Applied Economics and Statistics at the University of Delaware.

doi: 10.18278/jpcs.3.2.10

摘要

本文运用一项基于主体的建模（agent-based model， 简称ABM），扩大了关于“技术扩散”和“减少实际流域下非点源水污染”的经济实验结果。该结果为ABM中应用的猜想提供了基础。该模型的发起和参数化设置都是由地理信息系统和美国农业普查（Census of Agriculture ）的数据得出。整合后的模型能探索好几种政策干预对技术扩散和农业产量的影响，进而探索对农业非点源污染的影响。模拟结果证明，基于社会比较（social comparison）的信息“助推” （information nudges）能够增加基于环境（ambient based）的政策表现，同时还能提高效率。特别地，在以往类似情况下相似的他人行为，能对个人产生信息助推作用。

关键词：经济实验，基于主体建模，非点源污染，政策评估，农户行为

Resumen

En este artículo desarrollamos un modelo basado en agentes que amplía los resultados de experimentos económicos en el tema de la difusión de la tecnología y reducción de la polución difusa de agua bajo las condiciones de una cuenca verdadera. Los resultados de los experimentos económicos proveen las bases para las suposiciones utilizadas en el modelo basado en agentes. Los datos de los sistemas de información geográfica y el Censo de Agricultura de EEUU inicializan y parametrizan el modelo. Este modelo integrado habilita la exploración de los efectos de varias intervenciones políticas en la difusión de la tecnología y la producción agrícola y por consiguiente en la polución difusa de agua no agrícola. Los resultados del simulacro demuestran que los “codazos” de la información basada en comparaciones sociales incrementan el rendimiento de la política con base ambiental, así como la eficiencia, especialmente información adaptada a nivel individual sobre lo que otros como ellos han hecho en situaciones similares en el pasado.

Palabras clave: experimentos económicos, modelo basado en agentes, polución difusa, evaluación de políticas, comportamiento de agricultores

Introduction and Literature Review

Non-point source (NPS) pollution in water systems mainly comes from rainfall and snowmelt that move over and through the ground, bringing natural and human-made pollutants into waterbodies. NPS pollution, which comes mostly from nutrients and chemicals carried by agricultural runoff, is the primary cause of water pollution in the United States today. Unfortunately, regulation and remediation of NPS water pollution is a difficult task. It typically is hard and at times impossible to identify individual contributors to such pollution, and policies designed to address it must be designed to take polluters' hidden actions and asymmetric information into account. The cost of this type of individual monitoring and enforcement is often prohibitive (Xepapadeas, 2011).

Theoretical work (Segerson, 1988; Xepapadeas, 1992) has shown that policies based on ambient levels of pollution can lead to reductions of NPS pollution to a regulator-specified target level. However, since no program has implemented an ambient-pollution-based policy on a large scale to provide empirical data, researchers have often turned to economic experiment laboratory settings as test beds for such policies (Miao et al., 2016; Poe, Schulze, Segerson, Suter, & Vossler, 2004; Spraggon, 2002; Suter, Vossler, & Poe, 2009). In addition, since researchers must recruit and compensate participants in economic experiments, the experiments generally have been limited in scale and have restricted the ability to draw conclusions in contexts outside the laboratory. Thus, researchers have been interested in finding other ways to study the effects of these policies as part of efforts to improve their outcomes in terms of reducing NPS pollution.

Agent-based modeling (ABM) can help fill this gap by providing a mechanism for scaling up the findings in experiments to contexts that are closer to reality. With ABM, researchers can use findings from an experiment to create model agents that behave according to patterns identified in the experiment, and conduct simulations using an environment that better mimics a real-world setting. ABM also allows the researcher to observe the results of those agent interactions, which are extremely difficult to capture using other methods. Furthermore, we compared to traditional top-bottom methods such as econometric techniques, ABM imposes less distributional restrictions or assumptions.

ABM has been applied in various fields in recent years (Farmer & Foley, 2009), such as ecological modeling (Grimm & Railsback, 2005), population growth (Axtell et al., 2002), business strategies (Khouja, Hadzikadic, & Zaffar, 2008), land-use policy (Tsai et al., 2015), transportation policy (Zia & Koliba, 2015), and education (Johnson, Lemasters, & Bhattacharyya, 2017). In the context of agricultural and environmental applications, it has been used mainly for problems associated with changes in land cover to develop models that simulate land-use decisions by farmers facing multiple constraints (Matthews, Gilbert, Roach, Polhill, & Gotts, 2007; Veldkamp & Verburg, 2004), especially in studying coupled human and natural systems (An, 2012). In such systems, agent decisions generate environmental consequences, which could in turn affect human decisions and behavior. Recently, Tesfatsion, Rehmann, Cardoso, Jie, and Gutowski (2017) developed the Water and Climate Change Watershed (WACCShed) platform that allows the systematic study of interactions of hydrology, human, and climate in a watershed over time. Ng, Eheart, Cai, and Braden (2011) demonstrates an agent-based model of farmer decision-making on water quality in the context of first and second-generation biofuel crops and carbon trading. The ABM integrates a SWAT-based hydrologic–agronomic model.

In the bottom-up construction of an ABM, modelers need to assign decision rules to agents under specific scenarios. A major challenge lies in constructing credible decision rules for ABM (Zenobia, Weber, & Daim, 2009). Most of the previous work usually assumes perfect rationality, meaning that the agents could perfectly solve for utility maximizing problems in various and complex scenarios. However, behavioral economics have repeatedly shown that human behavior is often rationally bounded at best and that individuals often use heuristics instead of optimization when making decisions. As noted by Heckbert, Baynes, and Reeson (2010), combining economic experiments with ABM offers researchers many new opportunities. Experimental economics can be used to guide calibration of ABM so that the agents' behaviors and decisions reflect patterns identified by actions in experiments.

Some researchers have used survey methods to develop decision rules for ABM (Dia, 2002). Compared to using survey-based approaches to calibrate decision-making in ABM, we can use data collected through experiments that capture the "interpersonal" and "interplayer" dynamics that arise in experimental games (and are overlooked by surveys). Furthermore, Duffy (2006) pointed out that ABM projects also could facilitate researchers' ability to interpret the aggregate findings of an experiment involving human subjects.

Not many studies have combined experimental economics and ABM. Evans, Sun, and Kelley (2006) compared results from a spatially explicit laboratory experiment to outputs of a simulation from a land-use ABM involving utility-maximizing agents. They concluded that the participants in the experiment deviated from revenue-maximizing actions and that it was thus valuable to use non-maximizing agents in ABM. Heckbert (2009) also acknowledged the value of combining experiments and ABM, reporting a study in which a participant replaces the role of an agent and the participant's behavior under several treatments can be used to recalibrate the ABM. A few studies have attempted to integrate economic experiments and ABM in NPS pollution management context. Zia et al. (2016a) constructed agent-based models using an economic experiment documented in Miao et al. (2016). The agents were categorized to pursue different behavioral strategies under alternate policy and sensor information regimes, and a multi-level multinomial logistic regression model built from experimental data predicted the agents' type categories. Our research extends this idea by designing an experimental setting that includes technology adoption decisions and two layers of heterogeneity, meanwhile building a closer link between the experiment and the ABM.

We also include two information treatments to examine the ability of information "nudges" to induce desired outcomes from the participants. Originating from the social comparison theory by Festinger (1954), it has been shown that information "nudges" on social comparison and peer actions can promote environmental conservation behavior (Allcott, 2011; Ferraro & Price, 2013; Goldstein, Cialdini, & Griskevicius, 2008). These information "nudges" are attractive from a policy design perspective since they are more cost-effective compared with traditional monetary-based programs. However, not much research has considered incorporating information "nudges" in NPS pollution management. We are interested in if information "nudges" based on social comparison and peer action could help the performance of ambient-based policies. In the first information treatment, participants are provided with information about what people "like them" have chosen in a similar situation in the past. In the second treatment, their group in the preceding round provides participants with information regarding average production and average rate of adoption of technology. Participants' responses to the policy and the information treatments given the heterogeneity of production types are used to guide the agent's behavior in the models under various scenarios.

In this study, we scale up findings from an economic experiment with ABM in a spatially explicit watershed setting to provide insight into the effects of different policy interventions addressing NPS pollution. The models capture interactions among heterogeneous agents in terms of diffusion of technology adoption by farmers, which is difficult to model using other techniques. Specifically, we test how tax/subsidy policies based on ambient levels of water pollution work in scenarios involving heterogeneous production and pollution schemes and focus on cases in which the decision space of the agents is extended from making a single production decision to making a production and a technology decision. We also investigate how information influences people's behavior and whether policies can be designed to incorporate information "nudges" to induce more-desired outcomes. Our study contributes to the literature in two main aspects. In environmental and resource economics, our experiment investigates the effect of information "nudges" in an experimental setting that simultaneously incorporates an extended participant decision space and multiple layers of heterogeneity. Moreover, we use an ABM that features heterogeneous agents in a spatially explicit context to understand implications of the complex actions and interactions created based on experimental data. In the field of ABM, despite rising interest in using non-fully rational agents, not much work has actually done so. We are one of the first to introduce bounded rational agents into an ABM based on an economic experiment. The ABM agent decision rules are closely linked with human decisions in the economic experiment using an underlying game-theoretical model. Our research demonstrates that economic experiments can be useful to capture bounded rationality and guide ABM development. This study provides an example to incorporate human-based decision rules and a possible framework to integrate experiments and ABM in future research.

Experimental Design and Theoretical Foundation

In this part, we discuss the experimental design of our economic experiment. We first lay out the theoretical model, and describe the treatments in the experiment.

We build upon and extend the classic model framework in the environmental economics literature. Consider a group of agricultural producers indexed by $i = 1 \dots N$ operate farms or ranches adjunct to a common watershed. The farmers' operations generate pollution as byproduct. A regulator monitors water quality by a sensor at the downstream of the watershed. The farms may differ in both their capacity and their distance to the sensor. The farmers may choose to adopt a pollution abatement technology (e.g., buffer, cover crop) at a cost (τ) proportional to farm size. Each year, the farmers make two decisions:

a production decision x_i and a decision on whether to adopt an abatement technology a_i, $(\partial PE_i(x_i,a_i))/(\partial x_i)>0$, $(\partial PE_i(x_i,a_i))/(\partial a_i)<0$, indicating lower production and the adoption of the technology are associated with lower private earnings through $PEi(x_i,a_i)$. The environmental damage generated by each farm is $Di(x_i,a_i)=\alpha\beta_i x_i a_i + \beta_i x_i(1-a_i)$, where $(\partial PE_i(x_i,a_i))/(\partial x_i)>0$, $(\partial PE_i(x_i,a_i))/(\partial a_i)<0$, and β_i depends on the location of the farm relative to the sensor and α denotes the effect of the technology. We assume that the total environmental damage is $TD=\sum_1^N D_i(x_i,a_i)$. Without any regulation, a profit-maximizing farm will produce at their capacity level and not adopt the technology. The social planner's problem is to maximize social benefits (denoted as SP), where $SP=\sum_{i=1}^N PE_i(x_i,a_i)-\sum_{i=1}^N D_i(x_i,a_i)$. Suppose the regulator hopes to achieve a pollution standard D and imposes a tax/subsidy policy, where the tax/subsidy equals to the environmental damage minus the target level of pollution, $t(TD)=(TD-\bar{D})$. Following the literature, suppose $PE_i(x_i,a_i)$ takes a quadratic form $\gamma_0-\gamma_1(\gamma_{2i}-x_i)^2-\tau\gamma_{2i}a_i$, where $\tau\gamma_{2i}a_i$ takes into account whether the firm adopted the technology. Now the individual payoff function under the tax/subsidy scheme becomes: $\pi_i=PE_i(x_i,a_i)-(TD-\bar{D})$. We find the Nash strategy by backward induction. Consider firm i, given the pollution level of others in the group D_{-i}, its profit function from producing x_i and adopting the technology is:

$\pi_i^A=\gamma_0-\frac{(\beta_i\alpha)^2}{4\gamma_1}-\left(D_{-i}-\bar{D}+\beta_i\alpha\gamma_{2i}-\frac{\beta_i^2\alpha^2}{2\gamma_1}\right)-\tau\gamma_{2i}$, taking first order condition, the maximum is reached at $x_i^A=\gamma_{2i}-\frac{\beta_i\alpha}{2\gamma_1}$. The profit for not adopting the technology is $\pi_i^N=\gamma_0-\frac{\beta_i^2}{4\gamma_1}-\left(D_{-i}-\bar{D}+\beta_i\gamma_{2i}-\frac{\beta_i^2}{2\gamma_1}\right)$, and the maximum can be reached by producing $x_i^N=\gamma_{2i}-\frac{\beta_i}{2\gamma_1}$. The condition for a farmer to prefer to adopt compared with not adopt is therefore $C=\pi_i^N-\pi_i^A=\frac{\beta_i^2}{4\gamma_1}(1-\alpha^2)-\beta_i\gamma_{2i}(1-\alpha)+\tau\gamma_{2i}<0$. Thus, a unique dominant Nash strategy for a farm is defined as $\{C<0: x_i=\gamma_{2i}-\frac{\beta_i\alpha}{2\gamma_1}, a_i=1;\ C\geq 0: x_i=\gamma_{2i}-\frac{\beta_i}{2\gamma_1}, a_i=0\}$. This dominant Nash strategy is also the same as the social planner's optimal strategy.

Treatments

We consider two dimensions of treatments. On the within-subject level, we varied whether the tax/subsidy policy is in place and the complexity of heterogeneity that is in the experiment. For each of the policy treatment, we conducted four heterogeneity treatments, namely,

1. A homogeneous treatment where the locational impact on water quality and size of each farm is the same (Homo);
2. A first heterogeneous treatment where the locational impact on water quality vary, but the size of each farm is the same (Hetero1);
3. A second heterogeneous treatment where the size of the farms vary, but locational impact on water quality is the same (Hetero2);
4. A third heterogeneous treatment where both size and locational impact on water quality of farms vary (Hetero3).

To control for potential order effects, we randomly varied the order of the within-subject treatments that are presented. On the between-subject level, we provided participants with three information treatments. No Info serves as the baseline. In the Info1 treatment, we provide testimonial information on what production and technology adoption decisions people "like them" have made in the past. The information comes from the "no information" treatments. We find true decisions participants made that are closest to the Nash optimal strategies conditioning on their size and location. Therefore, this information differs by the location and the size of the firm and approximates the actual Nash optimal strategies. This resembles some policy recommendation on what people should consider doing based on their location and size. In the information treatment 2, we give participants information on the technology adoption rate and average production in their group in the last decision. This is similar to a policy that provides information on what others in the neighborhood are doing and has a self-evolving nature. Since each decision is independent and each participant has a unique dominant Nash strategy, theoretically the information treatments should not change participants' decisions. However, as noted before, human decisions often demonstrate bounded rationality and may follow simple heuristics or ad hoc rules.

Experiment Procedure

The economic experiment consists of 12 sessions conducted in late 2016, involving 192 participants recruited at a large public university in the northeastern United States.

Agent-based Model Setup

In this part, we discuss the ABM setup and initialization. We design the ABM to capture key elements of the economic experiment and an actual watershed while avoiding including unnecessary assumptions and processes. We first set the ABM to a spatially explicit context based on the Murderkill watershed located in the southeast part of Kent County, Delaware (Figure 1). The Murderkill[1] watershed is chosen mainly because it consists primarily agricultural land use and it is a typical coastal plain. Besides, it has promulgated TMDL regulations and has research efforts on the estuarine portion of the watershed. Moreover, the watershed is comprised of 68,000 acres of land, which is large enough to generate meaningful conclusions, but not too large to create computational obstacles.

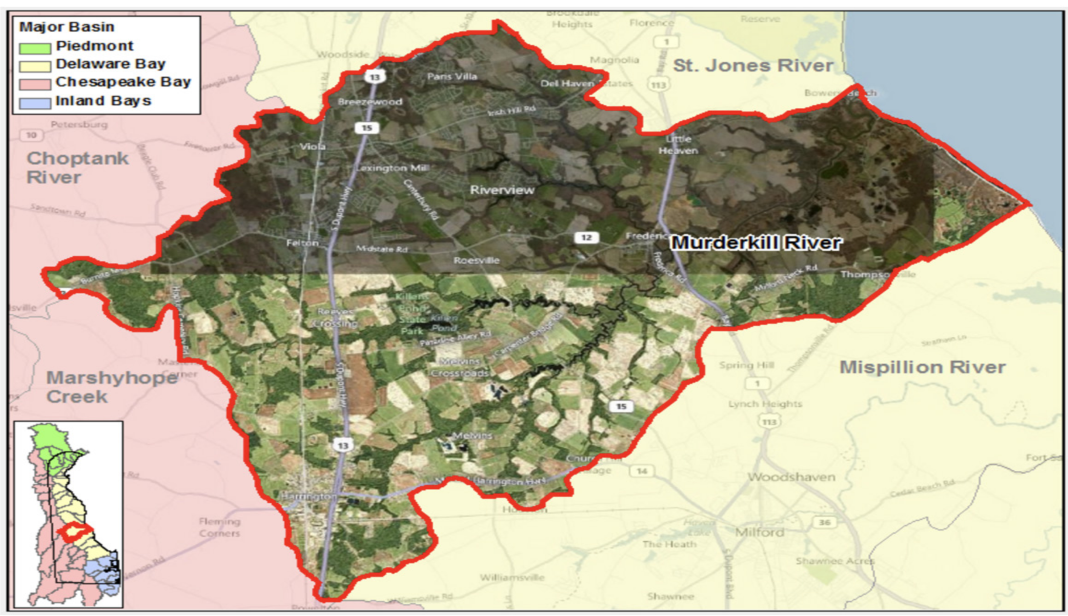


Figure 1. Murderkill River Watershed, Delaware, United States. Source: delawarewatersheds.org.

[1] Note that the origin of the name, Murderkill, has a Dutch origin as "moeder" means mother and "kill" means river or creek in Dutch. Thus, the rough translation of the name is "Mother River", and not a reference to a bloody past.

GIS Environment Setup

In our model, the agents are farmers operating farms in the watershed. However, since farm level data are not publicly available, we develop a method to simulate farm level agents from parcel level data. We obtain three sources of geographic information system (GIS) data for the Murderkill River watershed: (1) Parcel level size and location data for Delaware; (2) Watershed boundary data for Murderkill watershed; and (3) National Land Cover Database (NLCD, 2011). We combine these three data sources together to generate an estimate of the agricultural land for each parcel in the watershed.

Agent Initialization

By combining parcel-level GIS information with data on land cover for this watershed, we can estimate the amount of land used for agriculture within each parcel and the X–Y coordinates of the parcels. Since farms often consist of a constellation of parcels and we do not have data on the actual allocation of parcels to specific landowners, we initialize the size of each farm based on the probability density function from data from the 2012 Census of Agriculture (U.S. Department of Agriculture, 2012) for Kent County, Delaware. Using that information and the GIS information, we match a simulated landowner agent to various numbers of parcels. In this process, we first calculate a "distance matrix" that contains information on the geographic distance between the individual parcels and every other parcel in the watershed. We then create landowner agents by grouping the nearest neighboring parcels until they meet criteria identified by the probability density and average size of each category of farms in the Census of Agriculture. The result is that our agents constructed from neighboring parcels closely mimic the census data on farm size distributions. Figure 2 displays the farm size distribution of Kent County, Delaware, and Figure 3 shows our simulated farm size distribution.

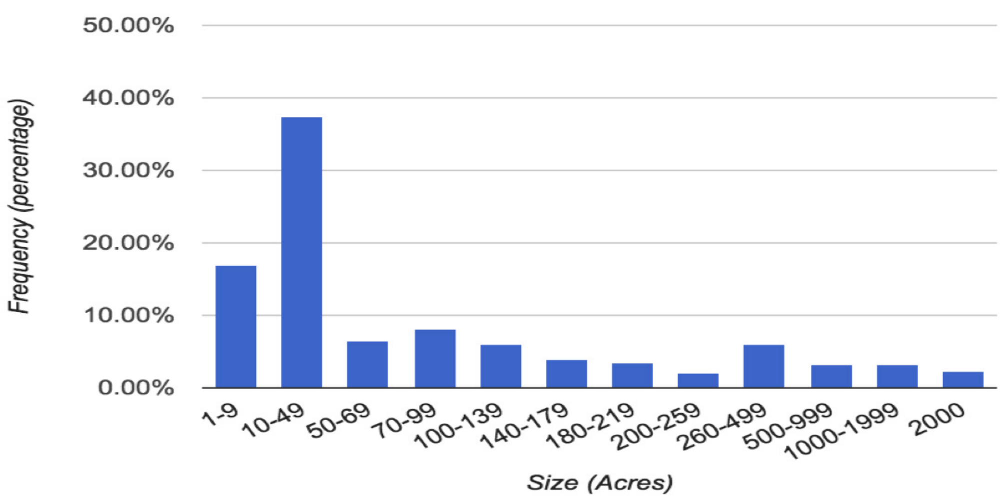


Figure 2. Census farm size distribution of Kent County, Delaware, USA.

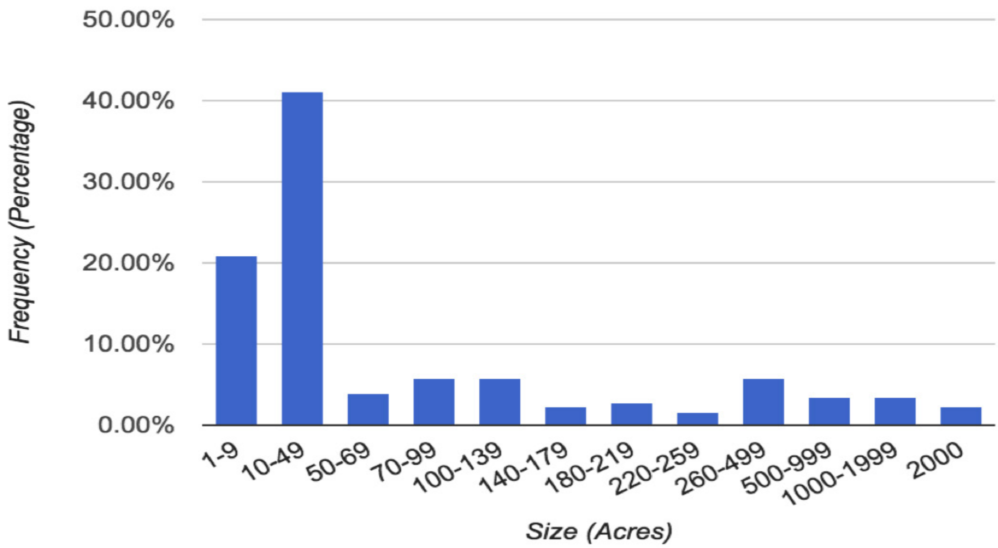


Figure 3. Simulated farm size distribution.

Network and Layout

An agent is assumed to operate a farm that consists of a number of parcels. The agents are placed at the center points of their farms, which are determined using GIS data. Each agent is connected to a number of neighbors based on geographic proximity and influences those neighbors. The modeler at the beginning of each simulation determines the number of agents in one neighbor group.

ABM Model Framework

In the ABM, we adopt the modification of the classic model in environmental and resource economics as documented in our previous section. Each agent operates a farm, generates income by producing an agricultural product (e.g., corn), and simultaneously generates byproducts that cause NPS pollution. The agents may choose to adopt a technology at a cost proportional to its size that could reduce byproducts. As explained before, an underlying dominant Nash strategy could be solved for every agent in the watershed. Since the dominant Nash strategy is the same as the optimal strategy for the social planner's problem, we can treat the Nash strategy as the "Theoretical Target" level of participants' response. A pollution monitor (i.e., sensor) is placed at the downstream end of the watershed, and amount of pollution contributed by each farm is based on the farm's distance from the monitoring point (our experiment measured individual contributions of pollution in the same way). Different policy and regulatory scenarios influence the agents' production and technology-adoption decisions based on results drawn from the experiment. Table 1 summarizes the variables used in the ABM.

ABM Model Process Flow

Figure 4 demonstrates the process flow of our ABM. Each agent makes two decisions, a production decision, and a technology decision. Both decisions influence the income received and the pollution generated by the agent. Combined with pollution generated by other agents, the total pollution is calculated. Depending on whether

Table 1. Summary of fixed, variable, and uncertain parameters

Parameter name	**Value**	**Description**
Fixed parameters		
Number_of_agents	About 174, depending on each realization	Number of agents in simulation
Simulation_horizon	25	Length of each simulation in years
Probability_of_farm_type	Based on cluster analysis and multinomial logistic regression	The probability that each agent will fall into each behavioral type
Target_adoption_rate	Depends on farm grouping results	The target probability that each agent will adopt the technology
Target_production_rate	Depends on farm grouping results	The target production rate for each agent
Number_of_connections	User defined	Number of neighbors of each agent
Factor_technology	0.5	Percent of pollution relative to original level if technology is adopted
Variable parameters		
Unit_corn_production	User defined	Weight of corn produced on one unit of farm size
Unit_pollution_generated	User defined	Average phosphorus generated by one unit of production
Uncertain parameters		
Percent_prod_deviation	Depends on experiment data	Adjusts amount of corn produced per unit of land based on agent type
Adoption_change_prob	Depends on experiment data	Adjusts amount of pollution per unit of production based on agent type
Adopted	Binary, value depends on each realization	Indicates whether the farm adopted the technology
Agent Type	One of several types depending on the cluster analysis of the experiment data	Different types of agents determine different production, pollution, and adoption probabilities

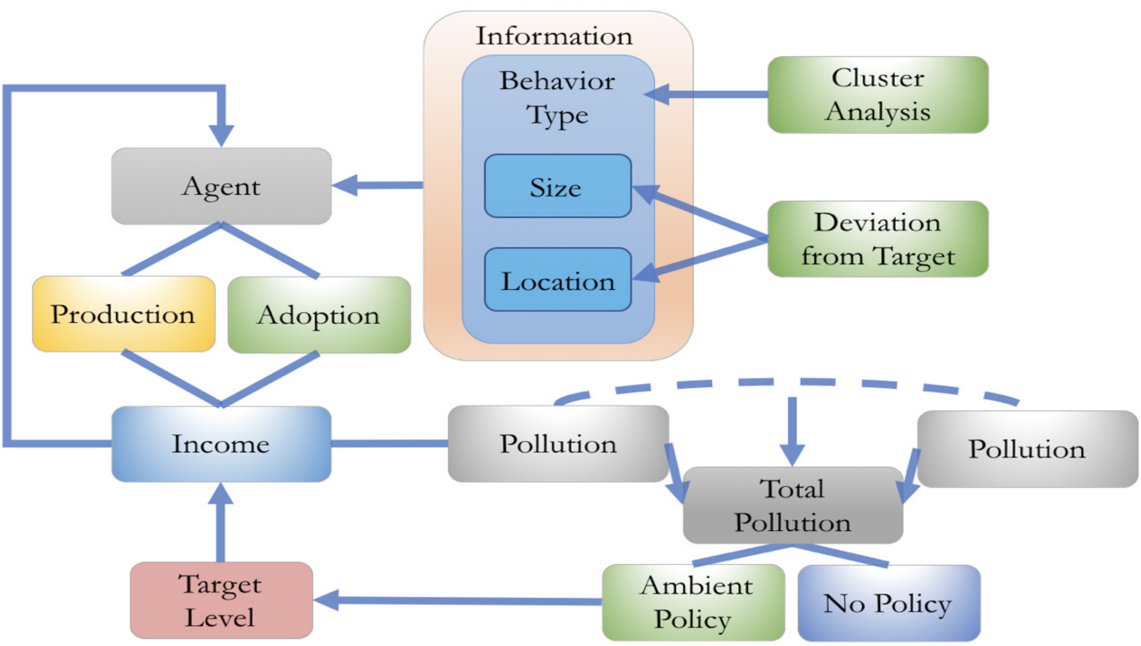


Figure 3. ABM model process flow.

an ambient-based policy is in place, the agent's income may be affected by a tax or subsidy based on the target level and the total environmental damage. This influence on income further affects agent decisions in the next year. An agent's production and adoption decisions are modeled based on the production and adoption deviations from the target levels. These deviations are modeled in two phases as demonstrated in the next section.

Experimental Data Analysis

We conducted statistical data analysis on data from the experiment as documented in Wu, Palm-Forster, and Messer (2017). The analysis was done in two phases. First, we are interested in classifying people into different behavior groups. The idea is to capture the inherent behavioral difference among people (e.g., some people are more environmentally friendly; some are more self-oriented, etc.) Second, after we classify participants into behavior groups, we estimate how agent production and adoption decisions are influenced by their location, size, information treatment, and type. We use the results to calibrate agent decision rules in the ABM model.

Cluster Analysis

Since we do not have any pre-defined knowledge or want to impose any assumption on how many groups participant behavior should be clustered into, the goal of this analysis is to identify the number of behavior types and cluster agents into that number of groups. With no pre-determined grouping structure, meaning that we do not observe the response variables, cluster analysis is suitable for this purpose. As a popular unsupervised statistical learning method, cluster analysis could generate grouping structures based on patterns in predictors. The first key question is to determine into how many clusters the agents should be grouped.

Clustering Metric

To account for the fixed effects of different treatments, the difference between an agent's actual pollution level and the Nash optimal strategy level in that treatment was considered as a measure of the agent's behavior at each round. Therefore, clustering analysis was implemented based on five variables (diff1, diff2, diff3, diff4, diff5), the agents' differences to Nash over five rounds. These variables are defined as

$$\text{Diff}_{ijt} = \text{Pollution}_{ijt} - \text{TargetPollution}_{ijt}\,.$$

where Diff_{ijt} denotes the difference of participant i's pollution level to the target pollution level in treatment j, round t.

There are a number of clustering methods available; the most popular ones include K-means clustering, hierarchical clustering and Gaussian mixture models. There is no definite right or wrong for each of the clustering methods. We selected to use K-means clustering because it generated the most informative grouping structure.

For K-means clustering, the most important task is to determine into how many groups to cluster. This depends on both statistical criterion and knowledge on what a sensible grouping structure is. We perform various statistical procedures to determine the number of clusters.

The Elbow Method

The most intuitive way to determine the number of groups is the "Elbow Method". Figure 5 depicts the within groups sum of squares versus the number of clusters. We can see that there is a sharp turn when the number of cluster is equal to three. Therefore, three appeared to be a reasonable number of clusters to divide into the agent types.

Calinski Criterion

Another popular method for this purpose is the Calinski Criterion (also known as the Pseudo F statistics). Figure 6 shows the results of applying Calinski Criterion to our data. The Calinski Criterion suggests that we should also use three clusters.

Majority Rule

Third, we applied 26 other indices on the same problem and used the majority rule to select the number of clusters. We consider up to 10 clusters as the possible number of clusters into which we could group agents. As shown in Figure 7, the Y-axis means the frequency that a number is selected as the best number of clusters chosen by the indices, and the X-axis is the possible best number of clusters. Eleven out of the 26 indices selected three as the best number of clusters. Therefore, according to the majority rule, we will assume three is

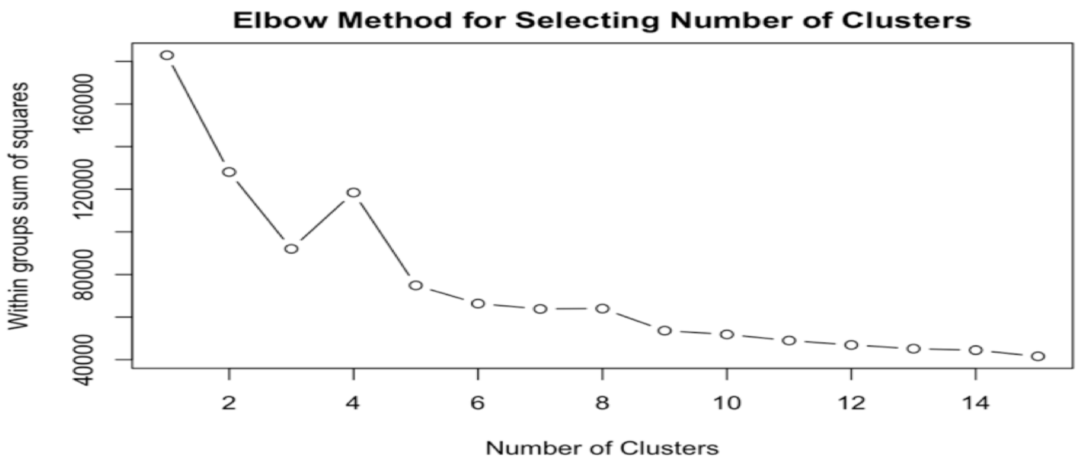


Figure 5. Within groups sum of squares versus the number of clusters.

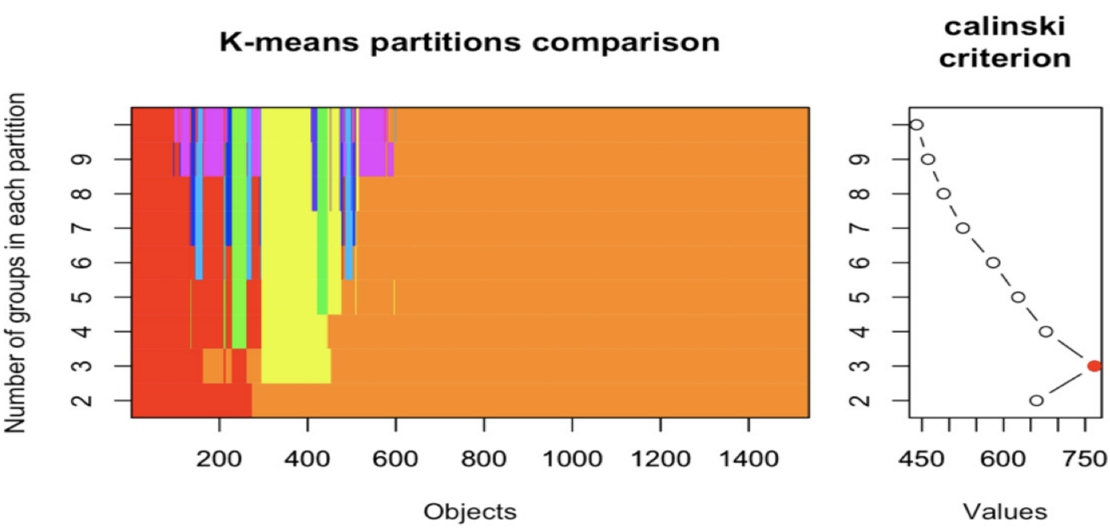


Figure 6. Calinski Criterion results.

the number of clusters we should use in the K-means clustering.

Separation Examination

We examine if the three clusters generated by K-means clustering provide reasonable separation for the data. We perform K-means clustering, assigning three as the number of clusters, and setting the seed to 20 to ensure reproducibility.

The results of the K-means clustering are summarized in Table 2. As we can see, the median values for Group 3 in all

Figure 7. Results of using a majority rule with 26 grouping indices.

Table 2. Group frequencies and median value by grouping variables and groups.

Cluster	**Frequency**	**Diff1**	**Diff2**	**Diff3**	**Diff4**	**Diff5**
1	171	–7.500	–5.60	–6.00	–6.16	–5.44
2	150	6.615	5.28	5.11	5.94	5.25
3	1,215	0.000	0.00	0.00	0.00	0.00

five rounds are equal to zero, meaning that group 3 is the group that tends to behave in accordance with the theoretical prediction. Group 1 and Group 2 have median values that are lower and higher than the target pollution, respectively. This means that Group 1 is the group that tends to generate less pollution than theoretically predicted and Group 2 is the group that tends to generate more pollution than theoretically predicted. We do not see obvious skewness or scarcity of any groups and the magnitudes of the separation seem reasonable. Next, we assign agents in the ABM into behavior groups using a multinomial logit model.

Mixed-Effects Multinomial Logit Model to Assign Group Probabilities

Based on cluster analysis, agents' behavior could be clustered into three categories. Cluster 3 corresponds to agents that tend to agree with theoretical predictions, and clusters 1 and 2 correspond to agents that tend to under and over pollute, respectively. In this part, we use a mixed effects multinomial logit model to estimate the cluster distributions among agents conditioning on the policy, heterogeneity, and information treatments.

The multinomial logit model could be formulated as follows:

$$\log\left(\frac{Pr(cluster=1)}{Pr(cluster=3)}\right)\Bigg|(Policy=j)$$
$$=f(u_{1,i},HeteroTreatments,InfoTreatments,HeteroInfo\_Interactions)$$
$$=X_iB_{1i}$$

$$\log\left(\frac{Pr(cluster=2)}{Pr(cluster=3)}\right)\Bigg|(Policy=j)$$
$$=f(u_{2,i},HeteroTreatments,InfoTreatments,HeteroInfo\_Interactions)$$
$$=X_iB_{2i}$$

where $u_{1,i}$, $u_{2,i}$ are the random effects on the intercept and are assumed to follow a normal distribution. J equals 1 or 0 and denotes whether the policy treatment is in place or not, respectively.

Therefore, the predicted probabilities for the three clusters could be calculated as

$$\Pr(cluster_i=1)=\frac{\exp(X_iB_{1i})}{1+\exp(X_iB_{1i})+\exp(X_iB_{2i})}$$
$$\Pr(cluster_i=2)=\frac{\exp(X_iB_{2i})}{1+\exp(X_iB_{1i})+\exp(X_iB_{2i})}$$
$$\Pr(cluster_i=3)=\frac{1}{1+\exp(X_iB_{1i})+\exp(X_iB_{2i})}$$

The results of the mixed effects multinomial logit model for both policy and no policy treatments are presented in Table 3.

In the no policy treatments, it is always in the agents' best interest to produce at the maximum and not adopt the technology, therefore, the theoretical optimal strategy is the upper bound of the pollution level. As a result, only two clusters exist in the no policy treatments, as reflected by having one intercept value in Table 3. Based on the above regressions, we calculate the cluster probabilities for each of the treatment cases to initialize the model.

Modeling Agent Production and Adoption Behavior

For production decisions, we calculate the percentage deviations from the target production decisions, taking into account the size of the farm. The metric is defined as

$$PerProdDiff=\frac{Production-TargetProd}{Size}.$$

Table 3. Mixed-effects multinomial logit model to assign agent behavioral types

	Policy		No policy	
	Coefficient	Std. Err.	Coefficient	Std. Err.
intercept 1	−2.0416***	0.1119	−10.5036***	1.3820
intercept 2	−0.8200***	0.1127		
hetero1	0.3183*	0.1691	0.5156	0.4567
hetero2	0.5703***	0.1562	0.5156	0.4567
hetero3	0.4495***	0.1688	−0.5666	0.4796
info1	−1.2637***	0.2073	0.8735	1.3800
info2	−1.1185***	0.2123	−1.3688	1.4154
info1_hetero1	0.7143***	0.2746	−1.5310***	0.6499
info1_hetero2	0.6048**	0.2647	−4.2738***	0.8099
info1_hetero3	1.2097***	0.2745	−4.4023***	0.9440
info2_hetero1	0.3696	0.2631	0.1780	0.7027
info2_hetero2	1.3090***	0.2661	0.7511	0.6957
info2_hetero3	1.4883***	0.2566	2.8191***	0.7137
Number of Observations	3840		3840	
Number of groups	192		192	

***, **, * denote significant as 1%, 5% and 10% level, respectively. All standard errors are clustered as individual level.

Table 4. Deviations from target production levels

	With policy			Without policy		
	No Info	Info1	Info2	No Info	Info1	Info2
constant	1.32*** (0.11)	0.16 (0.16)	1.41*** (0.13)	0.0015 (0.051)	−0.0039 (0.015)	−0.11 (0.14)
Size	−0.0095*** (0.00067)	−0.00086 (0.0031)	−0.0089*** (0.00081)	−0.00019 (0.00016)	−0.000073 (0.000074)	0.000083 (0.00012)
region	−1.19*** (0.30)	0.25 (0.42)	−1.41*** (0.30)	0.032 (0.15)	0.023 (0.038)	−0.015 (0.026)
cluster1	−0.20*** (0.040)	−0.19*** (0.04)	−0.13*** (0.038)	−0.12*** (0.042)	−0.068** (0.032)	−0.057** (0.024)
cluster2	0.26*** (0.031)	0.21*** (0.039)	0.19*** (0.036)			
info1_adopt		−0.21** (0.085)				
info1_prod		−0.000028 (0.0029)				
info2_adopt			0.064 (0.063)			0.00040 (0.013)
info2_prod			−0.0020* (0.0011)			0.00098 (0.0014)

***, **, * denote significant as 1%, 5% and 10% level, respectively. All standard errors are clustered as individual level.

In this case, we run a random-effects OLS model for each policy and information segment of data with standard errors clustered at the individual level.

For adoption decisions, we calculate the probability that an agent deviates from its target adoption decision, which means the probability that an agent changes its adoption decision away from the theoretical prediction. The metric is defined as the absolute difference of the actual adoption decision and the target adoption decision:

$$AdoptChange = |ActualAdopt - TargetAdopt|$$

Since the variable AdoptChange is binary, we run a random effects logit model for each policy and information treatment segments with individual clustered standard errors. The result of the model is shown below:

Based on the above regressions, we parameterize the agent's actual production

Table 5. Deviations from target adoption decisions

	With policy			Without policy		
	No Info	Info1	Info2	No Info	Info1	Info2
constant	6.76*** (2.17)	5.69** (2.33)	9.67*** (2.50)	−2.86 (2.05)	−5.17*** (1.09)	−14.96 (10.73)
size	−0.051*** (0.012)	−0.018 (0.042)	−0.050*** (0.013)	−.0096 (0.012)	−0.0041 (0.011)	−0.016 (0.017)
region	−7.09 (4.66)	−18.09*** (6.73)	−18.44*** (5.85)	−6.90 (5.96)	3.48 (3.35)	11.01* (6.61)
cluster1	−0.39 (0.44)	0.63 (0.57)	−0.27 (0.59)	2.13** (0.91)	3.48*** (0.54)	3.79*** (0.71)
cluster2	0.74** (0.33)	1.22*** (0.44)	1.10*** (0.38)			
info1_adopt		−2.29* (1.23)				
info1_prod		0.013 (0.040)				
info2_adopt			−0.59 (0.90)			4.48** (2.07)
info2_prod			−0.0012 (0.018)			0.079 (0.11)

***, **, * denote significant as 1%, 5% and 10% level, respectively. All standard errors are clustered as individual level.

and adoption decisions relative to their Nash optimal strategy levels.

Calibration

Prices

Agents are assumed to produce an agricultural good (corn) and act as price-takers. Given the constant fluctuation of corn prices in the United States, we conduct an OLS regression for the mean corn price from 1996 to 2016 on logarithm of year to capture the general price trend, and use a triangular distribution with maximum and minimum defined by the predicted mean and standard deviation of the prices to reflect the fluctuation.

Yield

In order to determine how much agricultural product (corn) is produced by each agent, we calculate the average yield of each unit of land. Similarly, we conduct an OLS regression for mean corn yield from 1996 to 2016 on logarithm of year to capture the general trend in corn production, and use a triangular distribution to reflect the fluctuation, with maximum and minimum defined by the standard deviation of the average yield.

Pollution

Following Zia et al. (2016b), we provide an estimate for the average Phosphorus leakage of cornfields based on the maximum and minimum Phosphorus loss estimates. During each simulation, the modeler has the option to modify the mean and standard deviation of average Phosphorus leakage. However, this value affects all simulation cases equally and therefore does not influence any relative comparison conclusions we draw.

Simulation Results

We present the results of the simulation experiment, and discuss the sensitivity of the results.

The Effect of Information Treatments

We compare how different information treatments would affect the performance of the ambient-based policy. Both location and size heterogeneity are included in this simulation. In Figure 8, the red line indicates the target level of pollution, and the blue line indicates the experiment simulation results. From left to right, the three subfigures indicate the results for the no information case, individual level information case, and group level information case. As shown in Figure 8, a gap exists between the simulated pollution level (blue line) and the target pollution level (red line) in the no information case; however, the gap is much smaller in either of the information treatments. This suggests that under this simulation scenario, both information treatments decrease the deviation between the target pollution level and the simulated level, which indicates that the effect of policy is stronger when ambient-based policy is coupled with information "nudges".

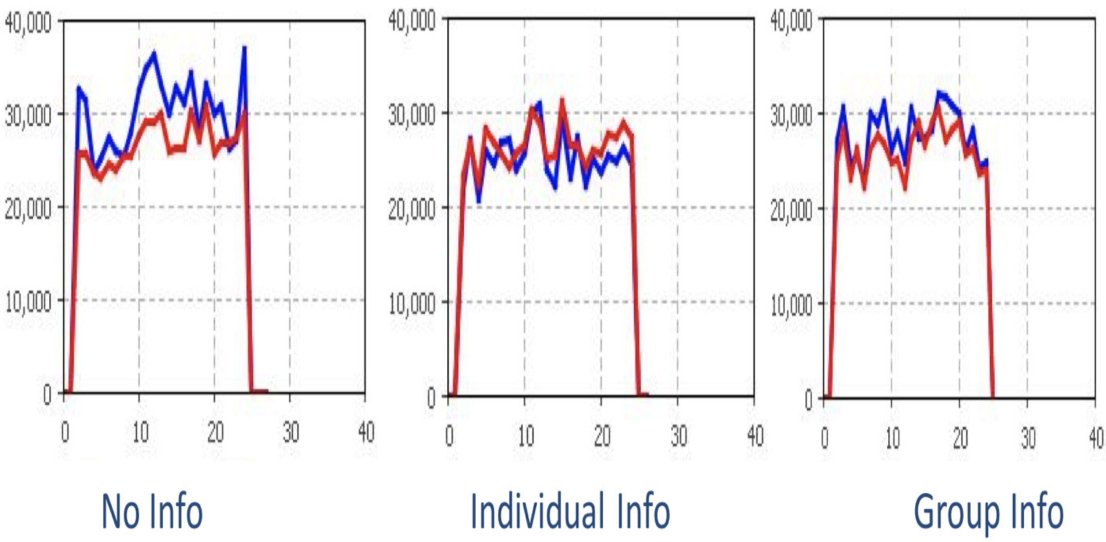


Figure 8. Effects of information on pollution level. The red lines indicate target levels and the blue lines indicate experiment simulated results'.

Comparing Individual decisions

In this section, we break up pollution by production decisions, adoption decisions, and size of the farms. We look at each information treatment separately.

No Information

Recall that under no information baseline, the aggregate simulated pollution level is mostly over the target level, but the deviation does not appear to be large. However, when we break up pollution into production and adoption decisions by farm size, we observe huge deviations in these decisions (Figure 9). The small farms are significantly over adopting the technology (blue lines), and the large farms are widely under adopting, even though it is not in their best interest to do so (as depicted by the blue lines). Similarly, the small farms are also over producing and large farms are under producing.

Individual Level Information

When we provide participants with individual level information on what people like them have done in the past, we observe that the deviations from participants' behavior to the target levels are much smaller (Figure 10). This clearly demonstrates that the individual level information induces participants to make better decisions, and improves policy efficiency.

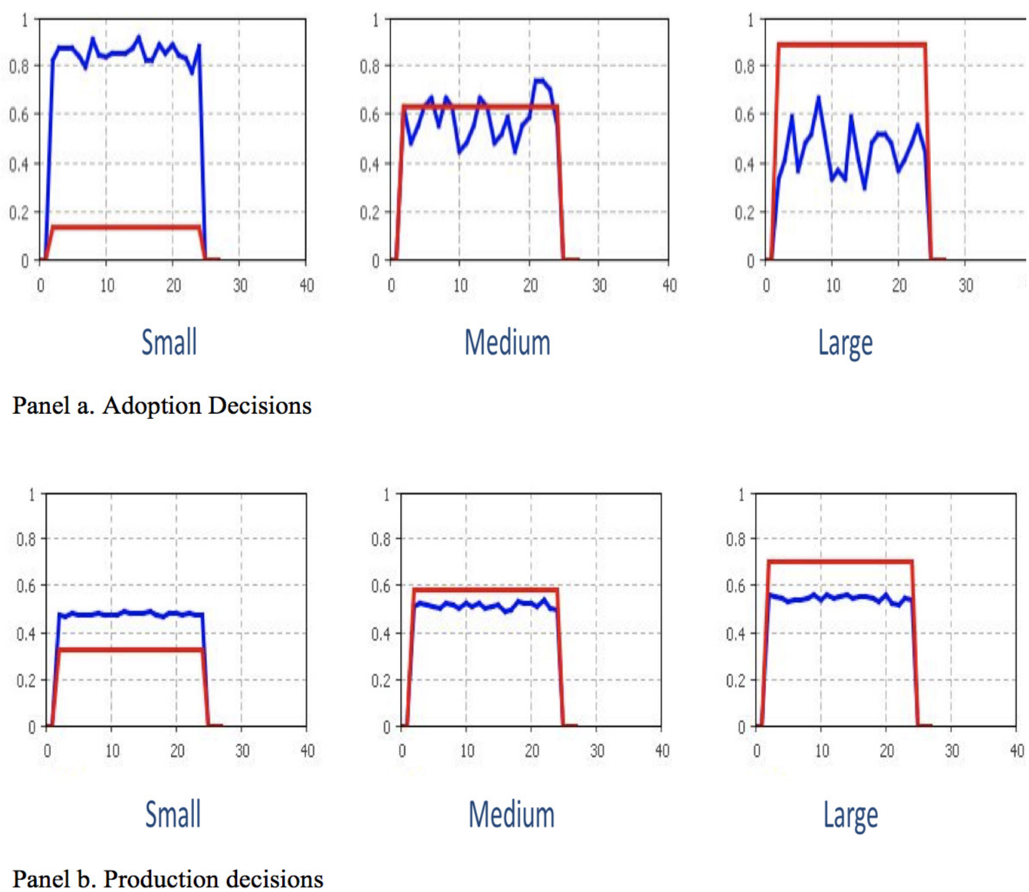


Figure 9. Adoption and production decisions by size under no information treatment. The red lines indicate target levels and the blue lines indicate experiment simulated results.

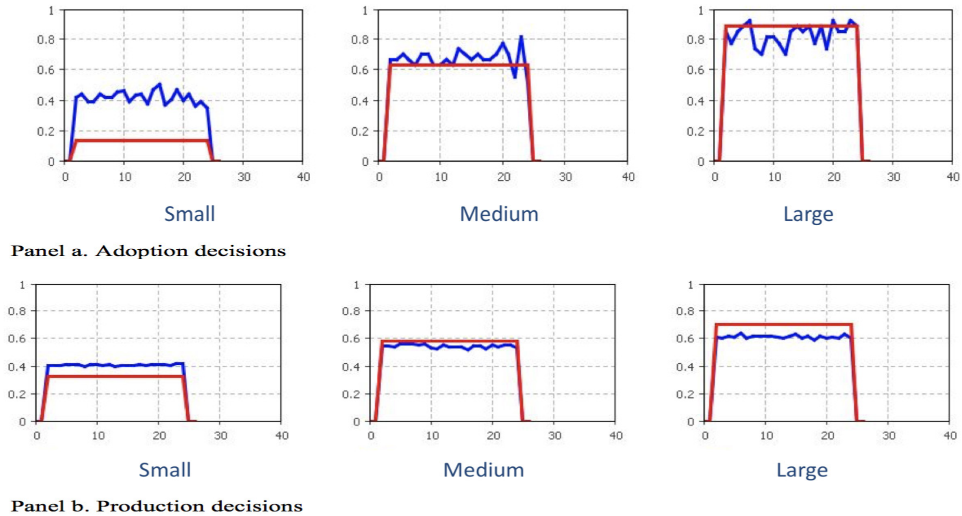


Figure 10. Adoption and production decisions by size under individual level information. The red lines indicate target levels and the blue lines indicate experiment simulated results.

Group Level Information

When participants are informed with group level information on average adoption and production decisions in their group in the last round, we find that this information helps participants make better decisions than the no information baseline, but the policy efficiency is lower compared to individual level information scenario (Figure 11). Furthermore, small farms tend to over adopt and over produce, and the large farms tend to under adopt and under produce. However, compared to the target levels, the deviations between the target and the simulated results are smaller compared to the no information scenario, but larger than when people were given individual level information. Therefore, group level information helps the policy performance and efficiency on an aggregate level, but the policy efficiency is lower than if people were given individual level information.

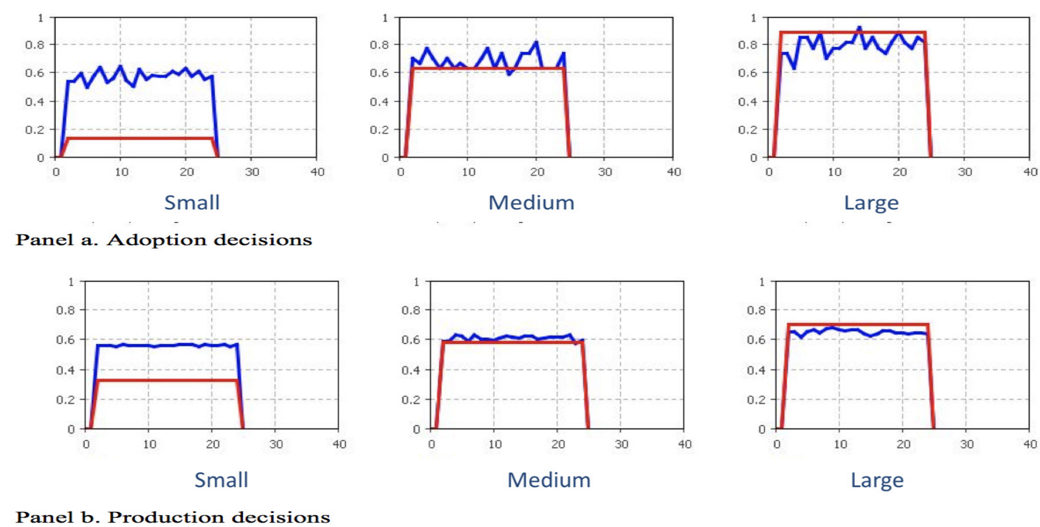


Figure 11. Adoption and production decisions by size under individual level information. The red lines indicate target levels and the blue lines indicate experiment simulated results.

Possible Explanations

Finally, we want to provide some discussion of the potential reasons for the patterns that were demonstrated in the simulation results. Under the no information treatment, the adoption of the technology is largely negatively related to the size of the farm. A possible explanation for this observation is that since the cost of adopting the technology is proportionally related to the size of the farm, participants may follow some heuristic decision rules that attribute significant weight to the cost of adopting in the processes. This clearly demonstrates that as opposed to always following profit maximizing decision rules, human behavior is often limited in their calculating ability and may be affected by various cognitive reasons and therefore demonstrate bounded rationality in terms of forming some rather heuristic decision rules. Furthermore, both information treatments seem to provide anchors for the participants. Knowing what people like them have done in the past and what others in their group have done provide people with a reference point in their decision process. Since individual level information provides people with tailored information, it helps people make better decisions compared to the myopic baseline case. Under the group level information where a group average is provided, we can observe that the absolute adoption and production decisions for farms with different sizes tend to be very close. This suggests that people might be anchored to the group level averages, or peer actions, even though it might not be in their best interest to do so.

Sensitivity Analysis

In this part, we discuss how our results would be affected by uncertain parameters in our ABM. Ideally, the result of an ABM should come from complex agent interactions and adaptions in a concise model rather than from complex assumptions about individual behavior and free parameters (Axelrod, 1997). Most of the parameters that influence the observed results in the ABM are calibrated and validated based on experimental data. Therefore, the uncertainty only results from realization of the randomness in each simulation experiment, which is stochastic in nature and should not generate any systematic biases. Meanwhile, if an uncertain variable affects each scenario of the simulation in an equal magnitude, the relative comparisons between the scenarios will not be affected. Therefore, one uncertain parameter that would possibly affect the result is how many farms the participants consider part of their group. This parameter affects the grouping structure and the group level information that is shown to the participants. In our baseline scenario presented before, we assume five people are considered to be in one group. We increase this parameter to 10, 15, and 20 in this part and the result is shown in Figure 12.

As shown in Figure 12, as the number of people that the participants consider themselves to be in the same group with increase, the deviation from the target pollution level and the simulated pollution level is not largely affected under individual level information treatment, but increases under the group level information treatment. This suggests that individual level information not only generates highest policy efficiency, but also is more robust to participant perceptions on their group size.

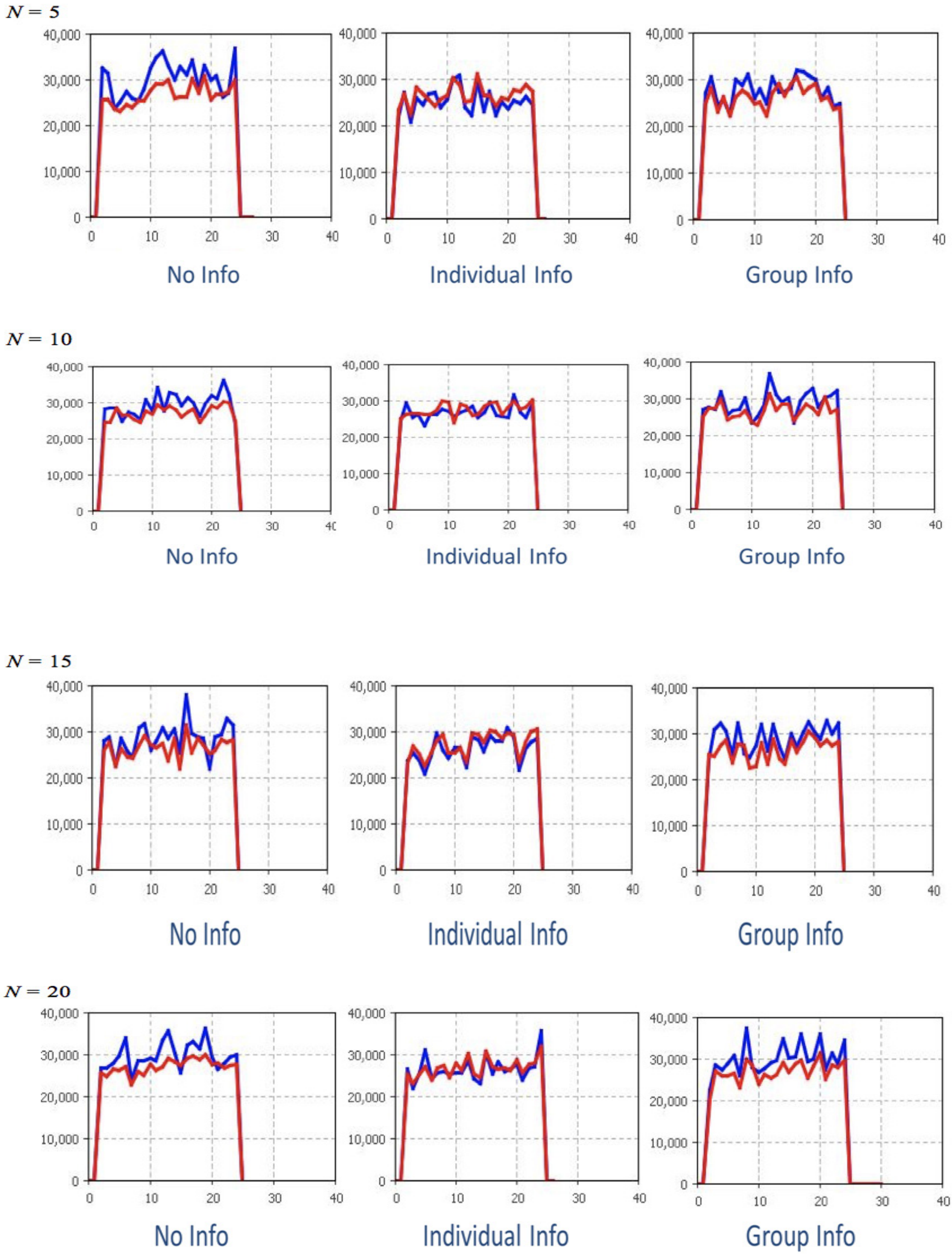


Figure 12. Sensitivity test on group Size. The red lines indicate target levels and the blue lines indicate experiment simulated results.

Conclusions and Discussions

Our study is one of the first that integrates economic experiments with agent-based modeling in a nonpoint source pollution setting. The ABM extends and scales up the findings from the economic experiment by providing a spatially explicit simulation environment based on an actual watershed. Instead of assuming full rationality, the economic experiment calibrates and validates the ABM by defining human-based bounded rational decision rules for the agents. We apply a modification of a classic game theoretical model from the environmental economics literature to the ABM and the experiment as the core underlying model in both scenarios. We define the target level (fully rational theoretical level) by solving for unique dominant Nash strategy. Using experimental data, we first identify the number of behavioral groups using exploratory cluster analysis and then group agents into the three identified groups by multinomial logistic model; second, we define agent decision rules by estimating adoption and production deviations from the target levels based on the information treatment, type, size, and location of each agent. The result of our simulation experiment demonstrates that both information "nudges" help the performance of the ambient-based policy. Individual level information induces higher policy efficiency compared to group level information, where the individual decisions tend to be anchored to the group averages, even though it may not be in their best interest. Our results show in a spatially explicit watershed setting that ambient-based policies, coupled with information "nudges" to provide guidance to people's behavior, have the ability to induce group level compliance, and the policy efficiency is higher when individual level information is being provided. Therefore, it is important to use informational "nudges" to help people make better decisions, especially under complex heterogeneous scenarios.

There are a number of limitations and directions of future work based on our research. First, a more complicated hydrological model may be developed and incorporated in the ABM and the experiment. Examples of such models include the WWACShed model by Tesfatsion et al. (2017) and the SWAT model used in Ng et al. (2011). However, if one attempts to also include bounded rationality in the agent decision processes and use economic experiments to capture these irrationalities, it is crucial to ensure that the conclusions from the experiment could be safely carried over to the ABM. In our experiment, this link was built by adopting the same underlying model and therefore the same incentives around the dominant Nash strategies. If a more complicated model were in place, it would be hard to solve for a perfect rational utility maximization prediction, and therefore would be difficult to have a baseline to compare with actual human behavior. Additionally, the more complicated a model is, the more information burden is introduced to the participants and the harder for the participants to generate informed decisions, so one needs to think carefully about the tradeoff.

Second, another extension of this research is to use farmer sample instead of a sample from university students in the experiment, aiming to increase externality validity of the experiment. The majority of research comparing samples from students and professionals generally find the two samples demonstrate similar responses in both agricultural (Cummings, Holt, & Laury, 2004; Fooks et al. 2016; Messer, Kaiser, & Schulze, 2008) and non-agricultural (Vossler, Mount, Thomas, & Zimmerman, 2009) contexts, it may still be a valid extension since the decision process of farmers is likely different from that of students. However, one also needs to note that the farmers may treat the experiment as a pre-policy evaluation and therefore behave strategically in hopes to influence policy makers (Suter & Vossler, 2013).

There have been very few articles in the literature trying to integrate experiments and agent-based modeling even though the integration would benefit both fields. This could probably largely be attributed to the interdisciplinary nature of the field, and the challenge to build a credible link between the two. The method in our study could be used as a framework to combine these two fields and help motivate future research in this area.

References

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, *95*(9), 1082–1095.

An, L. (2012). Modeling human decisions in coupled human and natural systems: Review of agent-based models. *Ecological Modelling*, 229, 25–36.

Axelrod, R. M. (1997). *The complexity of cooperation: Agent-based models of competition and collaboration.* Princeton, NJ: Princeton University Press.

Axtell, R. L., Epstein, J. M., Dean, J. S., Gumerman, G. J., Swedlund, A. C., Harburger, J., ... & Parker, M. (2002). Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley. *Proceedings of the National Academy of Sciences of the USA*, *99*(3), 7275–7279.

Cummings, R. G., Holt, C. A., & Laury, S. K. (2004). Using laboratory experiments for policymaking: An example from the Georgia irrigation reduction auction. *Journal of Policy Analysis and Management,* 23(2), 341–363.

Dia, H. (2002). An agent-based approach to modelling driver route choice behaviour under the influence of real-time information. *Transportation Research Part C: Emerging Technologies*, *10*(5), 331–349.

Duffy, J. (2006). Agent-based models and human subject experiments. *Handbook of computational economics*, 2, 949–1011.

Evans, T. P., Sun, W., & Kelley, H. (2006). Spatially explicit experiments for the exploration of land-use decision-making dynamics. *International Journal of Geographical Information Science*, *20*(9), 1013–1037.

Farmer, J. D., & Foley, D. (2009). The economy needs agent-based modelling. *Nature*, *460*(7256), 685–686.

Ferraro, P. J., & Price, M. K. (2013). Using nonpecuniary strategies to influence behavior: Evidence from a large-scale field experiment. *Review of Economics and Statistics*, *95*(1), 64–73.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140.

Fooks, J. R., Higgins, N., Messer, K. D., Duke, J. M., Hellerstein, D., & Lynch, L. (2016). Conserving spatially explicit benefits in ecosystem service markets: Experimental tests of network bonuses and spatial targeting. *American Journal of Agricultural Economics*, *98*(2), 468–488.

Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research*, *35*(3), 472–482.

Grimm, V., & Railsback, S. F. (2005). *Individual-based modeling and ecology* (Volume 2005). Princeton, NJ: Princeton University Press.

Heckbert, S. (2009, July). Experimental economics and agent-based models. In *18th World IMACS/MODSIM Congress* (pp. 2997–3003).

Heckbert, S., Baynes, T., & Reeson, A. (2010). Agent-based modeling in ecological economics. *Annals of the New York Academy of Sciences*, 1185(1), 39–53.

Johnson, L., Lemasters, L., & Bhattacharyya, P. (2017). A complexity context to classroom interactions and climate: ABM forecasting achievement. *Journal on Policy and Complexity Systems*, *3*(1), 28–47.

Khouja, M., Hadzikadic, M., & Zaffar, M. A. (2008). An agent based modeling approach for determining optimal price-rebate schemes. *Simulation Modelling Practice and Theory*, *16*(1), 111–126.

Matthews, R. B., Gilbert, N. G., Roach, A., Polhill, J. G., & Gotts, N. M. (2007). Agent-based land-use models: A review of applications. *Landscape Ecology*, *22*(10), 1447–1459.

Messer, K. D., Kaiser, H. M., & Schulze, W. D. (2008). The problem of free riding in voluntary generic advertising: parallelism and possible solutions from the lab. American *Journal of Agricultural Economics*, *90*(2), 540–552.

Miao, H., Fooks, J. R., Guilfoos, T., Messer, K. D., Pradhanang, S. M., Suter, J. F., … & Uchida, E. (2016). The impact of information on behavior under an ambient-based policy for regulating nonpoint source pollution. *Water Resources Research*, *52*(5), 3294–3308.

Ng, T. L., Eheart, J. W., Cai, X., & Braden, J. B. (2011). An agent-based model of farmer decision-making and water quality impacts at the watershed scale under markets for carbon allowances and a second-generation biofuel crop. *Water Resources Research*, *47*(9), 1–17.

NLCD (2011). *National Land Cover Database*. Multi-Resolution Land Characters Consortium. Retrieved from https://www.mrlc.gov/nlcd2011.php.

Poe, G. L., Schulze, W. D., Segerson, K., Suter, J. F., & Vossler, C. A. (2004). Exploring the performance of ambient-based policy instruments when nonpoint source polluters can cooperate. *American Journal of Agricultural Economics*, *86*(5), 1203–1210.

Segerson, K. (1988). Uncertainty and incentives for nonpoint pollution control. *Journal of Environmental Economics and Management*, *15*(1), 87–98.

Spraggon, J. (2002). Exogenous targeting instruments as a solution to group moral hazards. *Journal of Public Economics*, *84*(3), 427–456.

Suter, J. F., & Vossler, C. A. (2013). Towards an understanding of the performance of ambient tax mechanisms in the field: Evidence from upstate New York dairy farmers. *American Journal of Agricultural Economics*, *96*(1), 92–107.

Suter, J. F., Vossler, C. A., & Poe, G. L. (2009). Ambient-based pollution mechanisms: A comparison of homogeneous and heterogeneous groups of emitters. *Ecological Economics*, *68*(6), 1883–1892.

Tesfatsion, L., Rehmann, C. R., Cardoso, D. S., Jie, Y., & Gutowski, W. J. (2017). An agent-based platform for the study of watersheds as coupled natural and human systems. *Environmental Modelling & Software*, 89, 40–60.

Tsai, Y., Zia, A., Koliba, C., Bucini, G., Guilbert, J., & Beckage, B. (2015). An interactive land use transition agent-based model (ILUTABM): Endogenizing human-environment interactions in the Western Missisquoi Watershed. *Land Use Policy*, 49, 161–176.

USDA NASS (2012). 2012 *Census of Agriculture.* Retrieved from https://www.agcensus.usda.gov/Publications/2012/

Veldkamp, A., & Verburg, P. H. (2004). Modelling land use change and environmental impact. *Journal of Environmental Management*, *72*(1), 1–3.

Vossler, C. A., Mount, T. D., Thomas, R. J., & Zimmerman, R. D. (2009). An experimental investigation of soft price caps in uniform price auction markets for wholesale electricity. *Journal of Regulatory Economics*, *36*(1), 44–59.

Wu, S., Palm-Forster, L., & Messer, K. D. (2017). *Heterogeneous agents and information nudges in non-point source water pollution management* (APEC working paper RR 17-04). Newark, DE: University of Delaware Press.

Xepapadeas, A. (2011). The economics of non-point-source pollution. *Annual Review of Resource Economics*, *3*(1), 355–373.

Xepapadeas, A. P. (1992). Environmental policy design and dynamic nonpoint-source pollution. *Journal of Environmental Economics and Management*, *23*(1), 22–39.

Zenobia, B., Weber, C., & Daim, T. (2009). Artificial markets: A review and assessment of a new venue for innovation research. *Technovation*, *29*(5), 338–350.

Zia, A., & Koliba, C. (2015). The emergence of attractors under multi-level institutional designs: Agent-based modeling of intergovernmental decision making for funding transportation projects. *AI & Society*, *30*(3), 315–331.

Zia, A., Messer, K., Ding, S., Miao, H., Suter, J., Fooks, J., … & Koliba. C., (2016a). Spatial effects of sensor information in inducing cooperative behaviors for managing non-point source pollution: evidence from a decision game in an idealized watershed. *Ecological Economics (Under Review)*.

Zia, A, Tsai, Y., Turnbull, S., Ding, S., Miao, H., Koliba, C., … & Uchida, E. (2016b). Simulating the effects of alternate control strategies on heterogeneous farmer behaviors and water quality outcomes: An agent-based modeling application in the Missisquoi Watershed of Lake Champlain Basin. *SwarmFest2016: 20th Annual Meeting on Agent-based Modeling and Simulation.* Burlington, VT, July 31 – August 2, 2016.

APPENDIX A: RESEARCH DOCUMENTS AND PROTOCALS

Thank you for participating!

Please return the signed consent form to the administrator.

Please read and follow the instructions carefully and do not communicate with others during the experiment.

INTRODUCTION

This is an experiment about the economics of decision-making. You will earn money during this experiment if you follow these instructions carefully and make informed decisions; otherwise, you may end up losing money. Any money earned during this experiment will initially be recorded as experimental dollars. At the end of this experiment, we will convert your experimental dollars into actual US dollars that will be handed to you as you leave. The more experimental dollars you earn the more actual US dollars you will receive. At the end of the experiment, your earnings will be converted at a rate of $1 US dollar for 50 experimental dollars. Please read these instructions carefully and do not communicate with any other participants during the experiment.

General Instructions: Today's experiment has several parts. Each part will have five rounds. Each round is independent, meaning that decisions during a round do not affect future rounds in any way. The only value that gets carried over across rounds is the cumulative amount of money you earn, which will be used to calculate your cash earnings at the end of the experiment.

Your role: You own and operate a firm. You will make decisions that affect the amount of money your firm earns. This money will be called your **Firm Profit**.

Groups: Throughout the experiment, you will be in a group of eight people, each will play the role of a firm. Think of your firm and the seven other firms as being located near a river. Groups are randomly reassigned after each part of the experiment and you will not know who is assigned to each group.

Production and Production Income: Each business owner produces output that creates **Production Income**. Production income only depends on how much is produced. The more a firm produces, the more production income the firm will get.

Pollution: Production also generates pollution that goes into the river. In general, the higher the output being produced, the more pollution is being generated. Some concentration of this pollution is harmless. However, if the concentration is too large, the pollution has negative effects to the environment.

Total Pollution: This is measured by a sensor downstream and is the sum of pollution for everyone in the same group. Capacity: The firms may have a different production capacity, which is the maximum amount your firm can produce. Each firm's capacity will be shown on the calculator in the corresponding part for that firm. There are three types of capacities: Large firms with a capacity of 125; medium firms with a capacity of 100; small firms with a capacity of 75.

Technology: At the beginning of each round, the firms may choose to adopt a technology at a cost proportional to your firm capacity. When adopted, the technology will reduce the firm's pollution to a certain percentage of the original level for that round.

Location: The firms may either be located in the same location or at different locations along a river. As shown in Figure 1, when the region is separated by lines, it means the region is being divided into Region 1 to Region 4. In this case, Region 1 is the most upstream and Region 4 is the most downstream. The further downstream your firm is the more pollution per unit of production will be recorded by the sensor. As shown in Figure 2, when there are no lines separating the region, it means all of the firms are placed in the same region. The actual capacity and location of the firm that you operate will be shown on your computer screen.

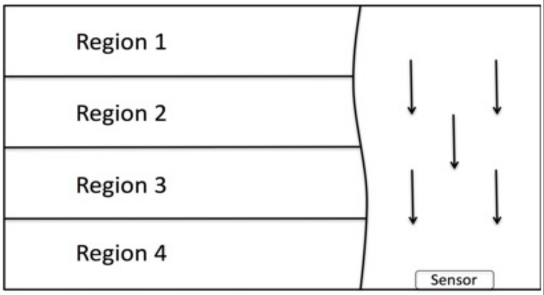


Figure 1. Different Locations

Figure 2. Same Location

Decisions: In each round, you will make two decisions:

(1) **Production Decision**—You will decide your firm's production level, between 0 and your firm's capacity.

(2) **Technology Decision**—You will choose whether to adopt a technology at a certain cost, labeled "Not Adopt" or "Adopt".

Pollution Table: To help you better understand the relationship of production, technology, location and pollution, you are given a **Pollution Table** that has pollution levels of a firm corresponding to different production decisions, technology decisions, and location. Use this table to understand how your production would affect pollution based on your location and technology decision.

Firm Profit: Your **firm profit** is calculated based on your production decision and technology decision and will be explained to you in further details in each part of the experiment.

Decision Calculator: A **Decision Calculator** is provided to test different scenarios to see how the decisions of other firms in your group could affect Total Pollution and your Firm Profit. Follow the instructions on how to use this calculator provided on the next page.

In summary:

- In each part of the experiment, you will be given additional instructions and all calculations will be described.
- Your earnings from the experiment depend on your cumulative firm profit.
- Use the decision calculator to test out different scenarios and determine your own production and technology decision.
- Choose your own production and technology decision and click "Confirm".
- Your production income is affected by your production decision, technology decision, and firm capacity.
- Your pollution depends on your production decision, technology decision, and firm location.
- A round of the experiment is complete when all eight players have made their production and technology decisions.
- After each part, participants will be randomly reassigned to a new group.

HOW TO USE THE DECISION CALCULATOR AND MAKE DECISIONS

In each round, you will be provided with a decision calculator like the one in the attached handout.

The layout of all firms and their corresponding capacity in your group is shown in the calculator.

Your firm is labeled "Your Firm" and marked with a black box.

Step 1. On the left part of the page, assume what everyone in your group will be doing by choosing a production and technology decision for every firm. To choose a production decision, move the slider or type in the amount that you think other firms will be producing; to choose a technology decision, simply choose between the "Not Adopt" and "Adopt" options. Note that your firm is labeled in the black box and you do not have to choose technology decision for your firm.

Step 2. On the top right part of the page, click "Calculate" and your pollution, total pollution and your profit of "Not adopt" and "Adopt" will be shown to you in the table right under the "Calculate" button.

Keep in mind that the decisions you make in the decision calculator are for informational purposes only and other firms can make their own decisions regardless of what you choose for them.

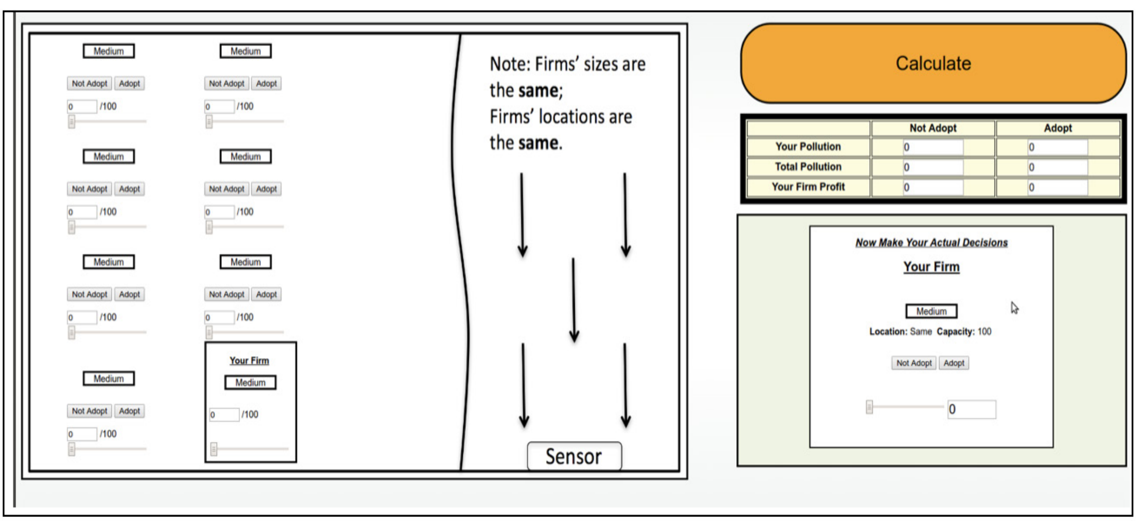
After you decide what your decision will be, make your actual decision in Step 3.

Step 3. On the bottom right part of the page, choose your actual production decision with the slider, and pick your actual technology decision. When you are done, click "Confirm". Once you have clicked this button, the button will turn gray and it is no longer possible to change your decisions for that round.

Results—While you are waiting for the other players to make their decisions, you can review the results of past rounds, which will be shown on your screen. After all eight players have clicked the Confirm button, the results of the current round will appear, including Your Pollution, the Total Pollution from all members of your group, your Production Income, and Your Firm Profit.

DECISION CALCULATOR

The images below are examples of the interactive Decision Calculator that you will use on your computer.



Pollution Table

This Pollution Table helps you to better understand how your firm's production decision, technology decision and location affect your pollution. Use this table along with the Decision Calculator to help you make more informed decisions.

How to read this table?

1. The first column (Production) indicates how much is being produced.
2. Find where your firm is located from the Decision Calculator. If every firm is in the same region, use the last two columns (marked as "Same Region").
3. Your firm's pollution for each level of production under "Not Adopt" and "Adopt" are listed in the columns corresponding to your region.

	Your Firm Pollution									
Production	**Region 1**		**Region 2**		**Region 3**		**Region 4**		**Same Region**	
	Not Adopt	Adopt	Not Adopt	Adopt	Not Adopt	Adopt	Not Adopt	Adopt	Not Adopt	Adopt
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	1.20	0.60	1.40	0.70	1.60	0.80	1.80	0.90	1.50	0.75
10	2.40	1.20	2.80	1.40	3.20	1.60	3.60	1.80	3.00	1.50
15	3.60	1.80	4.20	2.10	4.80	2.40	5.40	2.70	4.50	2.25
20	4.80	2.40	5.60	2.80	6.40	3.20	7.20	3.60	6.00	3.00
25	6.00	3.00	7.00	3.50	8.00	4.00	9.00	4.50	7.50	3.75
30	7.20	3.60	8.40	4.20	9.60	4.80	10.80	5.40	9.00	4.50
35	8.40	4.20	9.80	4.90	11.20	5.60	12.60	6.30	10.50	5.25
40	9.60	4.80	11.20	5.60	12.80	6.40	14.40	7.20	12.00	6.00
45	10.80	5.40	12.60	6.30	14.40	7.20	16.20	8.10	13.50	6.75
50	12.00	6.00	14.00	7.00	16.00	8.00	18.00	9.00	15.00	7.50
55	13.20	6.60	15.40	7.70	17.60	8.80	19.80	9.90	16.50	8.25
60	14.40	7.20	16.80	8.40	19.20	9.60	21.60	10.80	18.00	9.00
65	15.60	7.80	18.20	9.10	20.80	10.40	23.40	11.70	19.50	9.75
70	16.80	8.40	19.60	9.80	22.40	11.20	25.20	12.60	21.00	10.50
75	18.00	9.00	21.00	10.50	24.00	12.00	27.00	13.50	22.50	11.25
80	19.20	9.60	22.40	11.20	25.60	12.80	28.80	14.40	24.00	12.00
85	20.40	10.20	23.80	11.90	27.20	13.60	30.60	15.30	25.50	12.75
90	21.60	10.80	25.20	12.60	28.80	14.40	32.40	16.20	27.00	13.50
95	22.80	11.40	26.60	13.30	30.40	15.20	34.20	17.10	28.50	14.25
100	24.00	12.00	28.00	14.00	32.00	16.00	36.00	18.00	30.00	15.00
105	25.20	12.60	29.40	14.70	33.60	16.80	37.80	18.90	31.50	15.75
110	26.40	13.20	30.80	15.40	35.20	17.60	39.60	19.80	33.00	16.50
115	27.60	13.80	32.20	16.10	36.80	18.40	41.40	20.70	34.50	17.25
120	28.80	14.40	33.60	16.80	38.40	19.20	43.20	21.60	36.00	18.00
125	30.00	15.00	35.00	17.50	40.00	20.00	45.00	22.50	37.50	18.75

For Example:

1. A firm in Region 1, producing 75 units. Firm Pollution for not adopt: 18; adopt: 9.
2. A firm in Region 4, producing 75 units. Firm Pollution for not adopt: 27, adopt: 13.5.
3. A firm in Same Region, producing 100 units. Firm Pollution for not adopt: 30; adopt: 15.

ID# _________

UNDERSTANDING THE EXPERIMENT

This short exercise is designed to help you understand how the experiment works. The profit you earn in this section does not affect your real earnings.

Please use the decision calculator on the computer in front of you to figure out what your firm profit will be under the following scenarios:

You will be guided through Scenario A, and you will complete scenario B by yourself.

Scenario A:

Please fill in your profit for the following hypothetical decisions. The steps listed below will guide you through scenario A.

Everyone else		You		
Technology	**Production**	**Your Production**	**Your Technology**	**Your Profit**
Not Adopt	80	50	Not Adopt	
Not Adopt	80	50	Adopt	

Step 1: On the left part of the page, select "Not Adopt" for everyone else except your firm.

Step 2: Use the slider or type in the boxes to change everyone else's production to 80 units.

Step 3: Still on the left part of the page, find the box that lists "Your Firm", change the production decision to 50 units.

Step 4: Click "Calculate". Your pollution, total pollution and your firm profit should be shown to you.

Step 5: Find "Your Firm Profit" for "Not Adopt", which should be "33.75" in this case. Type in "33.75" in the first row under profit for scenario A.

Step 6: Find "Your Firm Profit" for "Adopt", which should be "25.55" in this case. Type in "25.55" in the second row under profit for scenario A.

Step 7: Click "Check answer for scenario A" when you are done. If the program asks you to try again, please check answers for the highlighted parts.

Now please complete scenario B on your own, please raise your hand if you have any questions.

Scenario B:

Please fill in your profit for the following hypothetical decisions on the computer screen.

Everyone else Technology	**Every else Production**	**Your Production**	**Your Technology**	**Your Profit**
Not Adopt	80	50	Not Adopt	
Not Adopt	80	50	Adopt	
Not Adopt	80	80	Not Adopt	
Not Adopt	80	80	Adopt	

Everyone else		You		
Technology	**Production**	**Your Production**	**Your Technology**	**Your Profit**
Adopt	100	100	Not Adopt	
Adopt	100	100	Adopt	

You may refer to instructions for Scenario A to help you complete Scenario B.

Input your firm profit for Scenario B on the computer program and check if it is correct by clicking "check answers". When the program asks you to "try again", it means your answer is not correct and will be highlighted. In that case, please use the calculator to recalculate the answer.

When you get both scenarios correct, you may click the continue button to move on to the next part.

INSTRUCTIONS FOR PRACTICE

You will now play five practice rounds to learn how the experiment works. The outcomes of these rounds will not affect your cash earnings.

In each round of this part, you will make your Production Decision and your Technology Decision. Use the Decision Calculator to see how your decision and others' decisions affect your earnings.

In this practice part, pollution does not affect firm profits. The more you produce, the more your firm profit will be.

After everyone makes their decisions, you will see the results screen that will display your Firm Profit and Pollution. In this part, your Firm Profit will be calculated as follows:

Firm Profit = Production Income.

MOVING on to PART 1 through PART 8

After you have finished the practice rounds, you will participate in Part 1 through Part 8 of the experiment. In these parts, the experimental dollars you earn from your firm's profits in each round will affect your cash earnings.

In each round of Part 1 through Part 8, you will make a Production Decision and a Technology Decision. Groups will be randomly reassigned after each part.

INSTRUCTIONS FOR PART 1-4

1. In these parts, your Firm Profit only depends on your production and technology decisions; the production and pollution generated by other firms do not affect your Firm Profit.

2. Note that the location and capacity of firms may or may not be different. The capacity of each firm is shown on the calculator. When firms have different locations, the region will be divided in 4 sub-regions by solid lines; when firms have the same location, the region will not be divided. Refer to the **Pollution Table** to see how location influences pollution. We will indicate each scenario at the beginning of each part.

3. Use the **Decision Calculator** to make more informed decisions. Although the results are for informational purposes only, the location and capacity of each firm is the same as the real decisions.

4. To make your actual decision for this round, choose a Production Decision and a Technology Decision. Once done, click "Confirm".

5. In these parts, pollution does not affect firm profits. The more you produce, the more your firm profit will be.

In these parts: **Firm Profit = Production Income**

INSTRUCTIONS FOR PART 5-8

In these parts, an **environmental regulator** has set a **target total pollution level**. There will be a tax or subsidy based on the total pollution of your firm compared with the target level. The target will change between parts and the specific value will be shown to you.

Your profit will be adjusted by a tax or subsidy (from here on referred to as **tax/subsidy**). This tax/subsidy can be either negative (a tax) or positive (a subsidy) and is determined based on how much pollution is in the river relative to the **Target** determined by the regulator. The pollution level in the river is the aggregation of pollution from all firms. There will be a subsidy for zero concentration, but the amount of subsidy gets smaller as concentration increases. If the measured concentration level is exactly the same as the target, there will be neither a tax nor a subsidy. As concentration increases beyond the target, the tax gets larger.

Pollution in one round does not affect pollution in other rounds. However, at the end of the experiment, your earnings will be the sum of the profits you earned from all of the rounds.

In each round, you will make a Production Decision and a Technology Decision. **Total Pollution** in your group affects the profits of firms in your group.

The **Tax Payment for each firm** in your group is calculated as follows:

Total Pollution ≤ Target	Subsidy Received = Target – Total Pollution
Total Pollution > Target	Tax Payment = Total Pollution – Target

For example, if the target is set at 60, then

- If the Total Pollution in your group is less than or equal to 60, each firm in your group receives 1 experimental dollar in subsidy for every unit of total pollution under 60 units.
- If the Total Pollution in your group is greater than 60, each firm pays 1 experimental dollar in taxes for every unit of total pollution above 60 units.

The amount of the Tax/Subsidy Payment is determined by decisions of everyone in your group. Your Firm Profit in these parts will be calculated as:

If Total Pollution ≤ Target,
Firm Profit = Production Income + Subsidy Payment

If Total Pollution > Target,
Firm Profit = Production Income − Tax Payment

<u>Use the Decision Calculator to help you make more informed decisions, otherwise, you may lose money. Note that in these parts, it is not true that the more you produce, the more profit you will get.</u>

Citizen-centric Government Services Ecosystem: An Agent-based Approach to the Theory of Planned Behavior

James R. Thompson, Trevor Caskey, Alison Dingwall, Zoe Henscheid[A]

Abstract

To illustrate the potential benefits and challenges of modeling theories of human behavior in the face of messaging campaigns, we employ an agent-based model that captures individual decision-making of agents interacting with social and environmental pressures. Messaging campaigns are often based on demographic or regional data with the assumption that all members of a given cohort will behave the same. Conversely, Ajzen's Theory of Planned Behavior (TPB) suggests that behavioral intent is driven by internal attitude, perception of social normative pressures, and perception of control, which are generally heterogeneous across a given population. We present a general framework for modeling individual behavioral intent as a function of these factors and embed the framework in a model of the environment that both pressures agents to act and responds to the actions they take. We analyze the effects of message timing on average behavioral intent, as well as the impact that influential figures have when speaking out in favor of the messaging campaign. The model provides a baseline for operationalizing the TPB in a format that can be replicated and expanded upon in future research.

Keywords: theory of planned behavior, agent-based modeling, computational social science, behavioral intent

---

[A] MITRE Corporation

doi: 10.18278/jpcs.3.2.11

摘要

为阐明人类在面临信息传递活动（messaging campaigns）时的行为理论以及该理论建模的潜在挑战和益处，本文运用一项基于主体的建模（agent-based model，简称ABM），捕捉各主体在面临社会和环境压力时所做的个人决策。信息传递活动时常基于人口或地区数据，并假设既定人群都会有相同的行为表现。相反，阿耶兹（Ajzen）的计划行为理论认为，行为意图受内在态度、社会规范压力感知和控制感知三种因素所驱使，而这三种因素在既定人口中一般都是不同的。本文为个人行为意图的建模提供了一般框架，将其作为这些因素发挥的作用。同时将框架嵌入特殊环境模型，在该环境下，主体对压力所采取的行动以及该行动的回应都有所体现。本文分析了信息时刻（message timing）对平均行为意图的影响，以及当影响力人物公开声明支持信息传递活动时所产生的影响。模型为计划行为理论的操作提供了基准线，该操作方式可以被复制，并在今后研究中进行扩大。

关键词：计划行为理论，基于主体建模，计算社会科学，行为意图

Resumen

Para ilustrar los beneficios y retos de la modelización de teorías de comportamiento humano en cara a las campañas de mensajes, empleamos un modelo basado en agentes que captura la toma de decisiones individual de los agentes que interactúan con presión social y ambiental. Las campañas de mensajes están frecuentemente basadas en datos demográficos o regionales bajo la premisa que todos los miembros de un determinado grupo se comportarán de la misma manera. En cambio, la Teoría del Comportamiento Planificado de Ajzen sugiere que la intención conductual está impulsada o la actitud interna, la percepción de la presión normativa y la percepción del control, que son generalmente heterogéneas en una población determinada. Presentamos un marco general para la modelización del comportamiento individual como una función de estos factores e insertamos el marco en un modelo del ambiente que ejerce presión para que los agentes actúen y responde a las acciones que toman. Analizamos los efectos del cronometraje de mensajes en la intención conductual promedio, así como el impacto que las figuras influenciales tienen cuando están hablando a favor de la campaña de mensajes. El modelo proporciona una base para operacionalizar la teoría del comportamiento planificado en un formato que puede ser replicado y expandido en investigaciones futuras.

Palabras clave: teoría de comportamiento planificado, modelización basada en agentes, ciencia social computacional, intención conductual

Introduction

The field of social science attempts to understand and explain society as a function of the behavior people who constitute the society in question. To draw an analogy to mathematics, we may be tempted to model such a system with the central limit theorem, where the sum of many independent identically distributed random variables tends toward a normal distribution. Unfortunately, such an approach to social science is doomed from the start as people rarely make our choices independently of our social networks and we most certainly are not identical in all respects. These two assumptions that make central limit theorems possible are thus immediately and aggressively violated by human behavior and the societies we build, resulting in complex dynamical systems that seem unpredictable or even chaotic. The challenge to social scientists is to construct theories and models that are simple enough to be controlled in experimental settings and yet complex enough to capture the dependent, heterogeneous nature of human behavior.

From a practical perspective, the goal of many such models is to understand the actions that society might take in response to some stimulus as demonstrated in Epstein and Axtell's work on growing artificial societies (Epstein & Axtell, 1996). Governments, in particular, wish their citizens to engage in the political process, participate in national programs, and comply with laws and policies. In current practice, persuasive messaging campaigns designed to elicit such responses are often based on demographic or regional data that result in the use of static approaches to advertising and influencing behaviors (Gupta & Chintagunta, 1994; Kalyanam & Putler, 1997). These static approaches necessarily assume that society is homogeneous within the given region or demographic and thus fail to capture the complexity and heterogeneity of real societies.

In this paper, we offer a preliminary step toward modeling how a heterogeneous society might respond to a persuasive messaging campaign. We apply an agent-based model to the question of how persuasive messaging impacts behavior and behavioral intent of a given population. The agents in our model make decisions that are based on a collection of theoretical frameworks of social psychology. The first of these is Ajzen's Theory of Planned Behavior (TPB) that postulates *intention* is an indication of an individual's readiness to perform the behavior of interest. Intention is believed to be determined by one's attitude toward the behavior, one's perceived normative pressure (from their social connections) to engage in the behavior, and one's perception of one's own ability to perform the behavior (Ajzen, 1991). However, TPB makes no statements about how intention may change over time. It simply states that intention can be derived from a set of behavioral, normative, and control beliefs. The formation and change of these beliefs depend on our individual experiences, relationships, and exposure to various environmental pressures. To incorporate TPB into a set of virtual agents, we must embed the theory into a framework that allows agents to experience and learn from their environment. To accomplish this we incorporate a cognitive processing model put forth by Jager, Janssen, and Vlek (1999) and a conditioning model adapted from the approach used by Epstein in Agent Zero (Epstein, 2014).

The remainder of this article is organized as follows. In the literature review, we review the literature relevant to TPB and the additional cognitive frameworks we employ. We formally define the TPB components of attitude, subjective norm (SN), and perceived control, and address the theory's strengths and shortcomings. We also describe the necessary step of placing TPB in a broader framework that can evolve over time and subsequently introduce the cognitive processing model originally described by Janssen et al. (1999). We give a detailed statement of our agents and the modeling environment in which they interact in the model. We describe the preliminary experiments we performed using the model and the results they produced in the results and analysis. We conclude with a discussion of our findings and avenues for future work in the conclusion.

Literature Review

Agent-Based Models of Social Behavior

Agent-based models have been successfully employed to study sociology and social psychology in a number of settings. Macy and Willer (2002) review the different social processes that have been modeled with an agent-based approach, such as social convergence and divergence, social influence, diffusion of innovation, collective action, trust, and cooperation. Smith and Conrey (2007) provide an overview of how agent-based models have been used in social psychology for building theories regarding dynamic, complex, interactive social processes. They discuss agent-based applications such as attitude polarization in groups, escalation of intergroup conflict, formation of stereotypes, aggression, and unhealthy behaviors.

Agent-based models are bottom-up computational approaches for understanding macroscopic social structures and group behavior based on the local interactions of individual agents with bounded rationality that use simple rules to operate and adapt within a spatial environment. This is different from micro-simulations that have a top-down approach by modeling aggregate behavior through equations and statistical estimates of aggregate data, and which do not represent the heterogeneity of individuals nor the spatial environment as distinct from the agent population.

Thomas Schelling is considered by many to be a pioneer in applying an agent-based approach to model social science. His simple spatial model of neighborhood composition, which employed agents with simple rules for preferring a small proportion of their neighbors to be "like themselves," produced segregated neighborhoods (Schelling, 1971). Schelling (2006) expanded on this work and provided much thought leadership on how simple models of individual behaviors can lead to social complexity and aggregate behavior. This work has been leveraged by many social scientists.

Improvements in computing power in the late 1980s to early 1990s provided an avenue for creating computer models that could capture heterogeneity, which is a key feature of agents in agent-based modeling. In 1996, Epstein and Axtell published their seminal work on Growing Artificial Societies that laid a foundation for generative social science using agent-based models. In their book, Epstein and Axtell (1996) describe how agent-based models can be used to study social phenomena, including trade, migration, group formation, cultural transmission, disease propagation, and population dynamics. Miller and Page (2009) expanded on this effort by modeling complex adaptive social systems and introducing concepts such as emergence, organization, decentralization, robustness, adaptation, and feedback into their models. They demonstrate how agent-based models can be used to explore these complex social systems and processes. In his review on the varied motivations for using agent-based models in the social sciences, Axtell (2000) describes how agent-based models allow for a better understanding of intractable abstract formulations of social processes particularly those involving individuals changing state over time. Axtell (2000) concludes that since agent-based computational models are powerful analytic tools with effective presentation technologies, these models may someday be the first line of attack on new problems, with mathematics used later "only to 'tidy up' what the agent model has clearly presented to [be] a robust feature of the problem."

Based on all this past work, we deemed an agent-based model to be an ideal way to explore the population-level impact of different outreach programs on citizen response. We chose the TPB put forth by Ajzen (1991), which we further describe in the section the TPB, as a basis for this exploration as it theorizes that an individuals' intention to perform a behavior is a function of their attitude toward the behavior, their perceived normative pressure to perform the behavior, and their perceived behavioral control (PBC). Agent-based models have been used quite often in modeling TPB; however, these models have modeled TPB for very specific behavioral applications. For example, Zhang and Nuttall (2007, 2011a, 2011b) employed agent-based models (ABMs) and TPB to explore consumer purchasing behavior and smart metering technology adoption. Schwarz and Ernst (2009) used household lifestyle data and an ABM of TPB to explore the social diffusion of water-saving innovations. Richetin et al. (2010) applied both agent-based modeling and models of goal-directed behavior with TPB to analyze the utility of both types of models in understanding the consumption of soft drinks and the engagement in physical activity. They reached the conclusion that the agent-based modeling technique when used to define agents as psychological processes was more realistic and produced statistically better results than using goal-directed behavior models to study intra-individual decision-making processes. Ceschi, Dorofeeva, Sartori, Dickert, and Scalco (2015) applied an agent-based model of TPB to examine householders' recycling attitudes. Other applications of agent-based models of TPB cover a large range of human behavior, including diffusion of organic farming practices (Kaufmann, Stagl, & Franks, 2009), child maltreatment (Hu & Puddy, 2011), tobacco use (Verzi, Brodsky, Brown, Apelberg, & Rostron, 2012), and energy use (Mogles, Ramallo-González, & Gabe-Thomas, 2015).

There has been some recent criticism of the TPB (Sniehotta, 2009; Sniehotta, Presseau, & Araújo-Soares, 2014) that suggest that TPB was not intended to serve as a model of behavior change. With this in mind, and unlike models that have used TPB for studying specific behaviors, our intent is to build a generalized agent-based framework for using TPB to explore any type of citizen behavioral response to various outreach programs. The remainder of this paper describes our approach.

The Theory of Planned Behavior

The TPB was put forth by Ajzen to extend the Theory of Reasoned Action (TRA) (Fishbein & Ajzen, 1975). TRA posited that an individual's intention to perform a behavior is a function of one's attitude toward the behavior and one's perceived SN regarding the behavior. To this framework, TPB adds the third dimension of perceived control to perform the behavior. The concept of PBC is similar to the idea of self-efficacy; that is, an individual's perception of his or her ability to perform the behavior. In theory, all three components of behavioral intent are founded on an underlying set of beliefs. Individuals associate different outcomes with performing behaviors and weight those outcomes as either positive or negative in some way. Individuals also believe that people in their social network—their social referents—have expectations for the their behavior and the individual associates some motivation to comply (or not comply) with those referents' expected norms. Last, individuals hold beliefs about the presence or absence of factors that either facilitate or impede the performance of the behavior in question. Depending on how all of these beliefs come together, people formulate some level of behavioral intent that Ajzen contends is indicative of behavioral action (see Figure 1).

Formally, TPB is constructed as follows. An individual's behavior can be approximated by one's behavioral intent *BI*, which is defined as the sum of one's attitude (*AB*) toward the behavior, the one feels to perform the behavior, and one's over the behavior.

$$B \approx BI = AB + SN + PBC \tag{1}$$

Each of the three components are in turn functions of the individual's underlying beliefs relative to the behavior in question. That is, individual j's attitude toward behavior α is given by the sum product

$$AB_j^\alpha = \sum_{o=1}^{O} b_{j,o} e_{j,o} \tag{2}$$

where $b_{j,o}$ is the strength of j's belief in outcome and is j's evaluation of outcome *0*. The overall attitude toward the behavior

AB_i^α

is the sum of this product over all outcomes

$$o \in \{1, \ldots, O\}$$

associated with behavior. Simultaneously, individual j's SN regarding behavior α is given by

$$SN_j^\alpha = \sum_{r=1}^{R} n_{j,r} m_{j,r} \tag{3}$$

where $n_{j,r}$ is the individual j's belief that referent r wishes j to perform behavior α ,and is individual 's motivation to comply with referent r. The overall subjective norm is the sum of this product over all the rele-

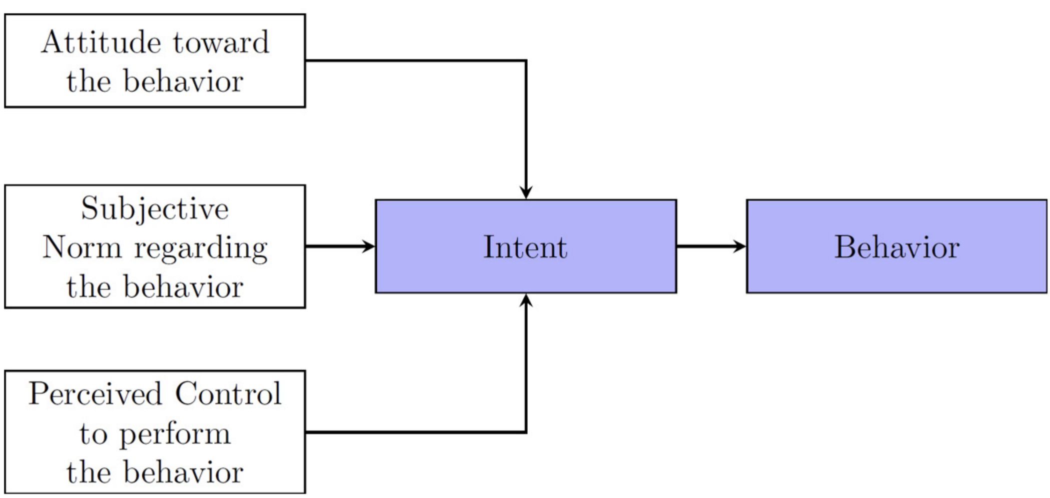


Figure 1. Components of the TPB.

vant referents in the individual's network. Last, the perceived behavioral control over behavior is given by

$$PBC_j^{\alpha} = \sum_{f=1}^{F} c_{j,f} p_{j,f} \quad (4)$$

where $c_{j,f}$ is individual j's belief that a control factor f is present and $p_{j,f}$ is the degree to which control factor f facilitates or impedes the performance of behavior α. The overall perception of control is then the sum over all factors $f \in \{1, \ldots, F\}$.

TPB has been widely used in areas such as health behavior and consumer behavior, where studies and surveys are constructed to elicit the three components of the behavioral intent in order to predict actual behavioral action (Carmack & Lewis-Moss, 2009; Pavlou & Chai, 2002). Godin and Kok found in their meta-analysis that TPB produced an average R^2 value of 0.41 when used to explain and predict behavioral intention (Godin & Kok, 1996). They also noted significant variation in the efficiency of the theory when compared across different categories of behavior related to health choices. This variation in results underscores the limitations inherent in assuming a population is homogeneous.

Note that there is no expression of time in the above formulation of TPB. Ajzen addressed the issue of predicting behavioral action with TPB by stating that the interval between measuring the components of TPB and observing the opportunity for action needs to be sufficiently short such that all the foundational beliefs remain stable (Ajzen, 1991). Ajzen went on to say that persuasive messaging has been correlated with changes in attitude, indicating that a reasonable assumption is that such messaging would have a similar relationship with SN and perceived behavioral control if the message content was directed at the normative beliefs or levers of control. So, although TPB does not elaborate on how beliefs, and therefore behavioral intention, change over time, it does allow that behavioral intention can—and most likely will—change as the individual refines their beliefs. Thus, in order to model how persuasive messaging affects the macrolevel response of a sample population, we must place TPB into a larger framework that evolves over time.

The Conceptual Model of Consumer Behavior

Janssen et al. (1999) investigated the well-known commons dilemma using an agent-based model of consumer behavior. Although behavioral actions related to persuasive messaging are a different environmental context from competition for common resources, we contend that the Janssen et al. conceptual model of consumer behavior is transferable to multiple environments of decision-making. The consumer behavior model is a multi-theoretical framework that draws on theories of human needs, motivational processes, operant conditioning, classical conditioning, decision theory, and TPB.

The consumer behavior framework builds on the idea that individuals exist in an environment where multiple systems interact and affect cognitive processing. There are four systems connected as shown in Figure 2. The Pressure system applies forces to the needs of an individual. For example, a down-turn in the economy or the decreased availability of a resource will exert some pressure on individuals in a society. These pressures are interpreted by the individual in the State system, which contains two dimensions of cognitive processing described in greater detail below. The result of cognitive processing is some action (or inaction) that is carried out through the Impact sys-

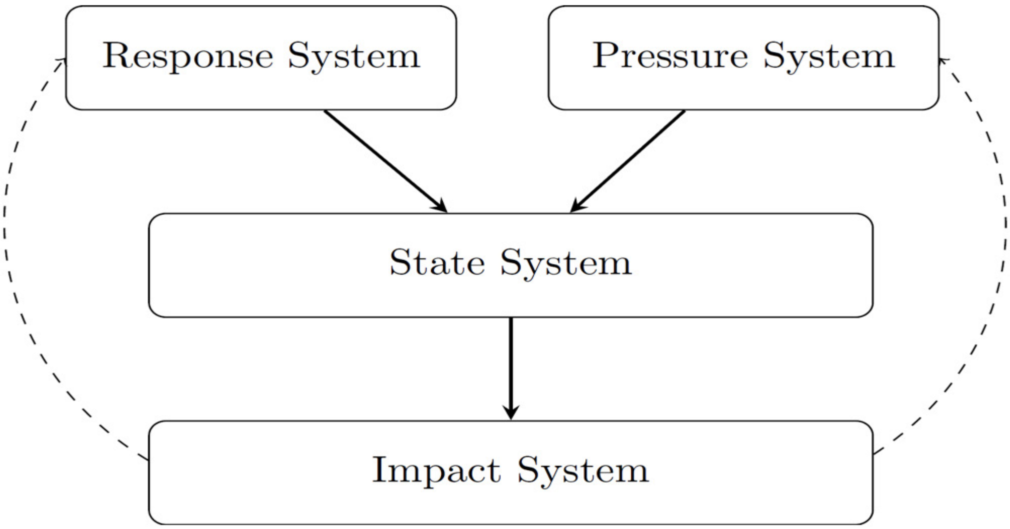


Figure 2. Conceptual model of consumer behavior system (Jager, Janssen, & Vlek, 1999).

tem. As many individuals go through this process, the Response system assesses the collective impact of behavioral actions and attempts to influence future behaviors to either change or continue in a manner that produces a favorable response. Figure 2 illustrates the closed-loop, iterative nature of this system.

At the core of the conceptual model is the individual's State system and Jager et al. (1999) employ two theoretical dimensions of cognitive processing in this system. The first is the degree to which an individual needs to elaborate on a decision. Reasoned behavior is associated with a high need to elaborate, while automatic behaviors are associated with lower needs to elaborate. The second dimension they use is the degree to which the cognitive process is social versus individually oriented. That is, social processing is more likely to be used when a behavioral action is visible to the public and the individual has a high degree of uncertainty about the outcomes or consequences of an action. Conversely, if the individual is certain about the consequences and the behavioral action is more private in nature, the individual will likely make a decision with less input from one's social network.

These two dimensions are assumed to be orthogonal and the intersection of the two produces four quadrants of decision-making as depicted in Figure 3. When an individual has a high need for cognitive elaboration about a behavioral action and she feels certain about the outcomes associated with the behavior, she is more likely to be individually oriented and is said to be *deliberative*. If one is uncertain about the outcomes associated with the behavior then one is more likely to be socially oriented and said to be *social comparative*. On the other hand, when the individual has a low need for cognitive elaboration about a behavioral action, the decision is more automatic. If one is certain about the outcomes, one is said to be *repeating* the behavior and if one is uncertain about the outcomes one is said to be *imitative*.

As discussed, the TPB has been employed to predict behavioral action based on the assessment of behavioral intent in a

Figure 3. The dimensions of cognitive processing.

number of domains, such as health-related behaviors and consumer decision-making. However, there is significant variation in the predictive power of TPB from one particular behavioral decision to the next, and there are a number of criticisms of TPB. One of the criticisms involves the concept of "inclined abstainers," who form what is measured to be an intention to act, but then subsequently do not act (Sniehotta et al., 2014). Central to this practical problem with TPB is the passage of time in an ever-changing environment. TPB was meant to equate intention to act with behavioral action in a relatively short interval of time. Although Sniehotta et al. advocate abandoning TPB altogether, we contend that when embedded in a larger dynamic framework, like the one set forth by Janssen et al., the usefulness of TPB becomes more apparent. Specifically, it provides us with a conceptual model for simulating the potential impact of persuasive messaging on behavioral intent as it evolves over time. The presence of inclined abstainers, in turn, gives us a potential measure of the threshold behavioral intent must reach in order to translate into behavioral action. In the next section, we describe how we combined these two conceptual models into a preliminary step toward simulating the collective response to persuasive messaging campaigns.

The Model

Model Components

Our objective is to create a dynamic model that can ultimately be used to demonstrate the relative effectiveness of a planned persuasive messaging campaign. As such, we require a messenger system that is representative of advertising media and a population of virtual agents who receive periodic messages from this system. Additionally, we would like to model different strategies of persuasive messaging. One such strategy often entails the use of celebrities or community leaders who speak out in support of the campaign's interests. These influential figures often have much larger social networks (e.g., Twitter followers) than normal. This also implies that our model needs a network structure for connecting individuals both to each other and to the messaging system. Finally, the agents are being encouraged by the persuasive messaging campaign to take some sort of action. When their behavioral intent reaches some threshold for action, they need a means by which to perform the action in question. All of these components taken together create our Pressure, State, Impact, and Response systems.

Adjusting the Theory of Planned Behavior: The State System

To incorporate TPB into an agent-based model, we embed the model into the cognitive processing framework established by Janssen et al. The driving factor for combining TPB with a dynamic framework is to add the index of time (t) to behavioral intent. This implies a mapping from $BI \to BI(t)$. This mapping is relatively simple when the individual has a low need for cognitive elaboration. Either the individual is certain of the outcomes of a given action one considers to be private and therefore repeats one's previous intent

$BI(t)=BI(t-1)$

or one is uncertain about the action and therefore imitates one's social network

$BI(t)=f(SN, PBC, t)$,

where $f()$ is a function of SN, PBC, and time t, but not AB.

The mapping is more complicated when the individual has a high need for cognitive elaboration. As previously discussed, the individual is now either in a deliberative state or a social comparative state. Ideally there should be a continuum between these states rather than a discrete determination of one or the other. Indeed, the imitative state is the extreme end of social comparison and the repetitive state is a logical evolution of successful deliberation. Thus, we combine the components of TPB as follows.

$$BI(t)=[\phi AB(t)+(1-\phi)SN(t)]PBC(t), \tag{5}$$

where $\phi \in [0,1]$ is derived from the angle θ as $\phi = \frac{1}{\pi}\theta$ in the cognitive processing plane (see Figure 4).

Note from Figure 4 and Equation (5) that if $\phi = 0$ then AB plays no role in the formation of $BI(t)$ and the individual is essentially in the imitative state. Conversely, if $\phi = 1$ then SN plays no role in the formation of $BI(t)$ and the individual is at the extreme of the deliberative state. Assuming further that $AB(t) = AB(t-1)$ and $PBC(t) = PBC(t-1)$, then the individual is in the repetitive state where $BI(t) = BI(t-1)$.

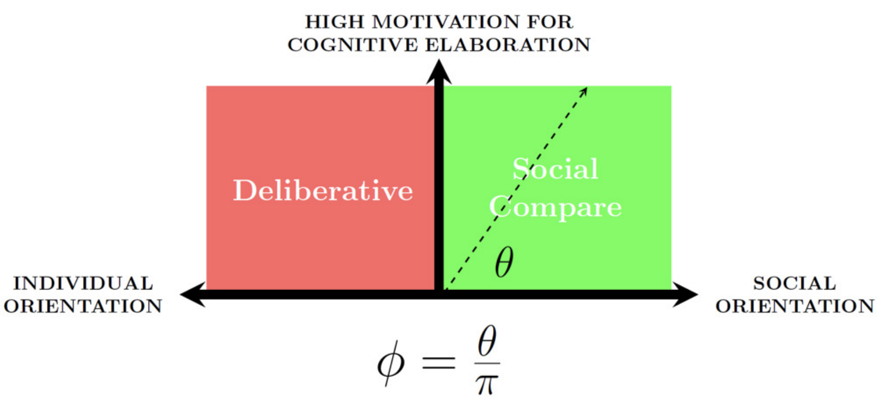


Figure 4. The definition of ϕ and θ.

This formulation provides two advantages. First, it reduces the weights on attitude (*AB*) and *SN* to a single parameter ϕ. And second, it explicitly ties those weights to the cognitive processing framework in a logical manner. As the Pressure and Response systems provide input and feedback to the individual, their theoretical position in the cognitive plane—as measured by θ—changes the level of importance the individual places on one's internal attitude and one's perception of the external SN.

The next step is then to adjust *PBC* in a manner that complements this formulation. The idea of perceived behavioral control came about as an addition to the TRA because its role is fundamentally different from *AB* and *SN*. That is, an individual who believes they have a high degree of control over a particular behavior does not necessarily perform that behavior on the basis of control alone. Conversely, if the individual believes she has no control over factors that facilitate a behavior, she may well abandon any attempt to perform the behavior even if the outcomes are desirable and her SN is urging her to take action. In this context, impedes action when it is low, *PBC* but when it is high it facilitates action only by not obstructing it. Under this set of simplifying assumptions, we restrict *PBC*(*t*) to the closed interval [0,1] and multiply it by the weighted sum of *AB*(*t*) and *SN*(*t*). The more factors the individual believes to be present that facilitate action, the closer *PBC*(*t*) gets to 1. If *AB*(*t*) and *SN*(*t*) are also high, then the individual will have a high value of *BI*(*t*) and may take action. Conversely, as *PBC*(*t*) approaches 0 the individual's *PBC*(*t*) will approach 0, albeit at a slower rate for higher values of *AB*(*t*) and *SN*(*t*). Given this adjusted formulation of TPB, the next section describes how each component of TPB changes in a dynamic environment.

Dynamic Components of TPB

Each component of TPB is the sum product of a set of beliefs and an associated weight or evaluation. In the case of attitude (*AB*), the belief is defined as the subjective probability that a particular outcome is associated with a behavioral action. The evaluation is an indication of how desirable (or undesirable) the outcome is to the individual. This relatively simple structure still allows for considerable complexity since multiple outcomes—some good and some bad—could be associated with a given behavior. For example, one may positively associate the outcome of improved health with exercise and may simultaneously negatively associate the outcome of exhaustion with exercise. The individual's attitude toward exercising comes down to which belief-evaluation product outweighs the other. Additionally, there is evidence that the salience of our beliefs and evaluations decay over time if they are not continually reinforced (Epstein, 2014; Rescorla & Wagner, 1972). As such, we adopt a variant of the Rescorla–Wagner equation for the belief and the evaluation. The Pressure and Response systems send signals to each agent in the model that either reinforces the belief or the evaluation component or both. Each agent then updates its *AB*(*t*) according to Equation (6).

$$AB_j^{\alpha} = \sum_{o=1}^{O}(1-\exp^{-\lambda_j b_o})(2(1-\exp^{-\lambda_j e_o})-1), \tag{6}$$

where b_o is the amount of reinforcement directed at the belief that outcome *O* is associated with behavior α and e_o is the amount of reinforcement directed at the evaluation of outcome *O*. The parameter λ_j determines the mindset of the individual agent. For higher (lower) values of λ_j, the agent needs less (more) reinforcement to reach high values of belief or evaluation. Note also that belief is defined as a subjective probability, so the first term in the sum is restricted to the interval [0,1]. Evaluation can be either good or bad, so the second term is confined to the interval [-1,1].

For subjective norm *SN*(*t*), we proceed in a similar fashion. However, the components are now the belief that a particular referent wishes or expects the individual to perform the behavior in question and the individual's motivation to comply with that particular referent's wishes. From the individual perspective each referent is part of the Pressure system, sending signals about the behavior. It follows that if a referent sends a signal regarding the behavior, then that referent wishes the individual to perform the behavior. The motivation to comply is more complicated. We therefore tie motivation to comply to the individual's change in attitude *AB*(*t*) - *AB*(*t-1*). If the individual's attitude toward the behavior has increased (i.e., become more positive toward the behavior), then the motivation to comply is increased. If the individual's attitude has decreased, the motivation to comply is decreased. As such, *SN*(*t*) is computed as

$$SN_j^{\alpha} = \sum_{r=1}^{R} (1-\exp^{-\gamma_j n_r})(2(1-\exp^{-\gamma_j m_r})-1), \tag{7}$$

where

$$m_r = m_r(1+sgn(AB(t)-AB(t-1))\zeta), \text{for} \zeta \in [0,1]. \tag{8}$$

In Equation (7), n_r is the amount of belief reinforcement sent by referent *r* and m_r is the amount of reinforcement to comply with referent *r*. Equation (8) illustrates that increases or decreases motivation to comply based on the change in attitude toward the behavior since the last time period. The parameter ζ determines how sensitive the individual is to changes in *AB*(*t*) and the parameter γ_j determines how much reinforcement of belief and motivation is required to produce a given amount of change in .

We update perceived behavioral control *PB*(*t*) in much the same way as *AB*(*t*). The Pressure and Response systems send signals about whether a particular factor is present and whether that factor facilitates or impedes the behavior in question. The only difference is that we now restrict *PBC*(*t*) to the interval [0,1] rather than allowing a negative component.

$$PBC_j^{\alpha} = \frac{1}{F}\sum_{f=1}^{F} (1-\exp^{-\lambda_j c_f})(1-\exp^{-\lambda_j p_f}), \tag{9}$$

where c_f is the amount of reinforcement that factor *f* is present and p_f is the amount of reinforcement that factor *f* facilitates the behavior α Here, we have also used the same mindset parameter λ_j as in *AB*(*t*) to represent how much reinforcement is required to change the individual's level of *PBC*(*t*).

The final step in making TPB dynamic is to allow for decay in the salience of all of the agent's belief. Recall that they hold beliefs about outcomes being associated with the behavior, referents wishing them to perform the behavior, and factors being present that facilitate the behavior. Here we make the assumption that individuals possess some degree of inertia in changing their beliefs. That is, if an individual has a mindset that allows them to increase with relatively little reinforcement, then that individual will decay slowly. Conversely, if they require correspondingly more reinforcement for the same marginal increase they will decay faster. Essentially this means that if a person is ready to believe in something, it will be harder to convince them it is not true. On the other hand, if they are skeptical, it will require less time for the salience of recent evidence to dissipate. Thus, if the current, cumulative reinforcement value is given by x then for a given amount of time Δt we decay the belief reinforcement according to

$$x = x\exp^{-(1-\lambda_j)\Delta t}, \quad (10)$$

where λ_j is the mindset parameter for agent j. For $SN(t)$, we replace λ_j in Equation (10) with γ_j. The cumulative reinforcement values are given by the parameters b_o for belief reinforcement in an outcome and n_r for belief reinforcement in what a referent wishes.

The Cognitive Plane

A critical piece of the above formulation is the determination of each agent's current cognitive state. The agent's cognitive state is assumed to be a function of the agent's need for cognitive elaboration, the agent's certainty in its beliefs, and the public or private nature of the behavior in question. We assume that this comes from its most salient evaluation of an outcome associated with the behavior. As such, we query the $AB(t)$ component that has the highest absolute value for $|(2(1-\exp^{-\lambda_j e_o})-1)|$ .If this evaluation is close to zero, we assume the agent has a low need for cognitive elaboration; i.e., the agent is close to indifference regarding the outcome. We then check the strength of the belief $(1-\exp^{-\lambda_j b_o})$ to determine if the agent is in the repetitive state (≥ 0.5) or the imitative state (<0.5). If, on the other hand, the highest absolute evaluation $|(2(1-\exp^{-\lambda_j e_o})-1)|$ is greater than , then we assume this agent has a high need for cognitive elaboration and thus the angle needs to be calculated as shown earlier in Figure 4.

Agent Algorithm

The final algorithm that ties all the parts of the State system together can be summarized as a set of processes and decisions that are evaluated at each time step the agent is activated. Specifically, at each time step, an agent is activated the agent performs the following actions:

1. Decay belief reinforcement according to Equation (10).
2. Process new messages to update *AB* and *PBC* according to Equations (6) and (9).
3. Process referrals from social network to update *SN* according to Equation (7).
4. Update $BI(t)$.
5. Check if action has been previously taken.
6. If no action has been taken, check if *BI*(t) exceeds the threshold for action . If so, attempt action.
7. Otherwise, check if action previously taken was successful. If so, update action status.
8. If previous action was not successful, update *PBC* and decide whether to try again or not
9. Schedule next time step to repeat these steps and stop.

The algorithm is shown graphically in Figure 5.

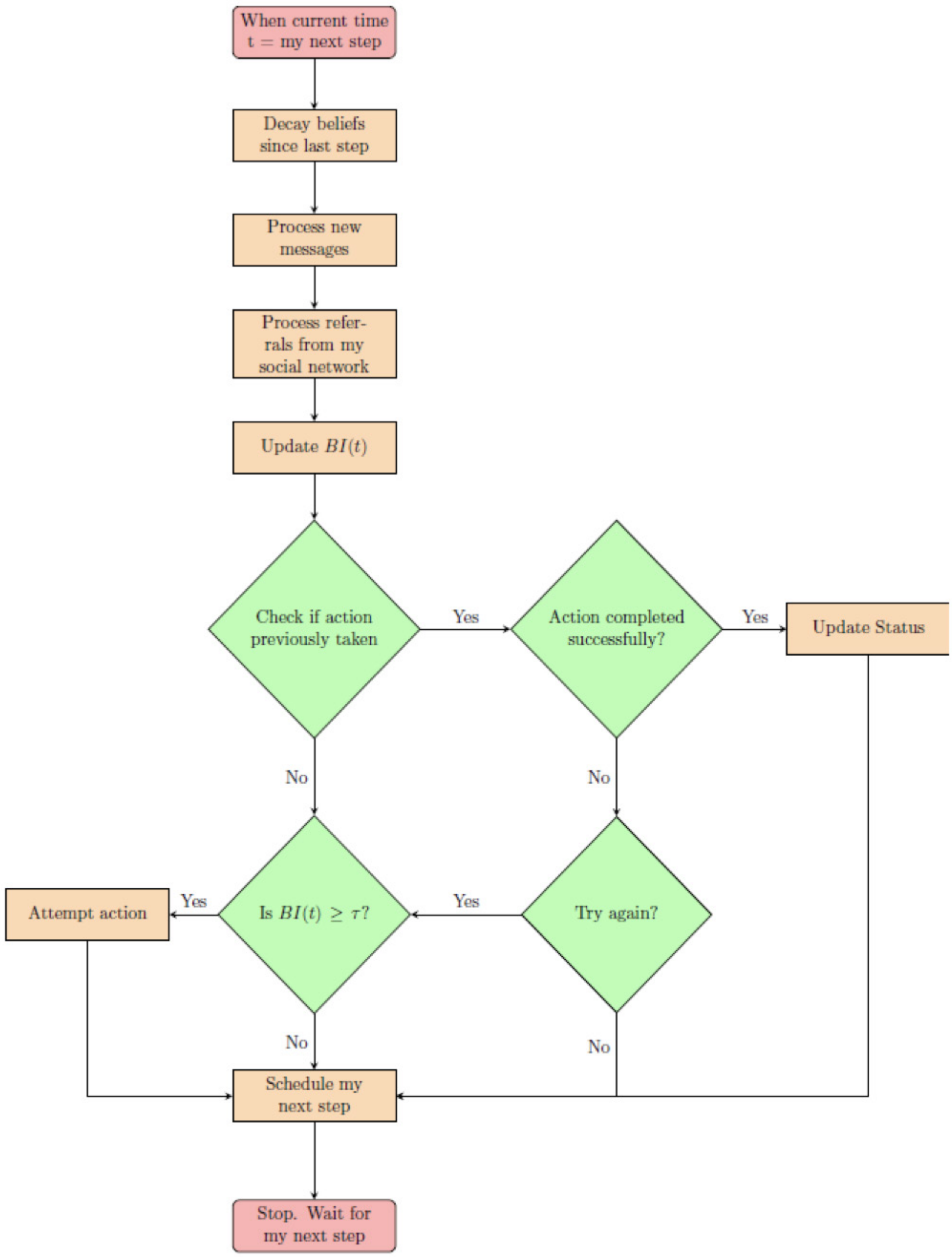


Figure 5. Agent algorithm.

The Pressure, Response, and Impact Systems

The remaining parts of the model define the environment that influences the cognitive processing of the population of agents. Each agent contributes to the Pressure system of other agents by acting as a social referent. When an agent's attitude toward a behavior increases, it sends out referrals to every other agent in its social network. These referrals contain the sending agent's level of belief reinforcement at the time. The receiving agent incorporates the reinforcement into its $SN(t)$ calculation, not its $AB(t)$. Thus, one agent's $AB(t)$ is impacting another agent's $SN(t)$. Note also that the receiving agent may not have a high motivation to comply with the sending agent, so the impact to the receiving agent's $BI(t)$ could be positive or negative. At the start of the model, each agent is randomly assigned a network of other agents and the configuration of this network can be controlled by the user.

The Response system has two components: the messengers and the influential figure. Each messenger is randomly connected to some subset of the population of agents and those agents in turn have a random motivation to comply with the messenger. Information sent by the messenger is incorporated into the receiving agent's $AB(t)$ or $PBC(t)$. The idea being that persuasive messaging seeks to alter an individual's attitude toward a behavior by informing the individual of associated outcomes or factors present or espousing the benefits of those outcomes or factors to increase the individual's evaluation. The user can instantiate as many messengers as desired and change the start time, duration, and rate at which messages are sent. The rate is used to parameterize a Poisson process for generating periodic messages.

The influential figure is an optional addition to the Response system. The influential figure is essentially a super-connected agent that consistently sends out referrals advocating on behalf of the messengers. The user can control what percentage of the population is connected to the influential figure, along with the start and stop times, and the rate at which the influential figure sends referrals. The user can also control the maximum and minimum values for motivation to comply with the influential figure. Since the influential figure is sending referrals, it impacts the receiving agent's $SN(t)$ rather than $AB(t)$ or $PBC(t)$.

The Impact system is simply a queuing service that is only open for a user-specified interval of time. When an agent reaches its threshold for action, it submits a service request that contains a randomly selected expiration date. The Service then checks the queue of service requests periodically and fulfills any requests it removes from the queue while it is open. The queue is first-in-first-out, so fulfillment of requests depends on how other agents are reacting to the environment. Many requests will result in a full queue that the Service may not have enough time to process completely. If the Service is closed when an agent attempts to submit a request or the agent's request expires before it is filled, there is a negative impact on that agent's $PBC(t)$. Conversely, if an agent is able to submit the request and it is subsequently completed, the result has a positive impact on the agent's $PBC(t)$.

Model Propagation and Run Time

A final feature of the model is that the agents can lengthen the run time by continuing to send referrals to their social network after the messaging campaign has stopped. The propagation of referrals and messages is what drives the model. When agents receive a message or a referral, they update their $BI(t)$ and schedule themselves to update again at some random time in the future. If they stop receiving messages and referrals, they will ultimately stop scheduling to update themselves and the simulation will stop. This propagation feature of the model allows us to test the salience of different messaging campaigns, since those with greater impact will run longer. Recall that cumulative reinforcement decays over time, so all configurations of the model will ultimately pass below the threshold for sending referrals and come to a stop. In the next section, we describe our test scenarios for the model and present the results.

Results and Analysis

To illustrate the practical application of our model, we designed three separate scenarios that test three hypotheses about the effectiveness of a particular persuasive messaging strategy. The purpose of these scenarios and the results that follow are intended to provide insights into how such a model, once properly calibrated, could assist in the design of persuasive messaging campaigns.

Scenarios and Experimental Design

Our first hypothesis is that the presence of messaging will have no effect on the average behavioral intent to perform a particular behavior within the population. Given the heterogeneous nature of our agents, this result is not trivial. If an agent evaluates the outcome negatively, increased reinforcement from messaging will have a negative impact on $BI(t)$. The heterogeneity of agents and the fact that belief decay occurs in the model also lead to a natural decline in $BI(t)$ in the absence of messaging. As a result, we varied the starting time of the messaging campaign in order to analyze the impact given different initial levels of the average population behavioral intent. We chose four different starting times while holding the interval of messaging constant (see Table 1).

We ran each experiment for 30 replications and collected statistics on the time average behavioral intent for each run. All 30 time series for each of the 4 experiments are shown in Figure 6. In the first experiment, messaging starts at t=0 and ends at t=100 as illustrated by the shaded area. It can be seen from the plot that the initial conditions of the model are impacting the level of $BI(t)$ and making it difficult to determine what impact is due to messaging and what

Table 1. Message starting and stopping times

Experiment	Message start	Message stop
1	0	100
2	100	200
3	200	300
4	300	400

is due simply to initial conditions. There is a secondary peak in $BI(t)$ that interestingly occurs around t=150, which is after messaging has stopped. In an attempt to separate initial conditions from messaging, we steadily move the starting time of the campaign later and later. The result is two-fold. First, we see that the peak for behavioral intent occurs consistently *after* the messaging campaign has stopped. Second, we note that the peak declines the longer we wait to start the campaign. This would imply a certain amount of inertia exists in both the increase and decrease of population-level behavioral intent.

The results of our four experiments are conclusive enough that we reject the null hypothesis that messaging will have no effect on behavioral intent. We can also see that the model approaches a relatively steady-state around t=150 if messaging starts after t=150. Thus, to eliminate the impact of the initial warm-up period, we establish scenario 1 experiment 3 with the message start time as t=600 as our base case for comparison against our next two scenarios.

In scenario 2, we wish to test the impact of an influential figure on behavioral intent. An influential figure is simply an agent in the model that has significantly more social connections than the rest of the population and advocates on behalf of the messaging campaign. However, only a portion of the population is connected to the influential figure and not everyone that is connected to them has a high motivation to comply with the figure. One can think of the influential figure as a celebrity that is highly visible, but not necessarily universally popular. In this case, we perform nine different experiments. In all nine, the influential figure is connected to 90% of the population and it speaks out on behalf of the messaging campaign from t=200 to t=300. For each of the nine experiments, we vary the motivation of each connected agent to comply with the influential figure. We accomplish this in the simulation by randomly assigning the motivation to comply with the influential figure from a constrained interval. Recall that motivation to comply is generally in the interval [-1,1], where 1 is highly motivated to comply and −1 is highly motivated to not comply. As such, we vary the interval for random assignment according to Table 2.

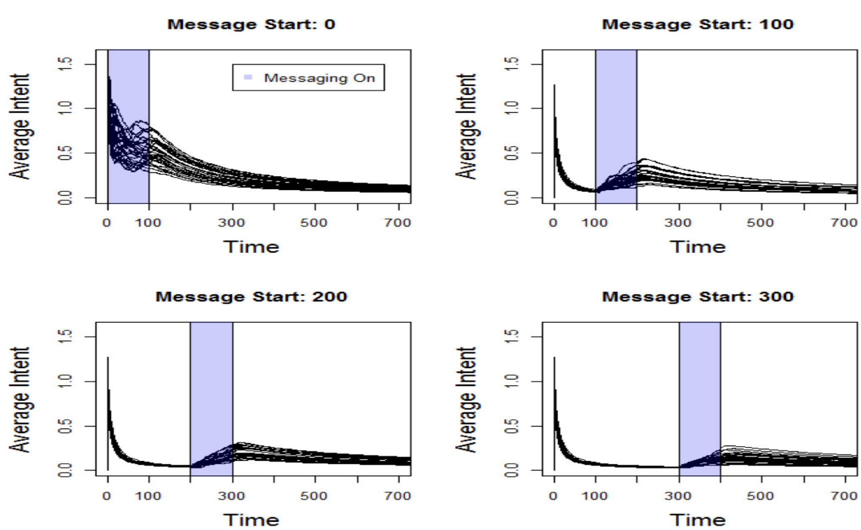


Figure 6. Time average behavioral intent.

Table 2. Motivation to comply selection interval

Experiment	Motivation to comply interval
1	[1.0,1.0]
2	[0.5,1.0]
3	[0.0,1.0]
4	[–0.5,1.0]
5	[–1.0,1.0]
6	[–1.0,0.5]
7	[–1.0,0.0]
8	[–1.0,–0.5]
9	[–1.0,–1.0]

For each of the 9 experiments, we ran 30 replications and compared the results to our base case from scenario 1. The box and whisker plots in Figure 7 show the average behavioral intent, with the boxes colored to indicate if the experiments are significantly different from the base case, where no influential figure is present. The blue plots indicate the experiments that are significantly higher than the base case and the red plots indicate the experiments that are significantly lower than the base case. The white plots indicate that there is no statistical difference between that experiment and the base case. The base case is shown in gray on the far right of the plot.

Our results indicate that an influential figure can increase behavioral intent even if some agents have a negative motivation to comply. Our results also indicate a lopsided relationship between having a positive effect and having a negative effect. That is, only when all agents connected to the influential figure had a large negative motivation to comply did the behavioral intent get worse. Otherwise, it either improved or was too similar to the base case to be conclusive. These results were confirmed by performing a paired t-test for each replication. The results of the tests are shown in Table 3.

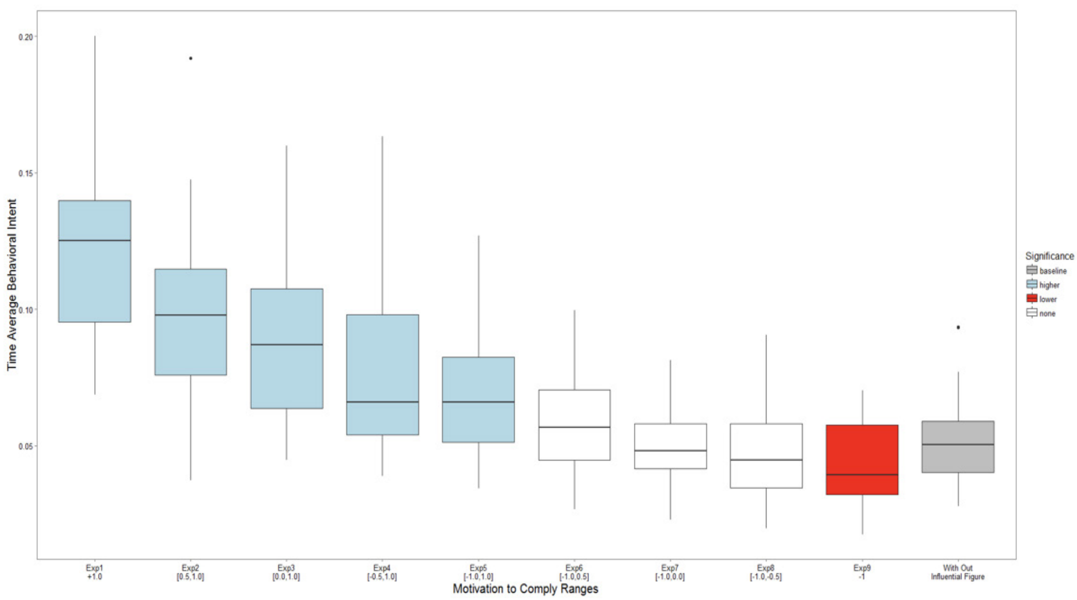


Figure 7. Comparison of with base case.

Table 3. Motivation to comply selection

Experiment	95% CI	*p*-value
1	(0.057,0.0848)	6.36E−13
2	(0.033,0.059)	7.25E−09
3	(0.024,0.049)	3.78E−07
4	(0.014,0.042)	0.0003
5	(0.0062,0.0265)	0.0021
6	(−0.0013,0.0172)	0.092
7	(−0.0098,0.0069)	0.727
8	(−0.014,0.0028)	0.186
9	(−0.0169,−0.0005)	0.038

Table 4. Varying the influential figure start and stop times interval

Experiment	Start of influential figure	Stop of influential figure
1	100	200
2	200	300
3	250	350
4	300	400
5	100	400

Since our focus has been to inform persuasive messaging strategies, our final scenario investigates how the timing of an influential figure speaking out relative to the timing of the messaging campaign impacts behavioral intent. Again, we operate from the base case where messaging starts at t=200 and lasts until t=300. We include the influential figure with 90% of the population connected to them and those connected all have the highest motivation to comply with the influential figure (i.e., +1). However, this time we varied the start and stop time of the influential figure in five different ways, as described in Table 4.

The results of this final scenario are shown in Figure 8. All five experiments showed an increase in the average behavioral intent across the population of agents. However, we also noted significant variation in the increase depending on when the influential figure spoke out. Experiment 5—where the influential figure was active from 100 steps before the messaging campaign started to 100 steps after it completed—showed the highest increase in behavioral intent. The smallest increase occurred in Experiment 1 when the influential figure spoke out before the messaging campaign started and stopped speaking out once the messaging campaign commenced. All of these experiments showed statistically significant increases in behavioral intent relative to the base case where no influential figure was present. These results are important because of the vast sums of money spent on endorsement contracts. The model suggests that timing of an influential figure's endorsement relative to a broader messaging campaign can have different levels of impact on the overall population's behavioral intent.

Discussion

Empirical studies show that static regional- or demographic-based models of human behavior do a poor job of predicting the impact that persuasive messaging campaigns might have on a given population (Gupta & Chintagunta, 1994; Wind, 1978). Social psychology indicates that, even within particular regions or demographics, populations are heterogeneous, and individuals formulate their behavioral intent based on a variety of factors. As

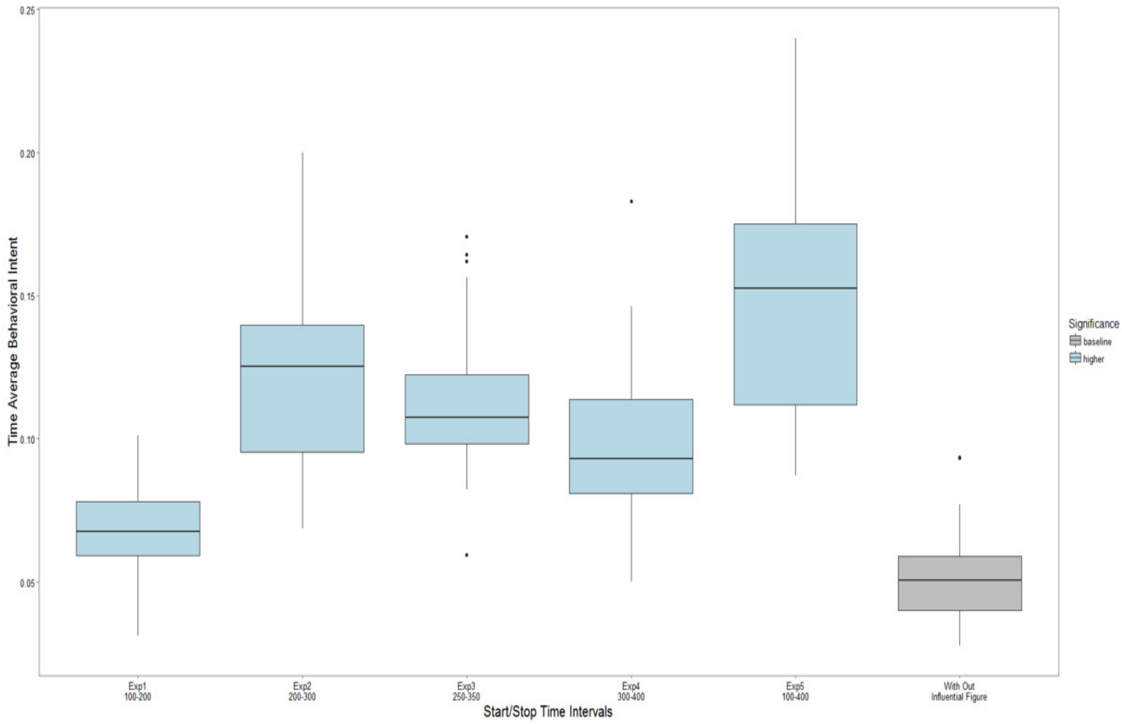


Figure 8. Time average behavioral intent.

a result, messaging intended to increase the evaluation of particular outcomes may, in fact, have the opposite effect or no impact at all. Behavioral theories, such as TPB, offer a common framework for representing the inherent heterogeneity in society, but they are still dependent on the experiences individuals have in a dynamic and often unpredictable environment. Fortunately, agent-based models provide us with a starting point for representing not only the heterogeneous nature of populations, but also the ever-changing environment they occupy. By moving away from static approximations of behavioral action and intent, we can begin to model more complex phenomena such as negative back-lash or "going viral." In this work, we have operationalized one of the most heavily cited theories of human behavior, TPB. Our goal was to illustrate the complexity involved in translating a long-standing theoretical framework into a working agent-based model and to underscore how such a model can be used to test various hypotheses pertaining to both behavior and the underlying theory. It is our contention that scientific advancement in this field will come from first embracing the merits of a particular theory and then rigorously testing any faults or shortcomings it presents. TPB is a theory designed to connect behavioral action to behavioral intent through the individual's current set of beliefs. As a result, it makes no mention of how those underlying beliefs change over time. To operationalize TPB in an agent-based model, the evolution of attitude, SN, and perceived behavioral control over time in a changing environment must be estimated. We chose to build on the existing structures of consumer behavior and the Rescorla–Wagner equations to represent this more dynamic setting for TPB. Our implementation of TPB in an agent-based framework offers a first step in this process. Our experiments qualitatively replicate the idea that increased messaging and the use of influential figures can lead to a net increase in the population's behavioral intent.

A natural next step is to perform more rigorous validation experiments that are coupled with traditional survey techniques that have been modified to incorporate the specific parameters of our model. Another avenue for future work is to test the impact of counter-messaging. Corporations, political candidates, and various organizations often engage in persuasive messaging designed to drive people away from their opponents. In our model, the choice for each agent was either to act or to do nothing. An additional action that is diametrically opposed to the first could easily be added. Proponents of each action could then choose between positive messaging for their product, negative messaging against a competitor's product, or a combination of the two.

One need only turn on the television or log on to social media to confirm that corporations and governments alike believe that persuasive messaging campaigns are an effective way to increase sales and participation. One website estimated that over $187 billion was spent on advertising in the United States in 2015 (Goodman, 2016). And it is clear that celebrities and athletes supplement their income by acting as influential figures on behalf of organizations hoping to increase their sales. Given the enormous sums of money already being spent on persuasive messaging, it follows that at least some measurable benefit comes from the effort. But how many messages are lost among the rest, failing to stand out, and how many have the exact opposite reaction from the original intent? In short, models can provide insights into this process, which in turn could potentially lead to organizations gaining a larger return on their investments in messaging campaigns. However, those types of results will likely only be obtained with continued experimentation and empirical analysis, and synthesizing the results of both can be made easier with agent-based models of human behavior and decision-making.

Acknowledgments

We thank our colleagues at MITRE Jon Cline and Sasha Lubyansky for their invaluable assistance in completing this research. We also thank Gary L. Klein and Chris Glazner for their continued input, peer review, and technical insights.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211.

Axtell, R. (2000). *Why agents?: On the varied motivations for agent computing in the social sciences.* Center on Social and Economic Dynamics. Retrieved from http://www2.econ.iastate.edu/tesfatsi/WhyAgents.RAxtell2000.pdf

Carmack, C. C., & Lewis-Moss, R. K. (2009). Examining the theory of planned behavior applied to condom use: The effect-indicator vs. causal-indicator models. *The Journal of Primary Prevention*, *30*(6), 659–676.

Ceschi, A., Dorofeeva, K., Sartori, R., Dickert, S., & Scalco, A. (2015). A simulation of householders' recycling attitudes based on the theory of planned behavior. In J. Bajo, J.Z. Hernández, P. Mathieu, A. Campbell, A. Fernández-Caballero, M.N. Moreno, V. Julián, A. Alonso-Betanzos, M.D. Jiménez-López, & V. Botti (Eds.). *Trends in practical applications of agents, multi-agent systems and sustainability* (pp. 177–184). Cham, CH: Springer International Publications.

Epstein, J. (2014). A*gent zero: Toward neurocognitive foundations for generative social science*. Princeton, NJ: Princeton University Press.

Epstein, J. P., & Axtell, R. (1996). *Growing artificial societies: Social science from the bottom up*. Washington, DC: Brookings Institution Press.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Boston, MA: Addison-Wesley.

Godin, G., & Kok, G. (1996). The theory of planned behavior: A review of its applications to health-related behaviors. *American Journal of Health Promotion*, *11*(2), 87–98.

Goodman, M. (2016). Advertising forecast: North America 2010–2021. Strategy Analytics. Retrieved from https://www.strategyanalytics.com/access-services/media-and-services/in-the-home/digital-media/digital-media/market-data/report-detail/advertising-forecast-north-america-2010---202

Gupta, S., & Chintagunta, P. K. (1994). On using demographic variables to determine segment membership in logit mixture models. *Journal of Marketing Research*, *31*(1), 128–136.

Hu, X., & Puddy, R. (2011). Cognitive modeling for agent-based simulation of child maltreatment. In *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction* (pp. 138–146). Cham, CH: Springer.

Jager, W., Janssen, M.A., Vlek, C.A.J. (1999). *Consumats in a commons dilemma: Testing the behavioral rules of simulated consumers*. Groningen: RUG-COV.

Kalyanam, K., & Putler, D. S. (1997). Incorporating demographic variables in brand choice models: An indivisible alternatives framework. *Marketing Science*, *16*(2), 166–181.

Kaufmann, P., Stagl, S., & Franks, D. W. (2009). Simulating the diffusion of organic farming practices in two new EU member states. *Ecological Economics*, *68*(10), 2580–2593.

Macy, M. W., & Willer, R. (2002). From factors to actors: Computational sociology and agent-based modeling. *Annual Review of Sociology*, 28, 143–166.

Miller, J. H., & Page, S. E. (2009). *Complex adaptive systems: An introduction to computational models of social life: An introduction to computational models of social life.* Princeton, NJ: Princeton University Press.

Mogles, N., Ramallo-González, A. P., & Gabe-Thomas, E. (2015). Towards a cognitive agent-based model for air conditioners purchasing prediction. *Procedia Computer Science*, 51, 463–472.

Pavlou, P. A., & Chai, L. (2002). What drives electronic commerce across cultures? A cross-cultural empirical investigation of the theory of planned behavior. *Journal of Electronic Consumer Research*, *3*(4), 240–253.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York, NY: Appleton-Century-Crofts.

Richetin, J., Sengupta, A., Perugini, M., Adjali, I., Hurling, R., Greetham, D., & Spence, M. (2010). A micro-level simulation for the prediction of intention and behavior. *Cognitive Systems Research*, *11*(2), 181–193.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology*, *1*(2), 143–186.

Schelling, T. C. (2006). *Micromotives and macrobehavior*. New York, NY: WW Norton & Company.

Schwarz, N., & Ernst, A (2009). Agent-based modeling of the diffusion of environmental innovations: An empirical approach. *Technological Forecasting and Social Change*, *76*(4), 497–511.

Smith, E. R., & Conrey, F. R. (2007). Agent-based modeling: A new approach for theory building in social psychology. *Personality and Social Psychology Review*, *11*(1), 87–104.

Sniehotta, F. F. (2009). Towards a theory of intentional behaviour change: Plans, planning, and self-regulation. *British Journal of Health Psychology*, *14*(2), 261–273.

Sniehotta, F. F., Presseau, J., & Araújo-Soares, V. (2014). Time to retire the theory of planned behaviour. *Health Psychology Review*, *8*(1), 1–7.

Verzi, S. J., Brodsky, N. S., Brown, T. J., Apelberg, B., & Rostron, B. (2012). *An agent-based approach for modeling population behavior and health with application to tobacco use.* Albuquerque, NM: Sandia National Laboratories.

Wind, Y. (1978). Issues and advances in segmentation research. *Journal of Marketing Research*, *15*(3), 317–337.

Zhang, T., & Nuttall, W. J. (2007). *An agent based simulation of smart metering technology adoption*. Cambridge, UK: University of Cambridge Press.

Zhang, T., & Nuttall, W. J. (2011a). Evaluating government's policies on promoting smart metering in retail electricity markets via agent based simulation. *Journal of Product Innovation Management*, 28, 169–186.

Zhang, T., & Nuttall, W. J. (2011b). Evaluating government's policies on promoting smart metering diffusion in retail electricity markets via agent-based simulation. *Journal of Product Innovation Management*, *28*(2), 169–186.

Plato's Cave to Unite Americans and Save Democracy but Not Without Analytics

Maury Seldin[A,1], Liz Johnson[B]

Abstract

Just as chemical combinations have emergent properties, so does the combination of certain concepts provide true education innovation, in order to expand the development of knowledge. With the tools and insights from analytics and complexity science, we can encode the concepts, as though we had a Periodic Table for the social sciences, and innovate in translational science, in order to spread the educational innovations to a broad electorate. Yet there is still a need to call on philosophical wisdom from Plato's Cave to provide foundational arguments for a unique innovative educational enterprise, one that serves as a catalyst for The Third Wave of the Internet, potentially uniting Americans and saving democracy in its wake. The goal of this research is to provide the vision and application of an Academy in the Cloud Enterprise (ACE), via integrated wiki sites, that capitalizes on the fact that learners start from different places. Furthermore, they have different commitments of energy to gain knowledge. Given there is great reliance on conclusions reached by trusted others, this provides an exceptional opportunity to leverage off the social networks of the motivated learners. Concepts can be sequenced to be credible at face value, but verifiable to some degree by further exploration going back to the ACE wiki site and following paths to literature texts, in whatever depth the learner chooses. Knowledge-based education, like in ACE, helps individuals make better policy decisions in their personal and professional lives. Part of that education relates to the functioning of a democratic society based upon better policy from the voters (and the eligible non-voters), as well as targeting elected officials making wisdom-challenged decisions, with excessive reliance on linearity of twentieth century thinking. The intent is to foster the use of the lens of complexity science appropriately for reasoning and research, using nonlinear analytics. ACE is a work-in-progress based off applications from complexity science, directed to the issues involved in the organic transitions occurring in the American political economy and other dimensions of its modernity. (Please note this is speculative work in process, as part of a larger research endeavor that is developing and will be formally defined and modeled later with an agent-based model).

Keywords: complexity, analytics, invisible hand, democratic structures, emergence, Plato's Allegory of the Cave, Third Wave of the Internet, justice, Enlightenment, Spinoza, motivated learners

---

[A] Counselor of Real Estate, Emeritus

[B] Complex Systems Institute, University of North Carolina Charlotte; Corresponding Author

[1] My coauthor selected the counselor status for identification. For present purposes I would add the following: Dr. Maury Seldin, a chaired professor emeritus of The American University School of Business Administration, was the founding president of the Homer Hoyt Institute created to serve as the research arm of two of the schools programs. He is Chairman Emeritus of the Homer Hoyt Institute (HHI), founded in 1967 and Chairman Emeritus of the Maury Seldin Advanced Studies Institute (ASI), founded in 1982 and serving as the corporate home for the Weimer School of Advanced Studies in Real Estate and Land Economics. Dr. Seldin is co-leader of the Seminar on Strategic Decisions at ASPEC) an interest group that he founded in 2002 and that in recent years has focused on applications of complexity science to a variety of strategic decisions ranging from personal health to organizational management and societal issues.

doi: 10.18278/jpcs.3.2.12

摘要

如同化学物质的组合会产生涌现性（emergent properties），不同概念的结合会产生真正的教育创新，进而扩大知识发展。有了分析学和复杂性科学提供的工具和见解，我们可以对不同概念进行编码（就如同社会科学的“元素周期表”一样），同时在转化科学（translational science）中创新，从而将教育创新扩散给更多人。然而，我们依旧需要柏拉图洞穴寓言中的哲学智慧来为独特创新的教育企业提供基本论点。这一教育企业充当了互联网第三次浪潮（The Third Wave ）的催化剂，并能通过潜在的方式团结美国人民，挽救民主。本文目的是通过整合各类维基网站，为“云企业学院”（Academy in the Cloud Enterprise，简称ACE）提供愿景和应用办法。由此学习者能投入精力从中获取知识。考虑到学习者在很大程度上依赖专业人士的结论，ACE能为积极学习者提供利用社交网络的绝佳机会。概念的可信任度可以依据表面价值进行排序，但学习者也可以在ACE维基网站及其给出的文本内容中进行进一步搜索，从而在某种程度上证实这些概念。ACE提供的基于知识的教育（Knowledge-based education），能帮助个人在其生活和工作中做出更好的决策。该教育部分涉及民主社会的运作和当选官员（elected officials）两方面。前者基于选民投票得出的更好政策，后者过度依赖20世纪线性思维，做出了挑战传统观念的决定。本文意图通过非线性分析学，培养正确使用复杂性科学的方法，进行推理和研究。ACE源于复杂性科学的应用，尚处于发展中。它针对的问题涉及美国政治经济和现代性（modernity）的有机过渡（请注意：本文提出的新概念尚处于推测阶段，需要更多的研究来进行发展，并在将来用基于主体的建模进行正式定义）

关键词：复杂性，分析学，无形的手，民主结构，涌现，柏拉图的洞穴寓言，互联网第三次浪潮，正义，启蒙，斯宾诺莎，积极的学习者

Resumen

Así como las combinaciones químicas tienen propiedades emergentes, también la combinación de ciertos conceptos provee una verdadera innovación educativa para poder expandir el desarrollo de la sabiduría. Con las herramientas e información de la analítica y la ciencia de la complejidad, podemos codificar los conceptos como si tuviéramos una tabla periódica para las ciencias sociales, e innovar en la ciencia traslativa, para poder hacer llegar las innovaciones educativas a un electorado amplio. Sin embargo, todavía se necesita hacer un llamado a la sabiduría filosófica de la Cueva de Platón para proveer argumentos fundacionales para una actividad educativa única e innovadora, que sirva como catalizador para la Tercera Ola del Internet, lo que potencialmente unificaría a los estadounidenses y desde su comienzo salvaría la democracia. El objetivo de esta investigación es proveer la visión e implementación de una Academy in the Cloud Enterprise (ACE), a través de sitios wiki integrados, que capitalice el hecho que los estudiantes empiezan desde niveles diferentes. Asimismo, tienen diferentes exigencias de energía para adquirir conocimiento. Dada la confianza que se le da a las conclusiones de terceros, esto otorga una oportunidad excepcional para aprovechar las redes sociales de estudiantes motivados. Los conceptos pueden ser secuenciados para ser creíbles en su valor nominal, pero verificables hasta cierto grado a través de una exploración adicional de volver al sitio wiki de ACE y seguir las vías a los textos literarios, a cualquier profundidad de contenido que desee el estudiante. La educación basada en conocimiento, como en ACE, ayuda a los individuos a tomar mejores decisiones políticas en sus vidas personales y profesionales. Parte de esa educación está relacionada con el funcionamiento de una sociedad democrática basada en una mejor política de parte de los votantes (y de los no votantes elegibles), así como el enfoque en oficiales electos que toman decisiones con retos de sabiduría, con una dependencia excesiva en la linealidad del pensamiento del siglo veinte. La intención es impulsar el uso de la óptica de la ciencia de la complejidad aproximadamente para el razonamiento y la investigación, utilizando analítica no lineal. ACE es un proyecto en curso basado en aplicaciones de la ciencia de la complejidad, dirigido hacia los problemas que surgen de las transiciones orgánicas que ocurren en la economía política estadounidense y otras dimensiones de su modernidad.

Palabras clave: complejidad, análisis, mano invisible, estructuras democráticas, aparición, La Alegoría de Plato de la Cueva, la Tercera, Ola de Internet, la justicia, la Ilustración, Spinoza, estudiantes motivados

Introduction

American Democracy is at the threshold of a new era. A significant marker is the 2016 elections. The electorate rejected the establishment, represented by the two major political parties. Each of the two major political parties is currently in a substantial transition, or will be soon. The nature of the new era is uncertain. However, there is substantial likelihood of a major discontinuity, or two. On the downside, there could be a disruption including one or more of an inflation and stock market bubble bursting along with the reversal of free trade progress and another major recession. On the upside, the shortcomings of modernity may be ameliorated by a transition equivalent to the evolution of tolerance that started in the Age of Enlightenment. In any case, the nation is suffering a great divisiveness from a lack of understanding of what is essential to resuming the two centuries of progress made toward the ideals articulated in the Declaration of Independence.

The potential for Plato's allegory of the Cave uniting Americans and saving democracy is in the knowledge phase of development and dissemination, with the critical aid of analytics. We assert that knowledge development, in the form of education innovation, can enhance policy outcomes. We argue to achieve this, a contextual perspective from philosophy, history, economics, complexity science, and analytics is essential to gain deeper understanding of America's current place in history and establish a foundation that takes into account the creative power emergence. Wisdom-challenged decisions, private as well as public, in recent decades provided the seeds of discontent. The disruption of the political structure can be viewed as a rebellion. Public policy in recent decades simply failed to avoid an avoidable Great Recession (2007–2009). Policy also failed to evolve to a sustainable political economy pursuing the ideals articulated in the Declaration of independence. Furthermore, there appears to be a reversal in the minimal progress that was being made in regulatory reform. We assert that the time has arrived for a *Declaration of Reform*.

A major purpose of this article is to declare that now is the time for a *Declaration of Reform,* and that *Declaration* needs to include an integration of knowledge drawn from nascent disciplines like complexity science and analytics. Policy decisions, private as well as public, have been indrawn from twentieth century paradigms built upon linear models of cause and effect without adequate attention to the nonlinearity of organic systems such as the political economy and individual pursuits of myopic self-interest.

The belief in free markets without providing a suitable structure falls short in producing adequate knowledge for sound public policy. The belief in mandating outcomes for desired values in public policy falls short in producing outcomes because it fails to understand enough about the process by which the system will actually operate. Divisiveness of recent decades has obliterated the traditional compromise and at times causing a freeze in public policy and misunderstanding of how the system operates, as well as some morality issues, brought about the Great Recession in the United States.

Philosophical and Historical Context Necessary for Education Innovation

The Pursuit of Justice

The pursuit of justice is a critical concept in understanding the operation of a democratic political economy. Understanding the nature of justice is essential, but not easy. Plato's (2008) book *The Republic* is critical to start. We will use the case of justice in the era of the Enlightenment as a precursor to what may return American Democracy back on its historic path to liberty and justice for all.

Ancient Views of Justice

Plato's (2008) book *The Republic* is mainly about justice. He viewed justice to be imbedded in a societal structure, led by philosopher kings and in modernity we add, philosopher queens. As for the populace, each was to fulfill his role in a form of meritocracy, somewhat akin to a track system. It took until the Age of Enlightenment for Western Civilization's justice to be spawned from the rights of the individual rather than from hierarchical authority.

That Age of Enlightenment was built upon a tripod of knowledge, justice, and morality. The knowledge component had roots in Plato's (2008) book, *The Republic*, particularly in its *Allegory of the Cave*. The allegory is about people chained in a cave seeing only shadows on the wall and not the reality of what was creating them. Their understanding of the system that existed was based on the shadows, not the underlying reality. As such, their knowledge was seriously deficient. Furthermore, they were not delighted to be enlightened.

The quality of knowledge used in decisions is critical to likely outcomes. Hammurabi's Code, the oldest extant written record of the laws for justice, tested the guilt or innocence of the charged violator by throwing him into the river. If he swam and survived, he was innocent. The population was classified as royalty, freemen, and slaves. Justice differed by class.

In addition, preceding Plato, his philosophical predecessor Socrates and his follower Aristotle, was Judaic law. Although Israelite structures evolved, by the time of the ancient Greek philosophers, justice for the Israelites was rooted in the individual's rights, and the pursuit of equality of treatment was especially noted. The law was divinely based and accepted in the Covenant. Justice was the application of righteousness, or what we may call morality. The community was structured upon law, sourced from the bible, the righteousness, and the justice dealing with the application of righteousness.

During the time of Aristotle, Epicurus had an explanation of the nature of things that had no divine providence. Justice was simply a relational agreement not to do harm between the parties. Again, justice is a component of societal structure. What differs over time and culture is the nature of justice.

In all the previous cases, we discuss a societal structure. Societal structures are frameworks of complex adaptive systems that are organic in nature—meaning they evolve. We look to deal with the evolution of American democracy and using justice and reason, values and knowledge, and capitalize on analytics to influence the ever-evolving structure and the outcomes. For democratic structures, we need a state that has the monopoly on the use of force for a sovereign territory. Additionally, a rule of law is required, and an accountability of the administration to the electorate. The source of authority is the electorate drawn from the eligible populace.

Democratic structures are organic in nature, as are other forms of political order. They evolve from the interactions of agents in the system. Those agents include the people as individuals with varied views of justice. Organizations are also agents in the system. Views of justice evolve both for the organizations as well as for the individuals.

The Enlightenment that spawned the transition to modernity provides a great case study of the organic process in which the nature of justice changed. In order to envision the potential of the next major shift in Western Civilization it is helpful to understand how Western Civilization underwent that change prompted by Radical Enlighteners who used the power of ideas. Research on the learning process of the minds of individuals, and the minds of groups of individuals, will help in understanding what is grasped by the learners. Additional research on the dissemination of the understanding of the concepts will help in the empowerment of the populace to improve the quality of their decisions as responsible participants in the electorate of a democratic society.

As a start, the focus is on concepts of justice, morality, and the *invisible hand*. An understanding of those concepts, as well as a selection of others, based on a realistic conception of reality, rather than being based upon shadows on the wall transmitted in 140 characters, will improve the quality of decisions. Fostering the process through research development and dissemination sets the stage for an evolution to what we propose calling, a New Age of Enlightenment. The start may be viewed as the case of justice for an agent of change that had the then outrageous idea that an individual had the right to think for himself and express those thoughts. That agent of change was Spinoza.

Justice in the Enlightenment Era

The case of justice under discussion took place early in the evolution of the Age of Enlightenment, also known as the Age of Reason. It was the excommunication of Spinoza from the Portuguese-Jewish community in Amsterdam in the middle of the seventeenth century. Other than the state religion, Judaism was the only acceptable religious practice in Amsterdam. The exception was made because of the faith being strictly pursued. Spinoza had the idea that within certain limits one had the right to think for himself and express those thoughts. That was not conforming to the state and religious law as practiced at that time. With due process, Spinoza was excommunicated.

Late in the next century, his ideas on the individual's right to think and speak became the foundation of the First Amendment to the Constitution of the United States of America. That constitution took more than a decade to evolve as the document declaring the previous thirteen colonies having been united as a confederation were now a federation. The Declaration of Independence was a revolt against the injustices of the British monarch. The authority for the revolt was unalienable rights. The transition to the federation was an emergent process.

The American independence was a middle ground in political restructuring in the eighteenth century. Scotland had a modest change in getting some legislative representation under a limited monarchy and added protections of individual rights. The French Revolution was at the other end of the spectrum.

The Power of Ideas

Ideas are powerful! Among the powerful ideas are a few upon which this discussion is focused. One just noted is that the source of authority for American independence was unalienable rights, the pursuit of which was prompted by a common concern of the 13 colonies to overcome injustices imposed upon them by the British monarch.

Another powerful idea is built upon the concept of an *invisible hand*. In both cases, knowledge of the processes by which the structural systems could effectively operate was primitive by today's standards. Furthermore, the prevailing contemporary knowledge, while advanced compared with that of centuries ago, is still so deficient that the discontent in America is at the threshold of another major transition with the path unknown and the approach of a brink not clear as to nature or timing. However, we need to learn from the Enlightenment Era in order to improve outcomes, especially to avoid deleterious discontinuities.

This research is moving towards allowing researchers to focus their attentions on the applications related to the transitions in the American political system, the American economic system, and other contemporary American issues (Seldin, 2016a). We argue it is critical to build on what we learned from the Enlightenment. However, we need twenty-first century science to provide education innovation to go beyond the development of knowledge and to encode the concepts, as though we had a Periodic Table for the social sciences and to innovate in translational science to spread the education to a broad electorate.

Building on Learning from the Enlightenment Era

A Knowledgeable Electorate

Returning to the views of justice in the ancient era, and recalling the idea that democracy requires an accountability of the administration to the electorate (in addition to the monopoly of force by the state and the rule of law); we may ask what Plato's view was. For a quote relating to the knowledgeability and the effectiveness of the electorate and the elected, we reference the ACE site, Plato's *Cave Unites Americans* (Seldin, 2016b).

Plato doubted that the democratic process would work, in part because of the capacity, but in part, it was as is, interpreted by Gardner Coates (2016) and presented on pages 26–27 of her *David's Sling: A History of Democracy in Ten Works of Art*. ACE's wiki page of *Plato's Cave Unites* Americans notes that with the following excerpt:

> "His [Socrates'] followers were less accepting [of Socrates accepting the hemlock]" (Plato, 2008, The Republic VIII-555bIX.580b). Plato, most notably, developed a skepticism of democracy after his teacher's execution. In The Republic, Plato ranks democracy near the bottom of his list of government types. He argues that blind pursuit of freedom can become a kind of slavery when the city is governed by those who know how to win elections, not those with the people's best interests at heart (Seldin, 2016a, pp. 26-27).

Some in the Age of Enlightenment reiterated doubts about electorate's capacity. Additionally, there was obscurantism at the state level and with the professions protecting their territory. At the other end of the spectrum was Van den Eden, a Radical Enlightener dedicated to education that would enable a republic serving the common good and freedom for all. His participation in a plan for an insurrection in Normandy to establish such a republic led him to the gallows.

Doubts by some of the electorate about others in the electorate surfaced in 2016 with the surprising result of the election of Donald Trump to the office of the President of the United States. Education efforts to improve understanding of reality and voter participation in the electoral process pursued by research and development of knowledge and in the sense of *translational science*, commonly referred to in medicine as "from bench to bed," is an innovation not an insurrection.

Societal Structures

Spinoza was part of a coterie of Radical Enlighteners driving the ideas of the individuals as the source of authority. The moderates were still dealing with Providence, but the radicals, including deists, were still looking to the individuals as the source of authority rather than royalty or church. The Scottish Enlightenment took a moderate approach and joined Britain receiving some legislative representation. There was a substantial comprise in Britain with the authority of the royalty.

France went to the extreme, starting with the French Revolution and continuing with more revolutionary discontinuities in political structure. The 13 British colonies had a long list of grievances about the treatment by King George III. The American Revolution was inspired by the lack of justice. The Declaration of Independence listed the specifications in some detail. However, the Declaration went on beyond declaring reasons and presenting facts. It made pledges in pursuit of independence based on what was claimed as unalienable rights.

The ideals were stated in the Declaration of Independence. They were ideals not reality. It took a couple of centuries of pursuing justice to get a lot closer to reality. Getting closer to the reality expressed in the values of the Declaration requires morality as well as knowledge. Indeed, the contribution of Smith with his book on what became known as *The Wealth of Nations* was built on his philosophical roots including morality. That book included his concept of the invisible hand, which he used in his first book, (2011) *The Theory of Moral Sentiments*.

The societal structure of the thirteen colonies after the Declaration of Independence was a confederation. Late in the following decade, the confederation became a federation, a federal government. The structure evolved to that great transition in order to deal with some issues that could not have been dealt with adequately at the time of the Declaration. The extent of understanding and the ability to adapt to the reality makes a significant difference.

Invisible Hands

The misunderstanding of Smith's concept of the *Invisible Hand* has impeded the advancement of America's pursuit of the ideals articulated in the Declaration of Independence. The concept was misunderstood going back to his (1904) An Inquiry Into the Nature and Causes of *The Wealth of Nations.*

Sheehan and Wahrman (2015) discuss a variety of misconceptions of the invisible hand at the time of the Enlightenment is presented in a book titled Invisible Hands: *Self-Organization and the Eighteenth Century.* According to the authors, the primary view of 18th century European thought transitioned from a mechanical and providential worldview to a more modern worldview of economics. Its focus is on what is called self-organization in the context of what is known as *emergence* in the nascent discipline known as *complexity science*.

The variety of misinterpretations included the *invisible hand* simply being a case of Providence. Its operations were in fact based on the idea of free markets. However, that idea was misunderstood then; and today it is still misunderstood by many. The major shortfall in the free market system comes from not understanding that the key to successful free markets is an equitable structure such as a level playing field.

Smith's (2003) *Wealth of Nations* was an interdisciplinary work built from a moral philosophy out of his concern for the detrimental effects of mercantilism. He provided enough in the birth of economics as a discipline to split it from political economy. A major component of his structuring economics was in his use of the phrase the *invisible hand*. Centuries later, the concept became understood as a metaphor for what is called emergence as it is used in *complexity science*.

As economics developed, it did so in the tradition of cause and effect analyses and linear models popularized out of the Scientific Revolution. There was a physics envy popular through most of the twentieth century. Then came the beginning of a trend to drawing from the biological sciences to deal with the nonlinearity of the political economy.

We are back to discussing shadows on wall as a source of understanding that fails to come to grips with reality. However, now, centuries after the idea was introduced, we are finally getting a handle on the reality that based upon our current knowledge, invisible hands is simply a metaphor for an emergent process stemming from transactions in a market producing characterizes at a macro level not present at the micro network. This has introduced us to the next section, Twenty-first Century Science.

Twenty-First Century Science

We argue that twenty-first century science provides a much more realistic understanding of the process that is called *emergence*. The process is present in neurological behavior of the human brain producing the operations of the mind, the brain at work. It is also present in the self-organizational behavior of agents in a societal structure producing the nature of the political economy. A brief overview of society as a complex adaptive system serves as a transition to the discussion of the mind and knowledge that underlies the self-organization of the human agents in the system as individuals and as agents, along with their organizations as agents, in the structure of the societal system.

Complexity Economics

Complexity economics is just one case of the social sciences that started the integration of complexity science into its main discipline. It functions as a branch called *complexity economics*. It takes a view of economics as a dynamic system with nonlinear relationships. It functions based on the interactions of agents (entities that are taking action; also known as nodes in network science). Those agents have biases, make errors, and learn over time; so they evolve as the biological entities that they are. The networks they operate in also evolve over time and the structure changes. That changing structure impacts the future decisions of the agents, so the structure continues to evolve influenced by changes in the decisions by agents, especially as influenced by education.

Political Structure

Further, the economic system functions in a political environment. The structure of that environment could be chaotic as in failed states. However, out of chaos, order may emerge. That is part of what complexity science is about. An ordered political structure, such as a democracy, has the monopoly on force for its territory. Nevertheless, it requires a rule of law and an accountability to the source of authority, the electorate. Given political structure is an organic entity, it is subject to decay. That is what has started late in the twentieth century with American Democracy. Organic systems emerge from interactions of entities, and grow through stages, sometimes to a great flourishment. However, the natural course of events in the process is decay.

Sometimes the system will be rejuvenated, and some systems simply generate offspring. There are various processes operating. In the case at hand, we have just mentioned the political and economic processes as with complexity economics. We also note the sociological processes. That discipline has also evolved.

The Whole of the Democratic System

The whole of the democratic system, economic, political, and social, is rooted in the agents. That gets down to the individuals. Those individuals evolve influenced by their genes and memes; or what is often termed nature and nurture. We can go back to Machiavelli (1469–1527) for the roots of political science and the view of dealing with reality of what is rather than what ought to be. Then we can move to the heart of the Enlightenment era with Kant (1724–1804). It appears in various translations from his Proposition 6 of (1963) *Idea for a General History with a Cosmopolitan Purpose* making the point that we loosely state as whomever is in charge is still operating under animal instincts and will take care of himself at the expense of others.

Yet there is a morality issue in whom to trust. It was easier to resolve in primitive societies where the social capital emerged from the necessity of operating as a team to hunt big game. The leader divides the spoils, but if he is not fair, he will have difficulty in assembling a group for the next hunting party. When the scale gets large, many participants work the system for their individual benefit at the expense of others in the system.

The invisible hand concept is that individuals working for their own benefit, but within the societal rules based on a level playing field, will enhance the wealth of the community. The incentive is important to the productivity of the individual, but the social capital emerging out of the self-organization is critical to the productivity. Working the system erodes the social capital and unfairly distributes the benefits. We argue it is a cancer in the system. In many systems, it is not moral.

The Mind in Human Evolution

In American democracy, we are free to believe whatever we choose to believe. Increasingly there is belief starting with the premise of science, but many people include civil religion. Others select faith based on a deity. However, the behavior makes the difference on how we live in a free society. In addition, that behavior is heavily influenced by the capacity we have in the use of our minds.

The capacity of the mind and willingness to learn in order to make use of it varies widely. As inferred, the success of democracy his heavily dependent on its use. Unfortunately, we are still "...hampered by the Paleolithic curse: genetic adaptations that worked well for millions of years of hunter-gather existence but are increasingly a hindrance in a globally urban and techno-scientific society" [Wilson, 2014, p. 176].

Doing a better job in the functioning of American democracy through the electorate learning is helped by the innovations being provided via the Internet and the research being fostered by the Academy in the Cloud Enterprise (ACE), which is what this research is about. This takes the current discussion to exploring a path beyond Modernity.

Beyond Modernity

Modernity as a State of Mind

Modernity as a state of mind includes perspectives of the humanities as well as the sciences, but human behavior still includes the remnants of the "... Paleolithic curse: genetic adaptations that worked well for millions of years of hunter-gather existence but are increasingly a hindrance in a globally urban and techno-scientific society." [Wilson, 2014, p. 176]. A major discontinuity in perspectives leading to Modernity occurred with the Scientific Revolution and the overlapping Age of Reason and Enlightenment.

The challenge facing American democracy may be viewed, at least partially, as the pursuit of dynamic balance in an evolving societal structure still faced with a present crooked *timber of humanity*, to borrow a phrase. Tyrannical regimes, hierarchal structured, use force to constrain the populace. Democracies have the power to use force, but the self-organization builds and operates the structure with the source of authority from the people. However, democracy is *not a spectator sport*, to borrow another phrase. It depends on the participation of the eligible electorate and the wisdom of their choices. The conditions facilitating the widespread presence of wisdom challenged decisions may be ameliorated by enhancing the education of the electorate.

Indoctrination by a centralized authority is at odds with the premises of the Declaration of Independence and runs contrary to the freedom of thought and expression protected by the First Amendment to the Constitution of the United States. The availability of public education for all is a first step in a free society's approach to an educated electorate. Beyond that first step, there is the idea of an opportunity for all to flourish, alternatively stated as the *right to life, liberty, and the pursuit of happiness*.

A Neurological Perspective of Learning

A neurological perspective of learning goes to the cutting edge of supersizing the mind. It is discussed in the eleventh chapter of a book titled *Supersizing the Mind: Embodiment, Action, and Cognitive Extension* authored by Clark (2008). It deals with the learning process that would enable individuals to grasp the advances in twenty-first century science. From Ridley's (2015) *The Evolution of Everything: How New Ideas Emerge*, Ridley's chapter on the evolution of education presents a viewpoint of how education would likely appear if its natural evolution occurred.

This article is part of an effort to develop and disseminate an innovative approach to education that would be a pioneering effort in what Case (2016) calls a *third wave of the internet age*. His book is The Third Wave: An Entrepreneur's Vision of the Future. One of those sectors discussed in his section on transforming is education.

The innovation being created in ACE stems from combining advancements in science discussed by Clark (2008) in his book *Supersizing the Mind: Embodiment, Action, and Cognitive Extension* with idea of the third wave discussed by Case in his book, The Third Wave: An Entrepreneur's Vision of the Future. The technology we are exploring for research and application, we view as related to what Clark writes in his chapter 3, Material Symbols. He has sections on Language as Scaffolding, Augmenting Reality, and Sculpturing Attention. This is further explicated in an internal item of ACE titled *Understanding the Analytical Tools for Complex Adaptive Systems*.

Providing the Mind with a Tool

We aim to provide the learner's mind with a tool that can serve as an extension of the mind. The mind is going to select information, process it in a paradigm, utilize a value system, and operate to defend or advance interests. Clark (2008), in his book S*upersizing the Mind: Embodiment, Action, and Cognitive Extension*, in his chapter 3, Material Symbols, have sections on Language as Scaffolding, Augmenting Reality, and Sculpturing Attention. In that third section, the discussion may help us understand how to empower the learner.[2]

To empower the learner, we as innovators, start with a foundation for the analytical tools. That foundation, assuming that we build it on the counterpart of the periodic table, specifying rules and tools for interaction of agents that foster favorable emergent behavior in developing and sharing knowledge, leads to choices that support emergent patterns with greater likelihood for favorable outcomes that other options. It is like choices in checkers or chess or in neural networks as discussed by Holland (1998) in his book Emergence: From Chaos to Order.

In short, as stated in the opening of the home page of ACE, "The mission of this enterprise is to improve policy at levels ranging from the individual level to the societal level through educational innovation dealing with the nature of complex adaptive systems utilizing the thrust of advancing research so as to better understand the emergence of patterns of interactions." It would be helpful to have a social science counter-

---

[2] Empowerment is what education is all about; in part for one's own life and in part for the communities within which one functions. A sense of place is one way to approach the empowerment, especially because it involves networks and emergence that impacts structures.
Sense of Place and Empowerment
This essay, Sense of Place and Empowerment, is the first of a two essay series. The second essay, Empowerment of a Society, carries the idea of empowerment from that of the individual in small-scale relationships to generating an impact on the structure of the environment.
Sense of place contributes to both self-empowerment and the empowerment of an individual by others. The intent of the analysis is to explore an innovative opportunity to get closer to an equitable representation of the American public in the development of the societal structure. That exploration starts with gaining a better understanding of how an individual's sense of place can impact the generation of power.
Empowerment of a Society
Individuals can gain empowerment by networking with others who have a sufficient commonalty of interests. It takes brains and commitment for an individual to progress through the structure that imposes constraints. There is a co-evolution of environment and mental content. The process generates empowerment. Empowerment as the exercise of power is a major route to changing the direction of the evolution of American Democracy. This essay is a segue to the six essay treatise that is intended to foster the process by fostering better analytics and an improved level of morality that contributes to building social capital and enhancing the presence of humanity in American Democracy.

part to the Periodic Table of the physical sciences. However, progress can be made by research revealing patterns of human behavior in the learning process. Complex event processing (CEP) has a role to play in the evolution of education in a democratic society.

The CEP is of great interest because it uses data from multiple sources to synthesize inferences about events and patterns that come from complex problems. We translate that to the jargon of complexity science to mean that *complex event processing* (CEP) is a computational model that processes interactions of agents to produce emergent properties in the form of potential patterns of outcomes. It is the patterns of outcomes and the patterns of activity from interactions within networks that are of particular interest. This is within the context of complexity science where altering structures has great impact on outcomes, but it is the processes that we need to understand better in order to get better outcomes.

Further to Holland's (1998) *Emergence: From Chaos to Order*, he uses neuroscience and checkers to explain patterns in emergence. We need research to identify the patterns in learning and the content of that which is learned to carry the innovation forward in the form of a pilot project. The entity of ACE can go beyond simply facilitating the self-motivated learners to pursue their educations. The going beyond is to use computer technology to foster the process. It is in the patterns that form the emergent process of education that is what the ontology is about in the applications at hand. The strategy for education innovation used in this enterprise is based on individuals developing their sense of place in community realizing the importance of balance in individual and community interests.

A departure point is a series of fine projects funded by the Bill and Melinda Gates Foundation in what Case would classify as second wave advancements in utilization of the web. Projects were launched with education apps on Facebook. The idea is building pathways to knowledge, but the focus was on facilitating the development of peer groups among students to allow them to get into college and to allow them to afford college (Case, 2016).

An extension in the second wave is to facilitate the interactions among the peer groups in which they educate each other. There is some research on that at the primary and/or secondary education levels. Our focus is at the post-secondary levels, but supplementing the hierarchical structure of curricula with self-selected paths of study, possibly supplemented with coaches or mentors.

The third wave calls for the integration of the web interaction as an extension of the mind leveraging off each other's knowledge. In order to facilitate that it would need to utilize some computerized shortcuts to patterns, possibly built on information science ontology. Progress is being made in the startup stage for the Academy in the Cloud with Motivated Learners via wiki. The primary concept is the evolution of the way the mind becomes an extension of the body especially in the use of language to extend our thought. Clark (2008), Supersizing the Mind, discusses that in the book: *Embodiment, Action, and Cognitive Extension*. The concept is built upon developing the ACE Motivated Learner Strategy.

In order to build on that concept, the enterprise utilizes the treatise *American Democracy, The Declaration, Pursuit, and Endangerment*, from the ACE wiki site. The site provides an introductory education to some nascent disciplines in order to aid learners to understand the system using the lens of complexity science.

That proprietary work, until now labeled with the phrase "All Rights Reserved," is placed in the public domain encouraging motivated learners to use it, modify it and share it with their contemporaries so that each of them may improve her or his quality of life, empowerment, and the functioning of a democratic society. It is provided using software that facilitates participation in chat rooms for the ACE site and hyperlinks within the ACE site that can retrace steps as well as moving forward on paths selected by the users. That same software should be readily available to motivated learners so that they can create or adapt their own website to build a network through which they can share their thinking based upon what they are learning. It makes it easier if the same server provider is readily available to the motivated learner.

Continuing with the primary concept that the mind becomes an extension of the body, we have from Case (2016), *The Third Wave: An Entrepreneur's Vision of the Future.* The author stresses content, context, and community as critical elements in the "reinvention of education" [pp. 51–52]. The strategy for education innovation used in this enterprise is based on individuals developing their sense of place in community realizing the importance of balance in individual and community interests. This is further explicated in *The Hybrid Project: Innovative Use of Software for Learning Through the Lens of Complexity Science*, which is also linked on the ACE site. That site mentions Ridley's (2015) *The Evolution of Everything: How New Ideas Emerge.* Ridley's chapter on the evolution of education closes his opening paragraph with "What would education look like if allowed to evolve." Allowing education to evolve is another of the concepts being blended in the development of the strategy. It is a shift from mainline structure of the hierarchical approach of selecting bodies of knowledge to be transmitted and using the classroom structure for transmission to a new mainline structure of taking post-secondary learners on a counterpart to a tour of a museum to a docent led tour of literature as a follow-on to discussion of some societal concerns starting with the individual's sense of place. The learner is free to select various paths beyond the traditional college structure of a major and a variety of electives with the bodies of knowledge structured as silos of disciplines, a categorization of disciplines similar to a categorization of denominations as in religions.

This new mainline is highly interdisciplinary. It integrates disciplines that have used linear models for analytics to understand likely outcomes, but they are only building blocks for patterns of choices by agents (individuals and organizations) in complex adaptive systems. This is not intended to supplant formal higher education; it is an innovative approach to what in higher education has been referred to as a liberal education and/or teaching people to think. The difference is that in the old mainline models curriculum is not prescribed in a hierarchical fashion, but rather emerges by the individual selecting in pursuit of knowledge.

Conclusion

The Academy in the Cloud Enterprise (ACE) is predicated on the participation of motivated learners, as a strategy to disseminate knowledge and thereby enlightening those trapped in the ignorance of Plato's Cave. It is through motivated people and people that care about strengthening democracy that innovations in education can be developed and disseminated. The innovation, currently under development, is an integration of multiple concepts, disciplines, historical contexts, perspectives, and technologies.

Another critical concept is that the emergence of the education of the community is a result of the interactions of individuals, each with his or her own talents with a sharing of knowledge. The results may be categorized using traditional disciplines, but blended as interdisciplinary when applied to complex adaptive systems.

According to Seldin (2016a), the mission of the ACE should be to improve policy decisions at every level of society by facilitating greater innovation in education (Seldin, 2016a). It would be beneficial to have a social science counterpart to the Periodic Table of the physical sciences, which this research hopes to achieve. We argue progress can be made by research revealing patterns of human behavior in the learning process.

Another step in that direction is the use of analytics as an exploration of the paths motivated learners choose in getting their education given ample links to expand on points made in a line of reasoning in dealing with a complex adaptive system such as American Democracy. It makes a difference in the context of differentiation among items of information, knowledge, and intellect. Consider information as data, bits that have meaning when combined with other bits of similar or diverse character. When information is mixed with other information in a fashion that reveals relationships, then there is knowledge. The quality of knowledge may vary widely, but it is an input to the process by which the mind forms some opinion of outcomes given whatever paradigm is used. Currently, we are pilot testing these constructs in a college level business ethics course taught in the context that business ethics contributes too and strengthens American Democracy.

Intellect is at a higher level than the assembly of knowledge is. It comes to bear when one is able to go beyond the pattern of the discipline and integrate widely diverse knowledge to grasp a holistic view of a system beyond that of available in a single discipline. It requires a greater cerebral capacity. Increasing cerebral capacity is akin to increasing height; it happens in a normal growth process. However, individuals naturally vary in height and cerebral capacity, and taller individuals commonly reach for items out of the unassisted grasp of shorter individuals. They also organize as in basketball teams to utilize different talents, and playmakers are important, as are rebounders. Lewis (2004) beautifully explains the differences in baseball in his book, *Moneyball*. The baseball manager could see how the pieces interacted rather than relying on the scouts that saw shadows on the wall.

That context was that the illusions of some provided an opportunity for others with a better understanding of reality to build a better baseball team for less money. The context for our dealing with the same fundamentals of behavioral science is to build a society closer to the ideals espoused in the Declaration of Independence. In order to do that, we turn to discussing a few aspects of injustice in the American political economy preceded by some comments on premises and perspectives.

The strategy is to leverage off those with the greater intellect, a higher cerebral capacity, higher interest, and quest for knowledge sufficient to get a better understanding of the system. The innovation is to provide an opportunity to all who care to explore the ACE site. Those who self-select as motivated learners to pursue knowledge to greater depths, revealed by the paths they choose to follow with hyperlinks may be offered mentors or coaches to assist them in attaining levels of understanding most people may not have the interest in pursuing or the capacity to absorb.

Those so selected may be encouraged to establish copycat sites that evolve to their interests and that of those in their social networks. Those in their networks that want to pursue similar depths are free to do so. Nevertheless, many are satisfied with tweets. Twitter is used to share thoughts, known as tweets, in 140 characters or less. The quality of knowledge gained by a tweet can be believed to be high based on trust. Alternatively, it may be a bit of information that tells more about the sender than about the purported knowledge conveyed. In any case, the receiver of information from social networks goes to some stage of reasoning and trust in selecting what to think.

The point is that there may be many wisdom challenged decisions being made based upon poor information and questionable knowledge. A better education of the populace is essential to a better functioning of a democratic society. It also holds that a better education of the leaders is essential. Furthermore, if they are not quite educable, there are likely to be more knowledgeable ones available.

That activity of learning and choosing is a process centered in the mind. The electorate in a democracy needs to use more of their capacity to enable them to produce better reasoning performance. Using more of that capacity is enabled through education in one form or another.

Now is the time for an innovation in education that goes well beyond the traditional hierarchical structures and moves into the *third wave of the internet* and capitalizes on Plato's wisdom, "No knowledge considers or prescribes for the advantage of the stronger, but for that of the weaker, which it rules" (Plato, 2008, p. 17).

Acknowledgements

This article is part of a new wiki site, Academy in the Cloud Enterprise, which features the treatise titled American Democracy: The Declaration, Pursuit, and Endangerment. That treatise calls for a Declaration of Reform. Research on the subprime crisis and the capital market freeze that led to the Great Recession led to a conclusion that the debacle was not just a real estate finance issue, it was a societal issue best understood through the lens of complexity science, especially complexity economics. See a newsletter Supplement– "Homer Hoyt Institute Research Initiative."

References

Case, S. (2016). *The third wave: An entrepreneur's vision of the future.* New York, NY: Simon & Schuster.

Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension.* New York, NY: Oxford University Press.

Gardner Coates, V. C. (2016). *David's sling: A history of democracy in ten works of art.* New York, NY: Encounter Books.

Holland, J. H. (1998). *Emergence from chaos to order*. Cambridge, MA: MIT Press.

Kant, I. (1963). *Idea for a general history with a cosmopolitan purpose in On History.* Indianapolis, IN: The Bobbs-Merrill Company.

Lewis, M. (2004). *Moneyball: The art of winning an unfair game.* New York, NY: W. W. Norton & Company.

Plato (2008). *The republic.* Overland Park, KS: Digireads.com Publishing.

Ridley, M. (2015). *The evolution of everything: How new ideas emerge.* New York, NY: HarperCollins Publishing.

Seldin, M. (2016a). *Academy in the cloud*. Retrieved from www.academyinthecloud.org

Seldin, M. (2016b). *Plato's Cave unites Americans.* Retrieved from www.academyinthecloud.org

Sheehan, J., & Wahrman, D. (2015). *Invisible Hands: Self-organization and the eighteenth century*. Chicago, IL: University of Chicago Press.

Smith, A. (1904). *An inquiry into the nature and causes of the wealth of nations* (5th ed). E. Cannon (Ed.). Library of Economics and Liberty. Retrieved from http://www.econlib.org/library/Smith/smWN.html

Smith, A. (2003). *Wealth of nations*. New York, NY: Random House.

Smith, A. (2011). *The theory of moral sentiments*. New York, NY: Gutenberg Publishers.

Wilson, E. O. (2014). *The meaning of human existence.* New York, NY: W. W. Norton & Company.