

Agent Modeling of Hispanic Population Acculturation and Behavior

Lyle Wallis
Dr. Mark Paich
Decisio Consulting Inc.
201 Linden St. Ste 202
Fort Collins CO 80524
(970) 416-1080
Lyle.Wallis@Decisio.net

Dr. Andrei Borshchev
General Director & CEO
XJ Technologies Company Ltd.
www.xjtek.com
andrei@xjtek.com
+7 812 2471674

Abstract

In recent US Census data widely reported in the press “Hispanics” have become the largest minority group in the US. Using simulation modeling technology we look at some of the structural forces that shape the characteristics of the Hispanic population. The model creates a simulated Hispanic population whose level of acculturation to the broader population of which it is a part dynamically varies according to individual choice. The modeling technique used draws on both System Dynamic and Agent based paradigms both supported by innovative AnyLogic software. The representative Hispanic population is disaggregated down to the individual level as individual agents. Each agent makes choices stochastically as modulated by its current state and the outside environment that it is in. The essence of being “Hispanic” is rooted in cultural attributes that are transmitted through the population by agent mobility and other mechanisms. While some aspects of an agent’s state are represented discretely we also draw on the well developed System Dynamics concepts of modeling “soft” variables to represent the accumulation and decay of cultural attributes within an agent. Also in the System Dynamics tradition we identify global level feedback structures that shape agent level behavior. Finally, we employ innovative data visualization to expose the system’s dynamic behavior to the audience in a compelling fashion. We find that dynamically complex behavior endogenously emerges in the population with temporally stable population sub-segments developing. The underlying dynamics that create these segments and the segments themselves are of interest to those who want a deeper understanding of the Hispanic population. In addition, most of the structures captured in this model are broadly applicable to studies of population and cultural dynamics and are not limited to the Hispanic population.

Hispanic Population Dynamics

There is no dispute that the Hispanic population is large and growing rapidly. There are 39 million Hispanics in the US and their population is growing at 3% per year vs. .8% for all other groups. Two generations from now, Hispanics will constitute 25% of the labor force as opposed to 12% today.

There is significant controversy over the likely social integration of the Hispanic population. A recent cover story in BusinessWeek titled “Hispanic Nation” highlighted two divergent viewpoints. Traditionally, large immigrant populations have gradually surrendered their native languages and blended into American culture. Although the assimilation process took several generations, these groups have not created separate subcultures divided by language or different values. Many believe that American culture will succeed in absorbing Hispanics just as it has for other immigrant groups.

A second viewpoint holds that the Hispanics will remain a separate and distinct subculture with their own language and institutions. In this view, much of Hispanic daily life will be conducted in Spanish. The availability of Spanish speaking businesses, schools and government offices will retard the adoption of English as the pri-

mary language. Slow adoption of English guarantees that the assimilation process will be slow or non-existent. At the extreme, some predict the growth of Quebec-like separatist movements in states such as Texas and California.

Individual-Level Model

To address the issue of Hispanic population dynamics, we integrated modeling techniques from micro-simulation with the feedback approach from system dynamics. Dynamic micro-simulation models include detailed representations of individuals or households as they move through the phases of life. These models include structure for aging, income, education decisions, and the choice of whether and when to have children. Dynamic micro-simulation models are calibrated with a variety of data sources including detailed population statistics from the Census, longitudinal data dealing with income and employment decisions, and behavioral studies on childbearing and educational choices. Recent examples of detailed micro-simulation models include Life-Paths and FamSim.

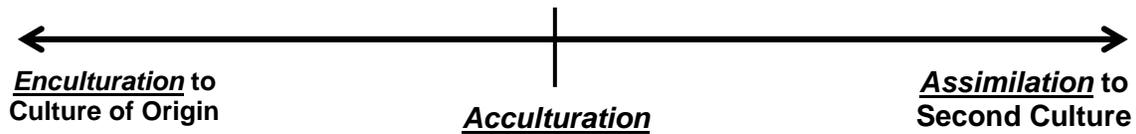
Our modeling approach integrates feedback processes at both the individual and the macro level. Model behavior depends on the decisions of individual but individual decisions depend significantly on the aggregate of choices made by others. At the individual level, with delays, decisions affect variables such as education and income, which, in turn, feed back to affect future decisions. For example, the decision to postpone childbearing increases the probability of staying in school, which, after a delay, increases income. Higher income further reduces the probability of having children. Our model includes many of the feedback relationships that are thought to connect education, income, fertility and aging.

At the macro level, aggregate individual choices feed back to affect the individual behavior that forms the aggregate. For example, individuals choose whether to live in a primarily Hispanic community. Individual location decisions determine the aggregate “Hispanic-ness” of the community that, in turn, affects individual decisions. A large Hispanic population makes it economical to provide Spanish speaking businesses and media driving greater Hispanic in migration. The impact of the linkage from aggregate to individual behavior is a tendency to concentrate similar individuals in certain areas. However, the tendency to concentrate isn’t necessarily the dominant factor in the system. Individual decisions to learn English and pursue distant educational opportunities can blunt the tendency toward population concentration.

Acculturation

Because “being Hispanic” is essentially cultural rather than racial we need a measure of culture. Market researchers use the concept of “acculturation” to describe the mix of cultural attributes that an individual carries that indicates his or her cultural affinities. Acculturation describes a dynamic balance between two cultures that implies a selected mix of first culture attributes and second culture attributes. An individual is said to be “assimilated” when a second culture is substituted for the first. It is important to realize that acculturation implies an aggregation of cultural attributes from both first and second cultures; an individual does not have to give up the first culture attribute in order to add the second culture attribute. A simple example is language usage. If Spanish is the first culture language, it is certainly not necessary to forget Spanish to learn English. The most direct indicators of an individual’s level of acculturation are his or hers language usage and residency time in the culture. Other indicators include:

- individuals media behavior
- ties to people in their country of origin
- interpersonal network composition
- and value expressions.



- Enculturation is the process of learning one's first and original culture
- Acculturation is the process of learning a second culture
- Assimilation is the substitution of the second culture for the first

Figure 1: "Acculturation" is a dynamic balance between two cultures that implies a selected mix of first and second culture attributes

Systems Hypothesis

Our hypothesis of the structure that drives the acculturation process is shown as a causal loop diagram in Figure 2.. In our model the rate of acculturation is driven by the level of exposure to first and second cultures. We believe that the level of exposure is significantly affected by the existence of "clusters" of like cultured people that live near each other. Referring to the causal loop diagram, if first culture clusters are available, and individuals choose to live in them, then the need to adopt second culture attributes is lessened and the rate of acculturation slows. This is represented by the arrow from "Availability of First Culture "Clusters"" to "Acculturation Level" in the diagram. The "O" indicates "Opposite" meaning that the causal effect moves in the opposite direction from the initiating cause. Further, as the level of acculturation increases, the desire and need for people to live in first culture clusters decreases, reducing their availability. As the diagram shows, this forms a "reinforcing loop." This reinforcing pattern is balanced by the selected choices of individuals to expose themselves to the second culture in order to gain advantages. Finally, while it is not dealt with in this study, the in-migration rate in the system is also affected by the availability of first culture clusters.

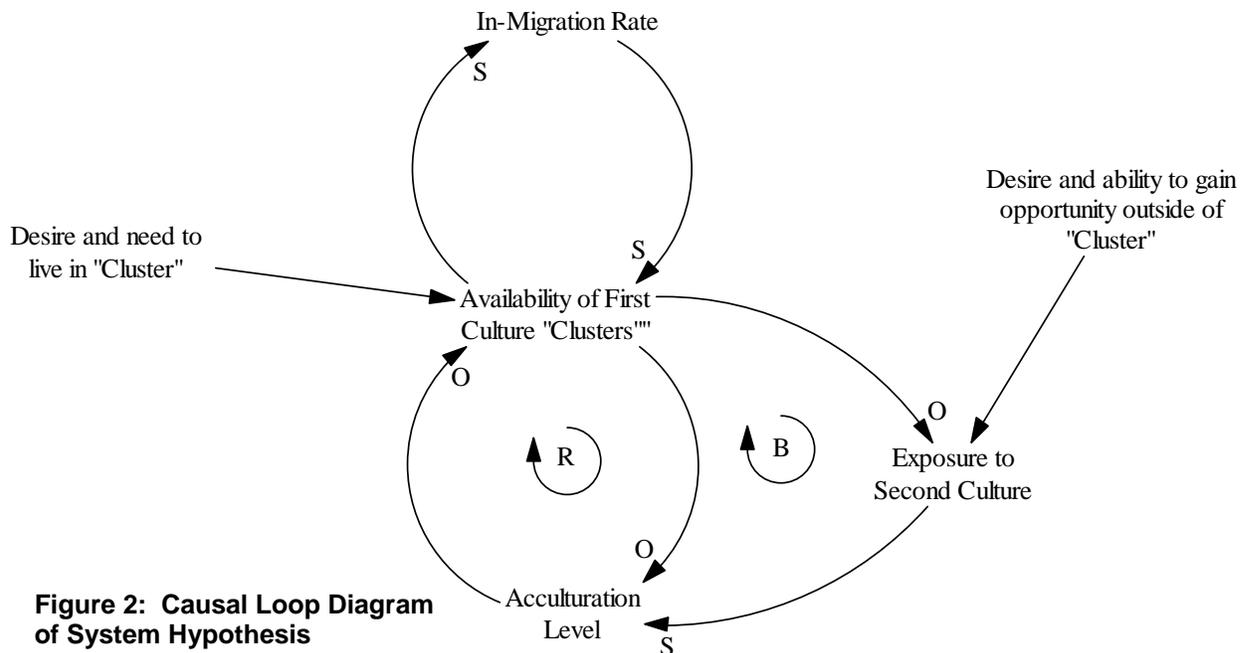


Figure 2: Causal Loop Diagram of System Hypothesis

Modeling Approach

We chose to combine this system perspective with an agent oriented modeling approach. One driver of this choice was that a desired outcome of the study is an understanding of how segments *emerge* in the system. Aggregated models, by definition, presume much of the segmentation which is not conducive to our objectives.

A second, powerful, reason is the potential to directly tie system structure and behavior hypothesis to market research activities. Another supporting argument for the agent approach is that the system hypothesis implies that individual mobility and locality need to be represented comprehensively. This is difficult to accomplish in aggregated models. Finally, modeling technology, in the form of the Anylogic modeling system produced by XJTechnologies, has emerged that makes it practical to implement such a model in a realistic time frame.

Describing Individual Behavior

The model structure is divided into the *agent* and the *environment* that agents live and interact in. The structure diagram for an agent identifies the important aspects of the individual that are included in the model. Characteristics such as the birthdate, income, and number of children are captured as variables, represented by the “blue dot” symbol. Simple dynamic behaviors that occur from time to time, such as household moves, are implemented as “timers” and are represented by red and white “clock” symbols.

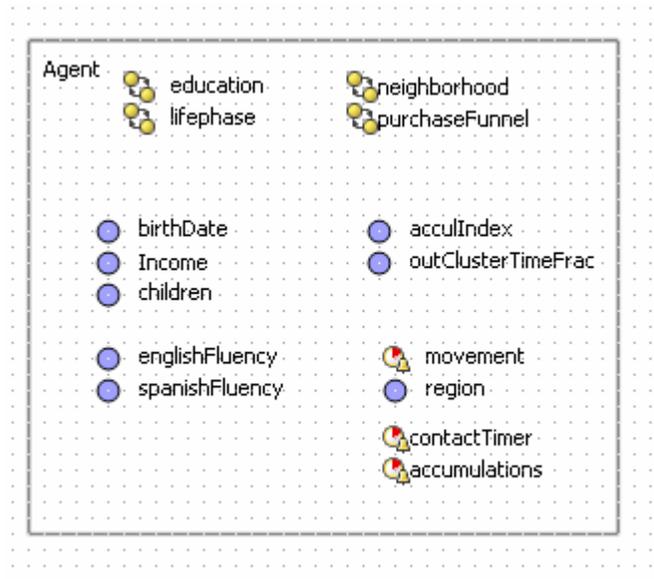


Figure 3: Structure Diagram of an individual agent

More complex dynamic behaviors, such as childbearing, educational choice, and so forth are represented by a powerful capability of Anylogic called *statecharts*. The statechart for “lifephase” shows how an individual’s life phase can be decomposed hierarchically and also represent the dynamics of change. Looking at the lifephase statechart it is easy to see at the topmost level of description the agent can be either alive, or dead. If the agent is alive it can be in one of four different life phases: child, adult, middle aged adult, or senior.

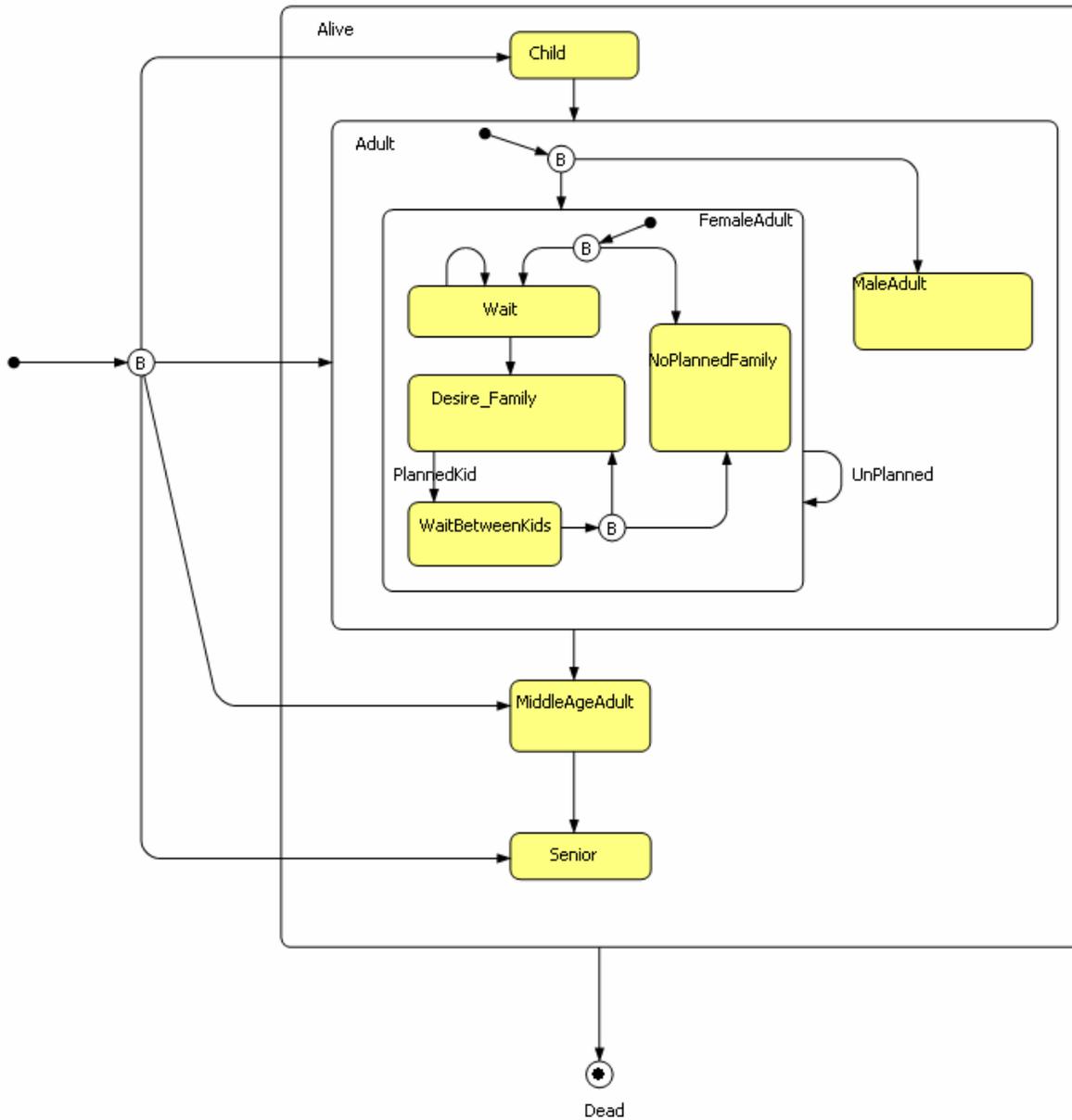


Figure 4: Statechart describing the lifephases of an individual agent

A stochastic function of age describes each state transition. For example, the transitions from “child” to “adult”, from “adult” to “MiddleAgeAdult” and so forth are all described by this simple “java” function: Calling this method in the “trigger” field of a transition causes Anylogic to schedule the state transition for the correct time in the future.

```
/// The method calculates the random time when the agent
/// should switch from one life stage to another
/// and returns the timeout left before the switch. "meanAge" is
/// the average age at which the transition should fire.
/// age() returns the current age of the agent.
/// If the the switch should already occur in the past
/// the method returns 0.
double getLifephaseTimeout( double meanAge ) {
    return Math.max( 0, DistrNormal.sample( .1*meanAge, meanAge ) - age() );
}
```

Life phases are further decomposed to describe childbearing behavior. Adults are either male or female. If female, the agent will stochastically “choose” to bear children or not. If the woman chooses childbearing then she will delay childbearing based on her level of education and cultural norms. Education levels are generated by the “education” statechart (not shown). Once again, transitions are described as stochastic functions of time, age, and other model state variables. Once the woman decides to have children then she advances through intermediate states to the “act” of childbearing which causes the model to generate a new agent that inherits characteristics of the mother.

Once a new child is added to the family then the woman may decide to try to have an additional child or not based upon stochastic properties of age, education, size of family, etc. Whenever the agent is in the female adult state there is a finite probability of the “unplanned” transition being taken. This represents the possibility of an unplanned child and generates a new agent and addition to the mother’s family.

Multiple statecharts can interact through the state variables of the agent creating complex behaviors. In this example, the probability of higher educational achievement is conditioned on lifephase state while the probability of childbearing choices is conditioned by educational achievement. This creates closed feedbacks and powerfully controls resulting behavior.

It is very important to note that the behavior of each state transition can be calibrated to studies of that specific behavior. For example, the probability of a woman desiring another child, given her age, family size, and education levels has been researched extensively. The results of these studies can be used to specify the behavior of the state transitions.

Dynamic System Behavior

The Hispanic Model employs animations to facilitate insights into how the system evolves over time and produces emergent behaviors. Additionally, several controls are available in the animation to allow the generation of alternative scenarios “on the fly.”

This version of the model is focused on the state of California and so the animations and datasets are focused on this geography. Several animations are produced on the animation results page. The first maps Hispanic population density by county over the state of California.

In the Acculturation index vs. Age plot each individual is represented by a color coded dot, with the color representing their generation. Acculturation is represented on a -2 to +2 scale. A -2 indicates that the individual has pure first culture attributes. A +2 indicates that the individual is entirely assimilated into the second culture. A 0 indicates a balanced acculturation state. Over time different patterns of acculturation develop in this representation.

Another view is the Acculturation Index vs Generation, color coded in the same way as the index vs age plots.

In Figures 5-7 snapshots of the animation pages of the baseline run are captured. In Figure 5 is the animation page at the beginning of the animation sequence. Note that the year reads 1950. By 1990, as shown in Figure 6, distinct patterns are emerging based on the inflow of new Hispanics, generational change, education, and so forth. A projection of 2013 in Figure 7 shows well established patterns.

Different sets of assumptions can produce radically different results. For example, by assuming that the tendency of people to “cluster” is low the snapshot in Figure 8 is produced.

One interesting model output is clusters of similar individuals. As discussed above, the model does not impose any population segments or clusters a priori. Rather, demographic and socioeconomic variables are tracked at the individual level and individuals can be assigned to clusters or segments, after the fact, as a technique for understanding the model results.

In most model runs, the adult Hispanic population bifurcates into two fundamental groups. The first group consists of first generation Hispanics who have a relatively low degree of acculturation. In 2002, these individuals spend most of their lives in primarily Hispanic communities, speak fluent Spanish, and have limited English skills. The average age of this group is 35-40 and income is relatively low.

The low acculturation segment persists in 2020. Average English fluency actually declines slightly due to the growth of concentrated Hispanic clusters that reduce the need to speak English. Average income remains low because many Hispanics remain in their community where opportunities are limited. Poor English skills are both a cause and a consequence of living mostly within primarily Hispanic communities. Education levels remain very low.

The second group consists of later generation Hispanics who have a much greater degree of acculturation. In 2002, the average age of this group is 28-32 and the average income is approximately twice as high as the low acculturation segment. English skills are much greater and individuals have lived outside Hispanic communities for 50-60% of their lives. This group is reasonably fluent in Spanish and Hispanic culture remains accessible.

The high acculturation segment also persists into 2020. English fluency increases by about 10% from an already high level and Spanish fluency remains the same. The average age increases to about 35, average education increases by almost 4 years and income grows at an average rate of about 3%. Significant income growth approximately doubles the income gap between the high and low acculturation segments. However, these high-income individuals have not abandoned Hispanic culture. They remain competent in Spanish and still live in Hispanic communities 40% of the time.

The bifurcation of the Hispanic community has been robust to all of the changes in external variables that we have tried. For example, in one test, we reduced the rate of in migration into California to pre-1970 levels. Obviously, the test reduces the size of the Hispanic population and the number of first generation Hispanics. The average acculturation level of the low acculturation segment is slightly higher than in the base case as is the fraction of time living outside Hispanic communities. However, the basic nature of the two segments remains the same. Income, education, and English fluency are dramatically higher in the high acculturation segment.

Eventually, the reduction of in migration would eliminate the low acculturation segment, but the delay is far outside the planning horizon of policy makers. Migration controls will have little short to medium-term impact on the acculturation issue. The demographic destiny has already been set by the several decades of large scale in migration. Rather, policies must be directed to education and measures to increase English literacy of the existing population.

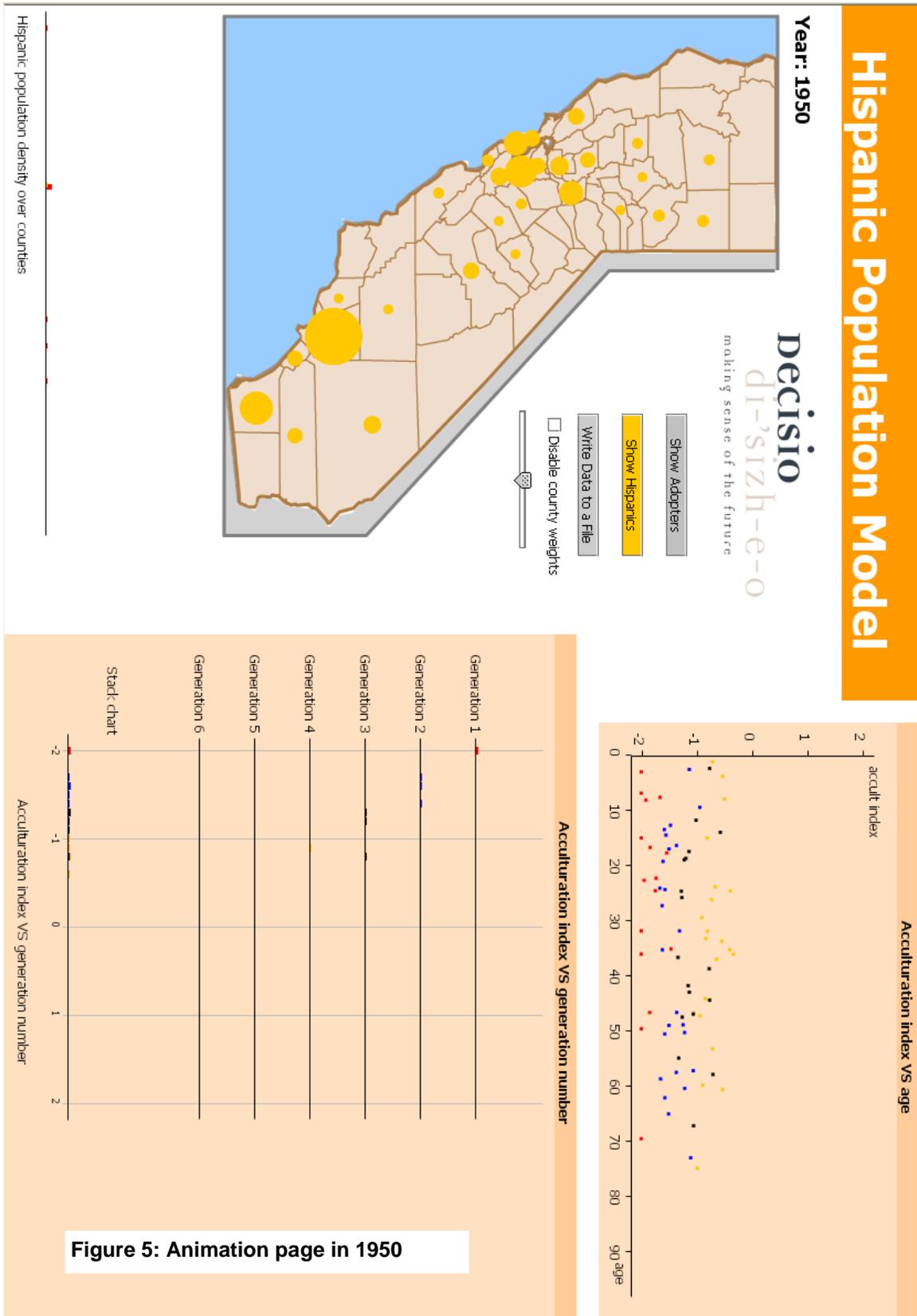


Figure 5: Animation page in 1950

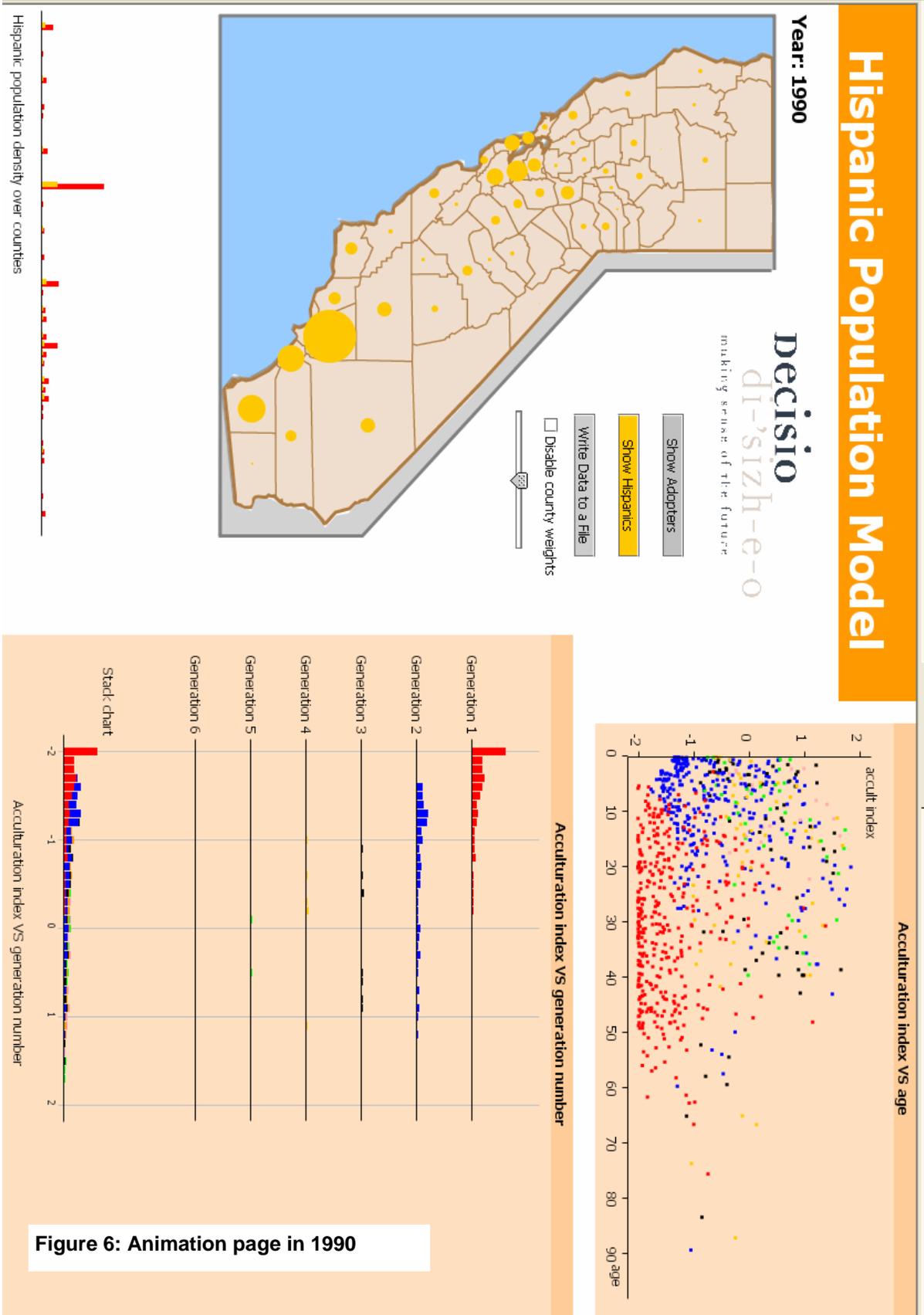


Figure 6: Animation page in 1990

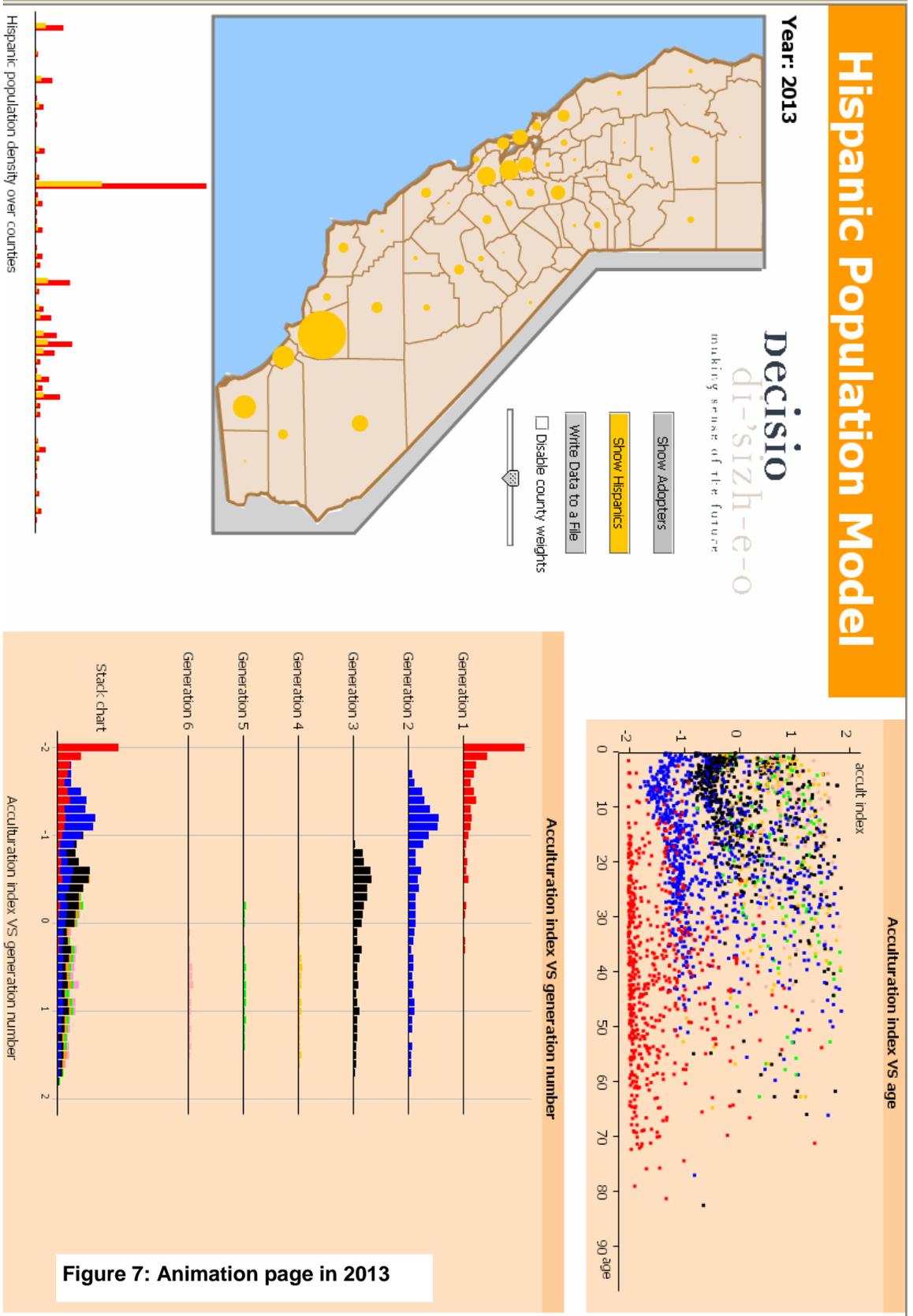


Figure 7: Animation page in 2013

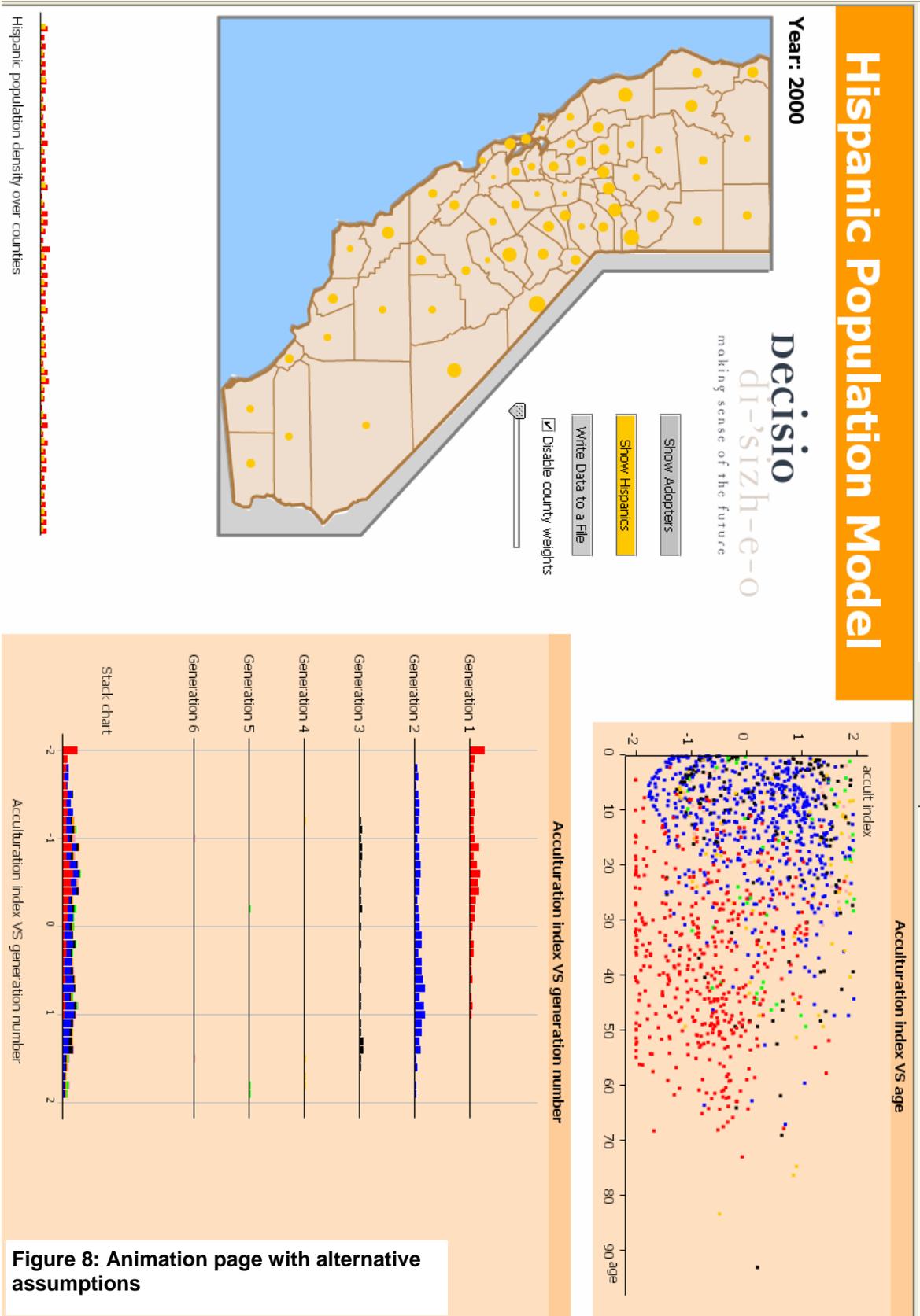


Figure 8: Animation page with alternative assumptions

What's Next?

This model demonstrates the practicality of creating agent models in a systemic context. By adding richness to the behavior of the individual agents more sophisticated systems behaviors can be modeled. For example, consumer purchase behavior could be captured through rich descriptions of an individual level purchase funnel. In addition, we also plan to examine, in more depth, the dynamics of cultural values. Although soft variables such as values are difficult to model, we believe that combining sophisticated survey methods with individual level modeling is promising.