Simulating the Cultural Evolution of Literary Genres

Graham Alexander Sack
Columbia University
English Department
New York, NY 10023
gas2117@columbia.edu

Daniel Wu
Harvard University
Sociology Department
Cambridge, MA 02138
danielwu@fas.harvard.edu

Benji Zusman
University of Florida
Genetics Research Laboratory
Gainesville, Florida 32610
bzusman@gmail.com

ABSTRACT

The purpose of this paper is to explore the evolutionary dynamics of literary genre: the development of the 19th Century British novel is used as a motivating case study. The author constructs an agent-based model in NetLogo consisting of two interacting levels: (1) a genetic algorithm in which cultural forms (e.g., works of literature, pieces of music, etc.) are represented as binary feature strings. Cultural forms evolve across generations via asexual and sexual reproduction. Genres are represented as hierarchical clusters of similar feature strings. (2) Cultural forms are subjected to the selection pressure of consumer preferences. Preferences are heterogeneous: each consumer's tastes are represented by an ideal point in feature space. Preferences are configured in landscapes that vary in their levels of structure, entropy, and diversity. Landscapes are dynamic and may change due to (i) exogenous demographic shifts (e.g., population growth, generational turnover) or (ii) endogenous feedback (e.g., conformity / anti-conformity effects).

Categories and Subject Descriptors

J.5 [**Arts and Humanities**]: Literature; i.6.3 [**Simulation and Modeling**]: Applications

Keywords

Cultural Evolution, Literary Theory, Agent Based Modeling

1. INTRODUCTION

The evolution of literary form and style is an emerging area of academic research and offers a valuable case study in cultural evolution generally. Several notable papers have appeared recently. In "Quantitative patterns of stylistic influence in the evolution of literature," Hughes et al [2] scale up methods traditionally used for authorship attribution to analyze stylistic shifts in the Project Gutenberg literary corpus. In the Genre Evolution Project, Simon and Rabkin at the University of Michigan postulate that literary genre is a complex adaptive system (CAS) and study its properties through the case study of science fiction [4]. Related efforts are underway to 'map the literary genome,' using topic analysis as well as the mining of databases such as Aarne-Thompson-Unter's folktale motif collection [1].

Critic Franco Moretti's essay collection, *Graphs, Maps, Trees* is a provocative and noteworthy example of recent literary evolution research. Based on an analysis of 19th Century British novels, Moretti offers the following speculative claims:

- Due to demographic changes and increasing literacy, the reading public grew substantially from the 18th to the 19th centuries, precipitating a 'phase change' in the form of the British novel circa 1820: novels became far more heterogeneous and generically differentiated, aimed at specialized niches rather than readers in general [3].
- The average lifetime of genres is 25-30 years, the same as a human generation. The reason for this historical rhythm is generational turnover in the reading public [3].
- Literary genre evolution is characterized by alternating cycles of divergence and convergence—that is, periods of increasing generic diversity and differentiation followed by periods of decreasing diversity and cross-over [3].

Statistician Cosma Shalizi argues in his response, "Graphs, Trees, Materialism, Fishing," that while Moretti identifies provocative historical patterns, he stops short of fully articulating the mechanisms underlying and driving literary genre evolution:

> I don't think Moretti's time series, by itself, is enough to begin to let us decide among these mechanisms (some of which are compatible), but I do think it lets us see that some mechanism is called for... One thing Moretti does not do, anywhere, is construct models linking individual behavior to aggregate patterns [5].

A similar criticism may be directed at the other papers cited, all of which are descriptive and based on *a posteriori* statistical analysis of corpora. The objective of this paper is to take up Shalizi's injunction by building a computational model of possible generative mechanisms driving genre evolution. We consider the following questions:

GECCO'13, July 6–10, 2013, Amsterdam, The Netherlands.

- How do the static characteristics and dynamic behavior of the "reading public" affect literary genre evolution?

- How is generic diversity affected by reader diversity? Is there a "phase change" as the reading public grows?

- Under what circumstances will the life cycle of literary genres parallel the life cycle of generations?

2. METHODOLOGY

2.1 Feature Strings and Preference Strings

The model contains two basic agent populations: (1) cultural forms (e.g., books); and (2) cultural consumers (e.g., readers).[1] The key attribute of agents in each population is a bit string of user-specified length. For cultural forms, this bit-string represents the morphological features of the work: for instance, in the case of literature, bits could represent attributes such as authorial style, length, plot, and theme. Because this paper is concerned generalizable dynamics, we have chosen not to calibrate the model against a particular feature set, though this is an objective for future research. For cultural consumers, the bit-string represents an individual's ideal preference. Each consumer has a tolerance for variation from this ideal represented as an acceptable hamming distance. For example, if a particular reader has a preference string [1,1,1,1] and a tolerance of hamming distance 1, then he would be willing to consume cultural works with feature strings [1,1,1,1], [0,1,1,1], [1,0,1,1], [1,1,0,1], or [1,1,1,0].

2.2 Preference Landscapes

Individual cultural consumers are in turn organized into larger preference landscapes, which vary in their levels of structure, entropy, and reader diversity (see figure 1):

- *Unimodal*: This landscape represents circumstances in which the preferences of cultural consumers are comparatively homogeneous, as with mass markets such as family film or Broadway musicals. Preferences are represented as an n-dimensional "sphere" of variable reader population density. The sphere is divided into segments—a randomly-assigned center, a layer of preferences at hamming distance of 1 away, a layer at hamming distance of 2 away, etc.—with the density of readers decreasing as one moves radially outward.

- *Multimodal*: This landscape represents circumstances in which cultural consumer preferences are clustered around two or more contrasting poles. Each individual mode has the same structure as a unimodal sphere of preferences.

- *Spikey*: The limiting case of the multimodal landscape. Preferences are organized into isolated "spikes," with high reader population densities concentrated on single bit-strings surrounded by empty space. This landscape—exemplified by specialist markets such as academic book publishing—represents a case in which reader preferences are sharply divided into separate niches with no overlap and in which the regions of high preference density are uncorrelated with one another.

- *Random Uniform*: Cultural consumers are uniformly distributed over the preference space. This represents a case in which preferences are highly diffuse and have minimal organization and structure.

2.3 Genetic Algorithm

Once the preference landscape has been constructed at set-up, a genetic algorithm is run on the cultural forms in order to simulate evolution. It is worth noting that the use of genetic analogies for the description of cultural artifacts is becoming increasingly accepted: the most well-known example is likely Pandora's Music Genome Project [6], in which songs are atomized into feature sets describing properties such as melody, rhythmic structure, and instrumentation. In this model, the genetic encoding is a bit-string representing the features of each book. Each reader can consume a user-specified maximum number of books in each time period. The fitness of each book is measured by the number of readers it receives in that period. High fitness books are selected by tournament and are more likely to survive and reproduce, increasing their influence on the genetic content of the next generation. Three reproductive mechanisms are used:

- *Survival*: books carry over from generation T to T+1 with no genetic change

- *Asexual*: individual bit-strings from generation T are copied with a user-specified probability of mutation to create a new generation of books at T+1

- *Sexual*: pairs of bit-strings from generation T are spliced in order to create a new generation of books at T+1

Each of these reproductive strategies has an intuitive interpretation in the context of cultural production. Survival corresponds to the case in which successful books are simply reprinted. Asexual reproduction corresponds to the case in which successful books spawn similar works with slight variation: that is, authors copy and modify the template provided by recently successful works. Sexual reproduction corresponds to what we might call "genre-crossing": authors take the features of two successful works and synthesize them in order to produce a new work.[2] The relative proportions of these reproductive strategies are user-specified. Users also specify the mutation rate, which is the probability that any bit will be mutated during either reproduction process. The

[1] The terms "cultural form" and "book" will be used interchangeably over the course of this paper, the latter being a special case of the former. The same applies to the terms "cultural consumer" and "reader." Although the model was designed with literary genre in mind, it is sufficiently abstract that it applies to cultural artifacts generally.

[2] The current trend of "mash-up" literature provides a salient example. Best-sellers such as *Abraham Lincoln: Vampire Hunter* splice the features of already-successful genres (e.g., historical biography and gothic). Lest we dismiss such works as gimmicks, it is worth recognizing that many high-prestige genres emerged through hybridization. Modernist works such as James Joyce's *Ulysses* self-consciously combined the features of the realist novel with those of the classical epic. *Pastiche*, *bricolage*, and the combination of high and low art were central to postmodern literature, epitomized by William Burrough's "cut-up" novels. Recombination seems to be a widely-used mechanism in literary production.

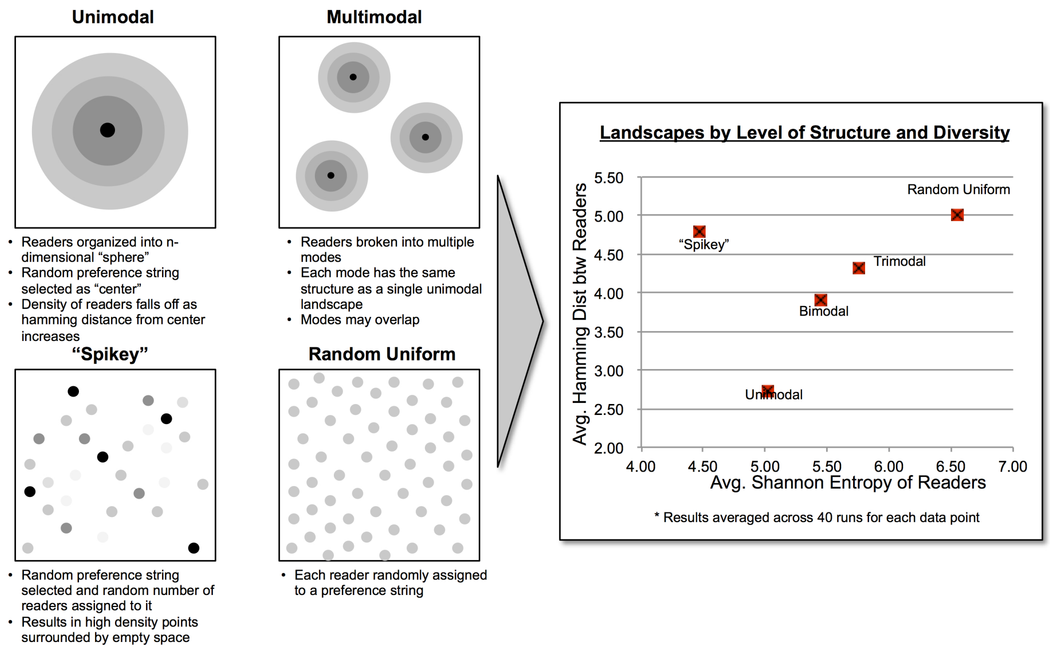

Figure 1: Cultural Consumer Preference Landscapes with Key Topological Features

mutation rate also has an intuitive interpretation in the context of cultural production: it characterizes the inherent creative experimentalism of the cultural field, that is, how far authors are generally willing to depart from established models.

2.4 Clustering Algorithm

The focus of this study is not individual literary works, but rather the aggregate genres into which they are organized. "Genre," however, is an ambiguous concept. It may be conceptualized in at least two ways: top-down or bottom-up. In the top-down case, genre functions as a generalized formula or set of conventions that precedes literary works and according to which they are constructed or judged. Alternatively, genre may function as a bottom-up phenomena: genres, in this case, are constructed after the fact as category labels for works that have similar characteristics.

For the sake of this paper, we are concerned only with genre in the bottom-up sense. To simulate this, we cluster books based on the statistical similarity of their feature-strings. A hierarchical clustering algorithm is employed to group books based on the average hamming distance between them. The algorithm works as follows: (1) the user specifies a minimum hamming distance to be used as a cut-off[3]: clusters that are separated by hamming distances above this cut-off are considered to be in separate genres; (2) each book is initially placed in its own cluster; (3) each iteration, the algorithm merges the clusters separated by the minimum pairwise average hamming distance provided it is less than the cut-off; if not, the algorithm halts.

To measure generic diversity, several metrics are used. These include the number of clusters as well as the following:

$$ShannonEntropy = \sum_i p_i log(p_i) \quad (1)$$

$$Diversity = \frac{1}{\sum_i p_i^2} \quad (2)$$

These metrics are calculated based on the probability distribution of books and are helpful for differentiating cases in which there are equal numbers of clusters, but books are distributed unevenly across them.[4] The Shannon entropy of clusters can be interpreted as the average number of questions one would have to ask to determine what cluster a book is in. Shannon entropy decreases as (1) the number of clusters decreases, (2) the frequency distribution of books across clusters becomes more skewed (and therefore more predictable). Diversity is measured by the Inverse of the Herfindahl-Hirschman Index (henceforth, IHHI), a measure of market fragmentations widely used in economics.

2.5 Dynamic Landscapes

Thus far, we have assumed one-way causality: the preference landscape is (1) static and (2) formed in a vacuum. This is unrealistic: preferences change over time. Moreover, the prevailing culture shapes preferences just as preferences shape the culture. To address this, the model incorporates several landscape updating processes.

First, we allow for demographic changes, such as growth in the size of the "reading public." We also include generational effects: sub-sections of the population update preferences synchronously at discrete time intervals.

[3]A user-specified cut-off is preferred to specifying the number of clusters (as is required by K-means), since this would defeat the purpose of cluster statistics as a metric for genre diversity.

[4]For example, 100 books may be distributed across 4 clusters as [25, 25, 25, 25] or as [97, 1, 1, 1]. Both cases have the same number of clusters, but the Shannon entropy in the first case will be 2 bits, while in the later case it will be approximately 1.

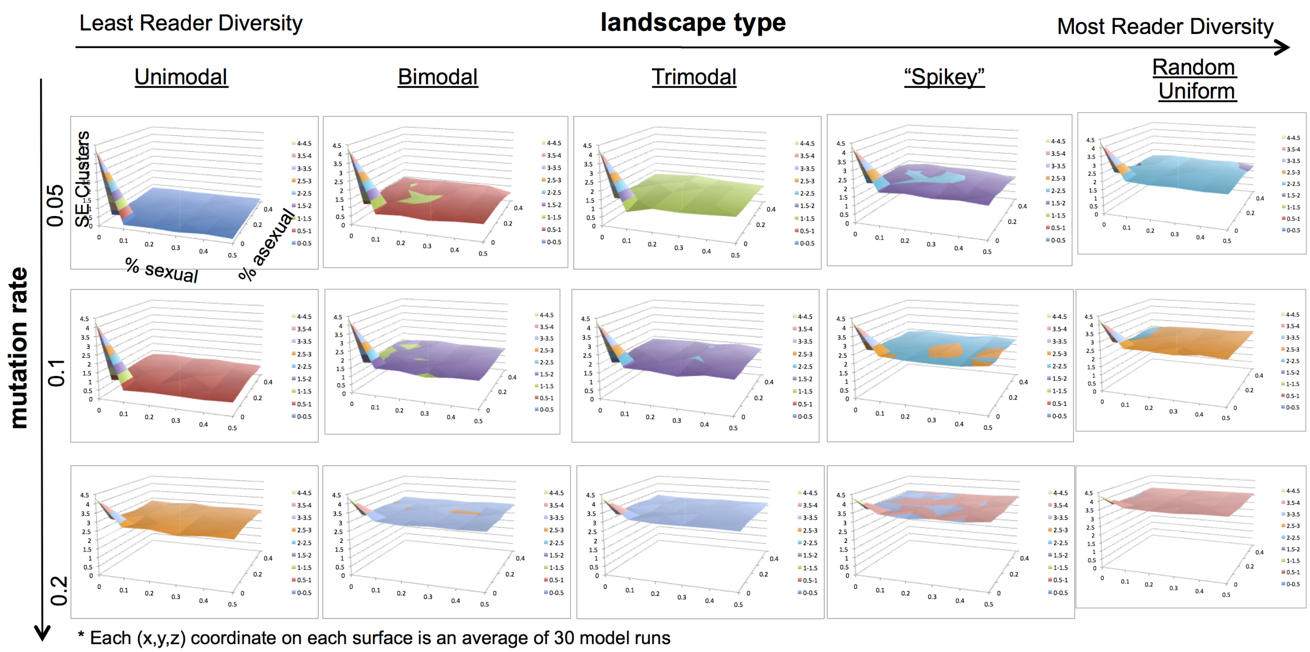

Figure 2: Shannon Entropy of Clusters vs. Preference Landscape and Mutation Rate

Second, we allow for feedback effects, whereby popular cultural forms influence preferences. We assume the population is divided into conformers and non-conformers. Conformers update their preferences towards forms that are currently popular. Non-conformers update their preferences away from forms that are currently popular.

3. RESULTS

The next several pages summarize insights concerning the impact on genre diversity from (i) preference landscapes features, (ii) demographic changes such as population growth and generational turnover, and (iii) feedback between reader preferences and dominant cultural forms. Briefly summarized, these insights are as follows:

- Generic diversity cannot be explained solely in terms of the preferences and characteristics of consumers: we also need to account for the *artistic process*, such as the level of creative experimentation, which is represented in the model by mutation rate.

- Growth in the "reading public" does not in and of itself guarantee either an increase in reader diversity or generic diversity. Under certain conditions, market growth may actually *decrease* generic diversity.

- The model predicts that dramatic changes in the preferences of cultural consumers—analogous to ecosystem disruption—lead to increases in creative experimentation (i.e., the cultural mutation rate).

- The preferences of conformist consumers, whose tastes are relatively stable and predictable, have a highly disproportionate effect on the level of generic diversity relative to the rest of the consumer population, producing "phase change" dynamics.

3.1 Static Landscapes

Figure 2 shows the results from evolving the cultural forms on different static preference landscapes. Results are displayed in a five-dimensional graph: the vertical axis of each miniature surface measures the shannon entropy of clusters (a metric for generic diversity), while the horizontal axes show the crossover rate and asexual reproduction rate, which measure the mix of processes by which new cultural forms are generated. The surfaces are organized along two larger axes: (1) landscape type and (2) mutation rate.[5]

Figure 2 yields several insights about genre, some intuitive and some surprising.

First, as we move horizontally from highly homogeneous preference landscapes such as unimodal and bimodal on the left to highly heterogeneous landscapes such as "spikey" and random uniform on the right, the surface representing the Shannon entropy of clusters shifts upward. Stated simply, as reader preferences become more heterogeneous, generic diversity increases. This makes intuitive sense: a small set of genres is adequate to meet the tastes of a homogeneous mass audience while a diverse set of genres is required if there are many separate niches.

More interestingly, figure 2 suggests that generic diversity is unaffected by the mix of reproductive strategies—that is, *how* authors create new works, whether by modifying vs. splicing existing templates, has little impact on the ultimate level of generic diversity. However, the *mutation rate* does have a significant impact on the generic diversity. This result is less obvious. The mutation rate, as noted above, captures the "experimentalism" of a particular creative market and can be thought of as the predilection of writers or

[5]The following values are held constant: # readers = 100; # books = 50; bit-string-length = 10; tolerance = 1 hamming distance; maximum # books a reader can consume in a period = 3.

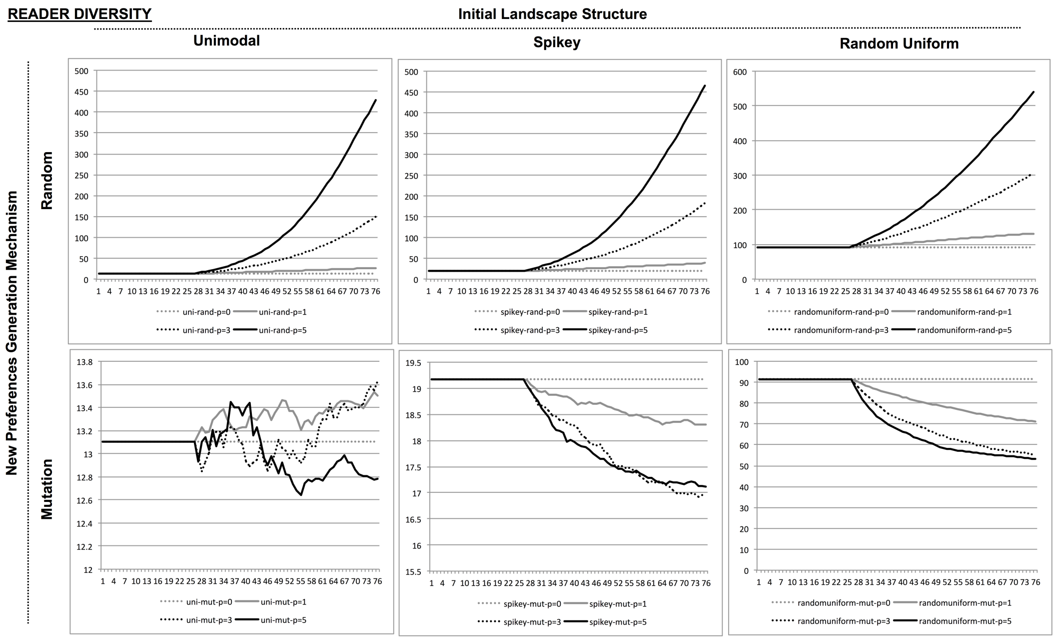

Figure 3: Impact of Population Growth on Diversity of Reader Preferences

artists to adhere closely to established, successful models or to depart from them significantly. What figure 2 suggests is that highly experimental creative markets will produce greater generic diversity *regardless of the structure of consumer preferences*. For example, a mutation rate of 0.2 on the unimodal landscape produces the same level of generic diversity as a mutation rate of 0.1 on a random uniform distribution (Shannon entropy of 2.5-3). This illustrates that increasing creative experimentation in an environment of homogeneous preferences has an effect on generic diversity equivalent to fragmenting the preference landscape.

This means that *generic diversity cannot be explained solely in terms of consumer preferences*: in order to explain the levels of generic diversity that we observe in cultural forms we also need to account for the *artistic process*, in particular, how experimental vs. conservative it is. This is a point underemphasized by Moretti, who explains the diversification of literary genres in the 19th Century exclusively in terms of the reading public.[6] The model shows that reader preferences are at best a partial mechanism and that the level of creative experimentation in a cultural market at a given historical moment is also a crucial input to explaining generic diversity.

3.2 Population Growth

The prior section concerned genre evolution on static preference landscapes. We next consider genre evolution on landscapes that shift over time due to demographic changes.

[6]"Normal literature remains in place for twenty-five years or so... but where does this rhythm come from?... The causal mechanism must be external to the genres and common to all: like a sudden, total change of their ecosystem. Which is to say: a change of their audience. Books survive if they are read and disappear if they aren't: and when an entire generic system vanishes at once, the likeliest explanation is that its readers vanished at once. This, then, is where those 25-30 years come from: generations." [3]

Figures 3 shows the impact of growth in the reading population. The graphs summarize various combinations of 3 key parameters: (1) the initial landscape, which can be unimodal, "spikey," or random uniform; (2) the rate of population growth (0%, 1%, 3% or 5%); (3) how new reader prefereces are generated. If the preferences of new readers are generated by random assignment, they will tend to differ from those of existing readers, whereas if they are generated by mutation they will tend to be similar to those of existing readers.

Figure 3 shows that the relationship between the initial landscape and the new preference generation mechanism is crucial. As noted above, the initial landscapes can be homogeneous or heterogenous: unimodal has minimal reader diversity (IHHI = 13), followed by spikey (IHHI = 19), and random uniform (IHHI = 90). If new readers have randomly assigned preferences, then all three landscapes respond the same way: reader diversity increases exponentially in time regardless of starting conditions. This is not the case, however, if new preferences are mutations of those composing the original landscape. In this case, the landscapes respond differently. For spikey and random uniform landscapes, reader diversity actually *decreases* as new readers are added. The intuition is as follows. For spikey and random uniform landscapes, preferences are initially highly diverse. Adding mutations of existing preferences to the landscape means adding readers that are *similar to those that already exist*, making these landscapes more homogeneous. For unimodal, the effect of adding new readers is uncertain: it may increase or decrease reader diversity.

Although not shown here due to lack of space, the effect of each growth scenario on generic diversity parallels the effect on reader diversity.

This analysis elucidates a gap in Moretti's claim that "the growth of the market creates all sorts of niches for 'specialist' readers and genres" [3]. *Market growth does not in and*

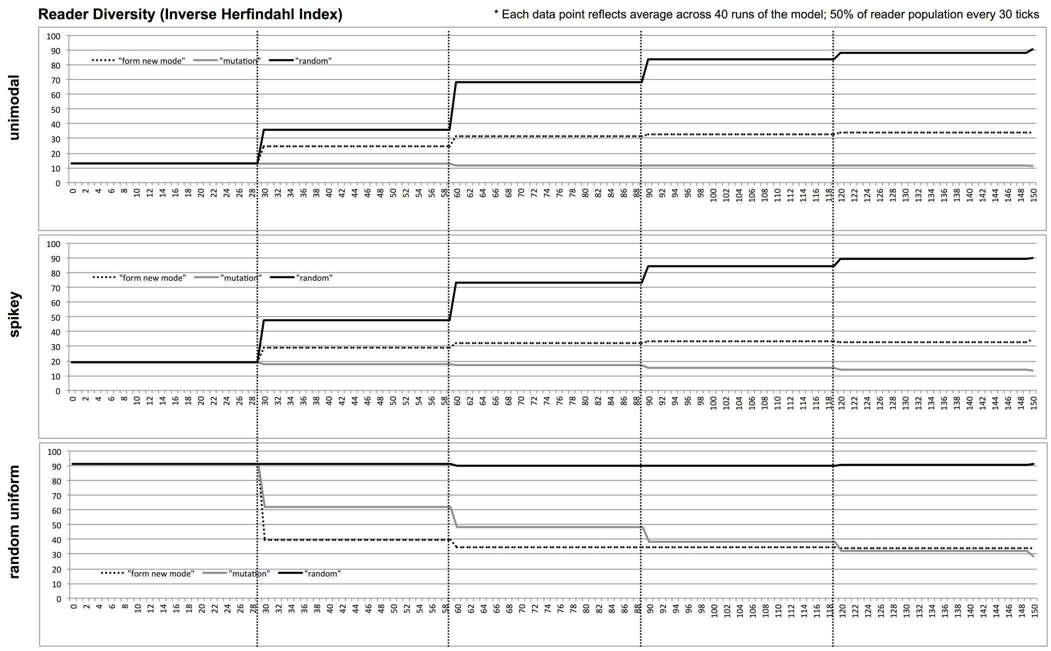

Figure 4: Impact of Generational Turnover on Diversity of Reader Preferences

of itself guarantee an increase in either reader diversity or generic diversity. In fact, market growth may actually *reduce* reader and generic diversity under certain conditions. Whether the growth of a literary market increases diversity or not depends crucially on (1) whether the initial condition of the market was homogeneous vs. diverse and (2) whether new readers have preferences that are similar to or different from the readers who already populate that market.

3.3 Generational Turnover

In the case of population growth, new readers are introduced gradually, but this is not the only kind of demographic change possible. Reader preferences may also change in a sudden and highly synchronized fashion. For example, an entire generation of readers may die off in a short span of time or some significant historical or technological event may alter cultural preferences simulateneously, as with the advent of radio or television.

To model generational change, we begin with static preference landscapes of varying heterogeneity (unimodal, spikey, random uniform) and then "shock" them every 30-50 ticks by killing half the agent population and replacing them with new readers with new preferences. These preferences are determined by (1) random assignment; (2) mutation of the preferences of older readers; or (3) formation of a new mode. The first two cases are the same as for the population growth scenarios. The third case accounts for the possibility—unique to generational turnover—that new preferences may be heterogeneous *inter*-generationally but homogeneous *intra*- generationally (e.g., the old generation prefers classical music while the young generation prefers rock and roll). Figure 4 shows the effect of generational shocks on different reader landscapes with various preference updating mechanisms.

The other key modification we make to the model is to endogenize the cultural mutation rate. Recall that the mutation rate represents the level of creative experimentation in a cultural market: the greater its value, the more formal variation there is between books at time T and time T+1. For previous simulation scenarios, this value was exogeneously determined. In the case of generation turnover, however, we endogenize it: cultural forms inherit both a feature string and a mutation rate from their parents and the population evolves to a steady-state mutation rate on its own accord in response to environmental factors.

Figure 5 shows the effect of generational turnover on the cultural mutation rate. Initially, the mutation rate converges to a steady-state level of 5%. Thereafter, periodic shocks to the preference landscape cause the mutation rate to temporarily spike to 7% before reverting to the steady-state. This result makes intuitive sense: a high level of creative experimentation is evolutionarily advantageous when the preferences of cultural consumers are uncertain or in transition, as is the case immediately after a shock to the preference landscape. In a biological context, increased mutation rates are known to be a defense mechanism for species level survival when habitats are disrupted: a sudden increase in diversity helps to improve the odds that at least some individuals will be well-adapted to the new environment and survive. In the context of this simulation, cultural forms fan out through feature space after a shock until the preferences of the new consumers are discovered. Thereafter, the genetic codes stabilize and the mutation rate falls: creative experimentation ceases to confer benefits and becomes detrimental as mutating forms drift away from the regions of high reader density already found.

The model's prediction that cultural "ecosystem disruption" leads to a temporary increase in creative experimentation invites a search for historical examples. One anecdotal case is the dissolution of the Hollywood studio system in the early 1960s. Many film historians argue that with cinema audiences in rapid decline due to competition from television, Hollywood studios, which had previously relied on for-

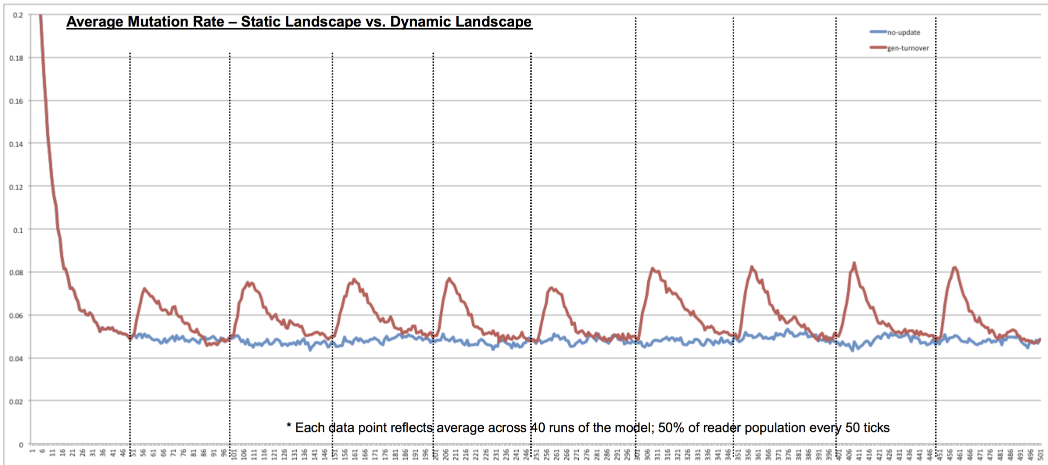

Figure 5: Impact of Generational Turnover on Endogenous Mutation Rate

mulas for box office success unchanged since the golden age of the 1930s, loosened their creative control and permitted auteur directors to experiment and take formal risks. This period of experimentation, however, was finite, ending in the early 1970s when studios hit on a new formula for box office success—the "blockbuster," exemplified by films such as "Jaws." The diversification of cinema in the 1960s might be regarded as a miniature Cambrian explosion in which the disruption of the traditional audience base precipitated the accelerated mutation of extant cultural forms.

3.4 Feedback Effects

The previous sections assume one-way causality *from* consumer preferences *to* cultural forms. In this section, we consider scenarios in which cultural forms evolve on a dynamic landscape with two-way causality: preferences affect cultural forms, but cultural forms also affect preferences.

We divide the reading population into two segments, (1) conformers and (2) non-conformers. Conformers update their preferences towards popular books, while non-conformers update their preferences away from popular books.[7]

Graphs (a), (b), and (c) in figure 6 show how key outputs—average books per reader, number of clusters, Shannon entropy of clusters—vary as the percentage of conformers is increased from 0% (everyone is a non-conformer) to 100% (everyone is a conformer).

The first salient result is that feedback effects extremize the outcomes. With a high percentage of conformers, the average number of readers per book converges to 3, the maximum allowed. The reason is clear: when 100% of the readers are conformers, their preferences all eventually collapse to the feature set of the most popular book. The opposite is true for high percentages of non-conformers.

[7]Each tick, conformers check if the most popular book is within a tolerable distance of their preferences (meaning they are willing to consume it). If so, they do nothing. If not, a conformer compares his preference-string to the most popular book and identifies bits where they differ. He then flips one of his bits, shifting his preference-string one hamming distance closer to what is popular. Non-conformers do the opposite: they compare their preference-strings to the most popular book and update so as to move one hamming distance away.

Second, the behavior of each landscape is effectively identical. Unlike with population growth, in which the dynamics differ between landscapes, for conformity-effects the dynamics overwhelm differences between the initial conditions. Each conformer vs. non-conformer combination has its own asymptotic landscape to which the preferences converge regardless of their original configuration.

Third, consistent with intuition, as the level of conformity increases, generic diversity decreases. However, it takes only a small percentage of conformers for generic diversity to collapse to its minimum. Both the number of clusters and the Shannon entropy of clusters have a pronounced L-shape indicative of a *phase change*: the generic diversity falls steeply as the percentage of conformers increases from 0 to 30%, but then levels off. Generic diversity barely changes as the percentage of conformers is raised from a minority of 30% to a majority of 100%. One explanation is that conformers are a consistent population, while non-conformers are transient. Once they converge on a popular book, conformers' preferences do not change any further due to positive feedback. For non-conformers, on the other hand, there is a negative feedback effect. Even if a book manages to capture non-conformer readers for a period, its popularity will cause those non-conformers to change their preferences shortly thereafter, so the book eventually loses its readers. Because the non-conformers cannot support a consistent book population, stable genres develop around the conformers, even if they compose only a small percentage of readers. This has interesting implications for cultural markets: it suggests that the preferences of a large segment of the population may have no effect on the diversity of cultural products. Product categories develop primarily around a small but tractable population of conformers.

4. CONCLUSION

The majority of agent-based models of cultural evolution are focused on what might be termed "internal culture"—e.g., belief propagation or the emergence of behavioral norms such as pro-sociality. Comparatively few models explore what might be termed "external culture"—that is, the artifacts, forms of expression, tools, and technologies that humans make. Whereas the former class generally understand

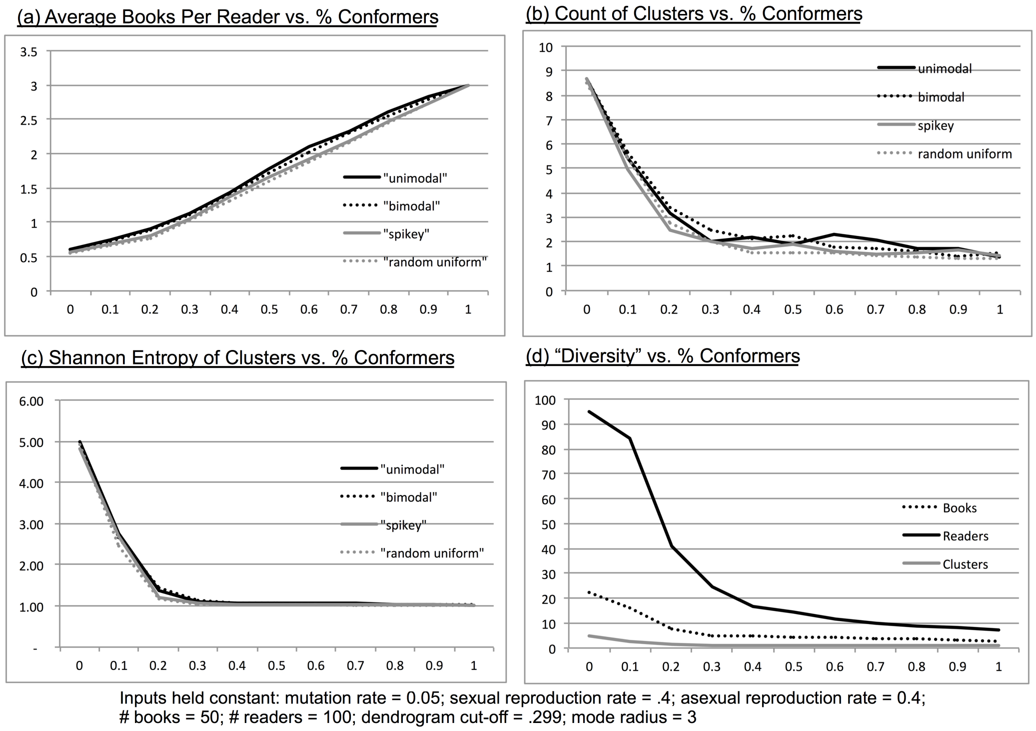

Figure 6: Impact of Percentage of Conformers vs. Non-Conformers on Key Metrics

cultural evolution as a process of social learning, the latter understand cultural evolution as morphological change in physicalized cultural artifacts over time. This approach calls for different methods. Whereas most models of internal culture use a single population of agents who hold particular beliefs or operate according to particular norms, modeling of external culture more naturally calls for two populations: (i) cultural forms that evolve under selective pressure and (ii) human agents that consume and/or create them. This study is intended as an initial effort to develop such a model.

The results suggest a number of preliminary insights about plausible mechanisms driving the evolution of cultural forms generally and literary genre specifically. First, generic diversity cannot be explained solely in terms of the characteristics of the reading public: we also need to account for the characteristics of the creative process, in particular, the level of experimentation in the cultural market at a given historical moment, represented in this model by the mutation rate. Second, we show that growth in the reading public is not sufficient to guarantee an increase in either reader diversity or generic diversity. In fact, market growth may actually reduce diversity under certain conditions. To determine the effect that growth will have, we need to know whether the preference landscape was initially homogeneous vs. diverse and whether new readers have preferences that are similar to or different from the readers who already populate that market. Third, the model predicts that disruptions to reader preferences, such as those caused by generational turnover, result in increased creative experimentation. Lastly, we find that genres and product categories develop based on the preferences of conformist consumers, who have more reliable and predictable preferences.

Although the model above addresses a set of claims about literary genre, the implementation is quite general, relying only on abstract feature and preference strings that could represent virtually any product—cultural or otherwise—that can be atomized into variable features: consumer preferences for toothpaste or automobiles could be modeled in a manner analogous to novels or films. The simulation developed here is, fundamentally, an agent-based model of product differentiation and in future research it will interesting to benchmark our results against those of established economic models, such as the classic Hotelling line.

5. REFERENCES

[1] Daranyi, P. Wittek, L. Forro. "Toward Sequencing 'Narrative DNA.'" *Proceedings Computational Models of Narrative*. LREC. Istanbul, 2012.

[2] Hughes, J., Foti, N., Krakauer, D., Rockmore, D. "Quantitative patterns of stylistic influence in the evolution of literature." *PNAS*, May 2012.

[3] Moretti, Franco. *Graphs, Maps, Trees*. Verso: New York, 2005.

[4] Rabkin, E. and Simon, C. "Culture, Science Fiction, and Complex Adaptive Systems." *Biocomplexity at the Cutting Edge of Physics, Systems Biology, and Humanities*. Bologna: Bononia University Press, 2008.

[5] Shalizi, Cosmo. "Graphs, Trees, Materialism, Fishing." *Reading Graphs, Maps, Trees: Responses to Franco Moretti*. South Carolina: Parlor Press, 2011.

[6] Westergren, T. Pandora's Music Genome Project. http://www.pandora.com/about/mgp, 2013.