THE UNIVERSITY OF CHICAGO

MANAGED ABUNDANCE: A QUANTITATIVE HISTORY OF AMERICAN FICTION, 1931-2009

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE HUMANITIES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF ENGLISH LANGUAGE AND LITERATURE

BY

JORDAN PRUETT

CHICAGO, ILLINOIS

AUGUST 2022

Table of Contents

List of Figures iv

List of Tables v

Acknowledgements vi

Abstract viii

Introduction: Problems of Abundance 1

0.1. The Conceptual Problem: Canons as Unrepresentative Samples 7

0.2. The Historical Problem: Literature in a Crowded Field 15

0.3. Chapter Summaries 23

Chapter 1. The *New York Times* Bestseller List and the Rise of the "Mass Market Hardcover" . 27

1.1. The *New York Times* in the Universe of Bestseller Lists 30

1.2. Historical Bestsellers, Bestselling Histories 36

1.3. The Mass Market Hardcover? 52

1.4. Conclusion: Seeing Like a Publisher 59

Chapter 2. Literary Brand Names and the Political Economy of Authorship 62

2.1. What Is An Author? 70

2.2. Authorship and the Theory of Value 77

2.3. The Author as Book Packager 87

2.4. Franchise Fiction 94

2.5. The Undeath of the Author 100

2.6. Conclusion: Literary Developers 108

Chapter 3. Dimensions of Prestige: A Cluster Analysis of the *Book Review Index*, 1965-2000 112

3.1. The American Book Review Market 117

3.2. Data from the *Book Review Index* 127

3.3. Journals Find Their Niche 135

3.4. An Alternative Model of Prestige 147

3.5. Conclusion: Clusters as Canons 158

Coda: On Literary Data 161

Appendix 165

Bibliography 169

List of Figures

1.1. Bestsellers in HathiTrust by year of first appearance 38
1.2. Genre breakdown of NYT bestsellers (five-year totals) 46
1.3. Top 200 words in Piper and Portelance (2016) as a proportion of all bestseller words 50
2.1. Dual-authored NYT hardcover fiction bestsellers, 1946-2020 62
3.1. Histogram of total reviews of each book in the *Book Review Index* data 129
3.2. Histogram of total reviews published by journal 133
3.3. Mainstream branches of the journal dendrogram 141
3.4. Network diagram of 15 most-reviewed authors within selected clusters 142
3.5. Authors' reviews in NYC magazines and mainstraim wide-coverage periodicals 154
3.6. Authors' reviews in conservative publications and liberal-progressive publications 155
3.7. Authors' reviews in literary magazines and major city newspapers 155

List of Tables

1.1. Top words and sample novel for Topic 3 41
1.2. Top 5 topics in three different five-year periods in the bestseller corpus.............................. 43
1.3. Top words by coefficient for genre classifiers.. 48
3.1. Most frequently reviewed books.. 130
3.2. Most frequently reviewed authors 134
3.3. Ten largest clusters discovered by hierarchical clustering 138

Acknowledgements

Many friends, colleagues, and advisors deserve credit for helping bring this dissertation to fruition. Though their names are too abundant to list in full, I want to thank some of them in particular. First of all, I would like to thank the three members of my dissertation committee: Patrick Jagoda, for his close readings and conceptual provocations, as well as advice and encouragement from day one of the graduate program; Sianne Ngai, for her conceptual clarity and for helping me avoid at least one embarrassing misreading of Marx; and Deborah Nelson, for many invaluable conversations about the publishing industry, and for many more about how to actually write a dissertation.

I owe a debt of gratitude to all those who read this project and provided feedback while it was being drafted. In particular, I would like to thank the participants of Deborah Nelson's dissertation writing group, including Joel Rhone, Dana Glaser, Sarah McDaniel, Sophia Sherry, and Vinh Cam. Their feedback greatly strengthened the argument of the dissertation, and their camaraderie eased the burden of writing it. Additionally, thank you to the participants of the 20th and 21st Century Cultures Workshop at the University of Chicago, for feedback on an early draft of Chapter One. Particular thanks to Hoyt Long.

For the third chapter in particular, thanks go to Dan Sinykin and Richard So, who generously provided the dataset on which the entire chapter is based. Thanks as well to you both for many generous conversations about quantitative methodology and the argument of that chapter.

Thanks go to Clovis Gladstone for providing me a wealth of advice, information, and practice with computational methods during my internship at the Textual Optics Lab. I could not have completed the research underlying this dissertation had I not first stumbled my way into

computational literary studies by way a position at the lab. Though I asked Clovis for that internship on a whim, it was probably the best whim I had in my six years of graduate school.

Above all, I wish to thank Katie Nolan, for many, many hours of conversation about every aspect of the dissertation, for encouragement, for proofreading, and for a dozen other forms of practical and personal support that I don’t have room to mention. To you I owe debts which—unlike novels—cannot be quantified.

Abstract

Over the past two decades, the adoption of computational methods by literary historians has changed the shape of literary history. Computer-assisted reading tools have permitted scholars to ask questions about cultural and linguistic trends at a scale that was previously impossible. Methods like topic modeling have enabled scholars to write new literary histories that consider a sample size of many hundreds or thousands of texts at a time. Large digitized university collections of popular, noncanonical, and out-of-print works, too vast to be sorted through and analyzed manually, can be summarized by the tireless eye of a probabilistic model.

But in this work, the scale of literary history has generally been approached as a conceptual rather than historical problem. Although computational literary scholars have considered the challenge of analyzing a digital library of 200,000 texts, they have paid less attention to what it meant that actual, historical literary markets were similarly comprised of many hundreds of thousands of circulating texts. Unlike computational literary scholars, the social actors who encountered those texts—readers, publishers, and authors—have generally not had access to sophisticated computational tools to help them make sense of it all.

This dissertation considers how social actors managed this problem in one historical context: the literary market in the United States between about 1931 and 2009. This was a period of considerable expansion for American publishing, whether measured in terms of new titles or total books printed. As the number of books in print expanded, publishers faced a daunting problem: how could they reliably match readers, who have diverse tastes, with novels, which are hugely varied in their style and content?

I consider three different institutional adaptations that emerged in order to manage this problem. First, I consider weekly bestseller lists, which manage abundance in the simplest way:

by advertising what is currently popular. I describe the history of the most influential bestseller list, the hardcover fiction list of the *New York Times*, and map out the rise and fall of its most prominent genres in the postwar period. I argue that the types of books likely to appear on its pages gradually shifted as mass market genres began to be published in hardcover. Second, I consider what has been called the “brand name author,” or an author whose name becomes a stand-in for a larger literary franchise. I show that such arrangements became more prevalent in mainstream hardcover fiction towards the end of the twentieth century. I argue that brand-name authors became increasingly attractive to publishers because they provided name-recognition and an avenue for repeat consumption, qualities which were historically difficult to establish and sustain in publishing. Finally, I consider the book review economy, which handles the problem of abundance by delegating the task of sifting through new books to professional reviewers. Working with a dataset derived from book reviews indexed by the *Book Review Index* between 1965 and 2000, I explore aggregate patterns of shared taste across different popular reviewing outlets. Borrowing methods from cluster analysis, I show that review publications tended to cohere into specialized clusters of like-minded reviewers who predictably reviewed the same book.

Introduction: Problems of Abundance

This is a dissertation about the recent history of a simple fact: the United States in the twenty-first century produces an astonishingly large number of books every year. According to estimates from the book industry analyst R.R. Bowker, in 1995, approximately 113,000 new titles were published in the United States, not counting books published by "nontraditional" means, such as self-published books.[1] By 2005, that number was closer to 190,000, and by 2009, it was something like 288,000. If one counts self-published books, the large majority of new titles, that number is even bigger: over 1,050,000.[2] These remarkable numbers sometimes slip under the radar in reporting on the state of textual culture in the United States, which generally prefers to gloomily speculate about the meaning of the latest Gallup poll showing that Americans appear to be reading less. Whatever the significance of such polls, the statistics on the production of new books are unambiguous: the United States in the twenty-first century publishes a huge and growing number of books.

In this dissertation, I attempt to make sense of what this fact means for the history of literature. But I don't approach this question from the perspective of those disciplines in which statistics about publishing trends are most often cited, such as studies of the economics of the publishing industry. Instead, I do so from the perspective of literary history, a field traditionally more invested in understanding the rise and fall of literary movements than in producing or interpreting statistics about books. As I will argue, taking account of the magnitude of book

1. Editors, "Fewer Titles = Good News: Some Call for More Cuts in Output," *Publishers Weekly*, May 15, 2006): 4.

2. Editors, "Big Gains for New Players," *Publishers Weekly*, April 19, 2010: 5.

production opens up new insights about American literary culture. Though many of the questions posed by this dissertation have relevance in multiple periods, I focus in particular on the latter half of the twentieth century, both because it is the context with which I am most familiar and because it was a period of substantial expansion in the publishing industry. By understanding the crowded field of competitors into which all aspiring authors must insert their work and from which all readers choose what to read, we can gain greater clarity about why American literature looks the way that it does. I give additional details about the period covered by the dissertation at the end of this introduction.

The magnitude of contemporary book production has not gone unremarked in literary studies. But it has mainly been a topic of discussion among scholars for whom it is more an object of methodological than historical interest: those applying computational methods to literary history, working in the field that has variously been called quantitative formalism, cultural analytics, digital humanities, or, simply, computational literary studies. A decade ago, it was sometimes common in discussions of computational methods to appeal to the scale of the historical literary archive—and its continual and rapid growth—as the *raison d'être* for making use of such methods. The logic of the argument is simple enough. Literary scholars face a problem of scale: the set of extant texts available for study even within a narrow cultural or historical context is simply too vast to be mastered by scholars, or even read by them. There is just too much—thousands, millions, of books, the majority of which would be unfamiliar to even the most well-read literary historians. How can we then claim to fully understand literary history? Matt Wilkens has called this the "problem of abundance."[3] Restricting himself to the

3. Matt Wilkens, "Contemporary Fiction by the Numbers," *Post45* (March 11, 2011). https://post45.org/2011/03/contemporary-fiction-by-the-numbers/

case of fiction, Wilkens points out that, by 2010, the publishing industry in the United States released at least 50,000 new novels every year. Of those many novels, scholars “read a couple dozen, see reviews of maybe that many again, and so could claim to know something about roughly 0.1% of the total.”[4] One tenth of one percent: only a small sliver of the full population, and unlikely to be a representative sample. Literary scholars therefore need methods by which “to extract information from and about books without spending much time on them individually.”[5] For Wilkens, computational methods fit the bill, allowing scholars to extract information about many texts in a manageable period of time, enabling them to study large-scale patterns that might not be otherwise visible.

However, Wilkens is focused on the problem of abundance as an epistemological problem rather than a historical one. The practical and material side of literary abundance goes unremarked: the ways that abundance structures the organization of the actual literary market, as well as the choices available to those who participate in it. Literary scholars aren’t the only ones who “can’t read it all.” Much as scholars necessarily encounter only a fraction of the literary material produced by a given culture, any would-be reader of literature cannot possibly consider more than small portion of all the books available to them when choosing what they wish to read next. This perhaps-obvious fact is hugely important for understanding the literary marketplace. It is something all-too-familiar to authors, booksellers, and publishers, who worry anxiously about how to maximize the chances that their voices will be heard amongst the cacophony of competing voices. This side of literary abundance has been relatively undertheorized in conversations about quantitative method, which have focused more on abundance seen “from the

4. Wilkens.
5. Wilkens.

outside," from the perspective of the scholar or analyst. But what does literary abundance look like from the inside: from the perspective of readers, authors, and publishers?

This dissertation contributes to recent debates about quantitative method in literary studies by theorizing abundance from both angles at once: as a methodological problem of how to represent the totality of literary production in a given time and place and as a practical facet of actual, historical literary markets that structures the behavior of the actors making decisions within those markets. Although some of the claims that I make in the dissertation are applicable in a variety of historical and cultural contexts, I restrict my analysis to literary history as it played out in the United States in primarily the second half of the twentieth century. American literary history in this period is especially fertile ground for considering how the problem of abundance changes how we theorize literary history. This was a period of rapid expansion of literary output—whether measured in terms of titles or copies—and also a period in which abundance was often explicitly theorized and debated by authors, publishers, and booksellers. It thus offers a particularly striking example of a time and cultural context in which abundance mattered. I argue that publishers, booksellers, and readers in this period developed strategies and institutions to manage the problem of abundance, and that these strategies and institutions affected the structure of the larger literary market.

I focus on three institutions for managing abundance: bestseller lists, the "brand-name" author, and book reviews. These institutions are useful sites for investigation because they were developed and elaborated by actors with a variety of social locations in the literary market. Brand-name authors manage abundance on behalf of publishers, by providing predictable sales that reduce the uncertainty inherent in publishing. Bestseller lists and book reviews, by contrast, are produced by periodicals that are generally independent of publishers; their primary motivation is to provide a service to readers, who read their publications in order to receive news or commentary about the state of the book market. All three institutions interface with readers, by reducing the overwhelming amount of information about new books on the market into some small and digestible sample. By focusing on these three institutions, I am able to consider abundance from a variety of perspectives. I give greater detail on my arguments regarding each of these institutions in the chapter summaries at the end of this introduction.

In the following two sections, I provide additional detail on the concept of “abundance” as theorized in this dissertation, as seen from the two angles to which I alluded above: an epistemological or literary-theoretical angle; and a practical or readerly angle. The first of these two sections summarizes abundance as it has been mobilized rhetorically in computational literary studies over the last two decades. In this scholarly literature, abundance presents a problem for literary historians, one which motivates a turn towards quantitative methods. As I show, many of the concerns expressed by computational literary historians were equally present in book history, a field that was interested in largely distinct historical questions but which found itself confronting many of the same epistemological problems. In the second section, I give a brief introduction to how abundance was theorized in the publishing industry in the historical context covered by this dissertation: the United States between 1931 and 2009. In this period, publishers often expressed anxiety about overproduction, concerned that the increased competition for readers caused by rises in title output would threaten the financial well-being of the industry. I conclude the introduction with a summary of the chapters of the dissertation and a discussion of how they relate to my broader argument about the “problem of abundance.”

Since the periodization used in this dissertation is idiosyncratic, before continuing, it is worth addressing the start and endpoints that I chose for the historical period that I cover. Why

1931 and 2009? There is always something arbitrary about selecting a start and end date for histories that span decades, and this is no less true for this dissertation than for any other project. But the dates I have chosen are not entirely random. They reflect boundaries derived both from the particular sources I consult and the general history I consider. The start date represents the earliest year featured in any of the datasets considered in the dissertation: it is the first year in which the *New York Times* published a bestseller list, though the list would not be expanded to cover nationwide sales until 1945. But a start in the 1930s is also relevant because of changes in the publishing industry. The paperback revolution in American publishing arguably began in 1939, when Pocket Books was founded and began producing cheap, pocket-sized paperbacks for the mass market. This innovation, quickly imitated by competitors like Avon and Dell, vastly increased the scale of book production in the United States.

On the other end of my periodization, my chosen end date also reflects cut-offs in the data: after 2009, coverage of novels in the HathiTrust Digital Library drops off considerably. Since my first chapter relies on text data and metadata derived from HathiTrust, I limit my analysis in that chapter to bestsellers published through 2009. But there are also historical reasons for ending the analysis in the first decade of the twenty-first century. In this decade, the Internet revolutionized the publishing industry at a scale not seen since the invention of the mass market paperback. eBooks created a format even cheaper to mass produce than paperbacks, giving new life—and volume—to highly specialized books and self-published books. Meanwhile, online retailers like Amazon changed both the balance of power and the economic calculus of publishing. Whereas even the largest bookstores can only keep a few tens of thousands of unique books on hand at a time, Amazon's system of distributed warehouses allow them to sell almost every book in print, as well as a huge number of used books that are out of print. Though abundance is still with us in the twenty-first century, its shape looks so different that understanding it would require another study.

0.1. The Conceptual Problem: Canons as Unrepresentative Samples

The idea that literary historians are faced with an unmanageably-large archive of historical literary works has been a common refrain in computational literary studies over the last two decades. In the following section, I review some examples of this rhetoric. In some cases, scholars appealed to the size of this archive as the *raison d'être* for adopting quantitative methods in literary studies. Early versions of this rhetoric appear in the work of Franco Moretti, whose early works on “distant reading” set the terms of much of the subsequent debate. In a frequently-cited piece, Moretti observes that the canon of all those nineteenth-century British novels typically read, taught, and analyzed by scholars is only a small fraction of the archive of all novels published in the century: no more than 200 novels, about 0.5 percent of the total.[6] The remaining twenty thousand or so novels comprise literary history’s “great unread,” a term he borrows from Margaret Cohen. But rather than propose a new, alternative canon, Moretti proposes that literary history adopt methods that allow it to “come up with a new sense of the literary field as a whole.” Such methods are diverse, and include “sampling; statistics; work with series, titles, concordances; incipits,” or even genealogical tree diagrams.[7]

Although computational text analysis does not appear in this piece, similar appeals to the scale of this “great unread” would crop up in debates about computational methods more than a

6. Franco Moretti, “The Slaughterhouse of Literature,” *MLQ: Modern Language Quarterly* 61, no. 1 (2000): 207-227.
7. Moretti, 207-209.

decade later. One notable example appears in the work of Mathew Jockers, who proposed computationally-driven "macroanalysis" (a name derived by analogy to "macroeconomics") as a method equipped to provide the "sense of the literary field as a whole" wished for by Moretti. [8] Jockers contrasts computational methods with the "anecdotal" style of close reading, which focused on only a small number of authors or texts. For this anecdotal style, Jockers cites Ian Watt as an example, whose influential *Rise of the Novel* narrated a general history of the rise of the novel by means of an analysis of three authors: Defoe, Richardson, and Fielding. Jockers expresses skepticism at the idea that we can give a full account of literary history even in a restricted social and historical context on the basis of just a few authors.[9] For one thing, the canonical authors we typically consider are not a "random" sample, and so may not be "representative" of the larger population, in a statistical sense. And though Jockers concedes that authors like Watt had no broader "yardstick" against which to measure their claims, he suggests that the growing availability of large digitized collections means that the "larger literary record can no longer be ignored."[10]

Although Jockers' idea that received literary canons may not be "representative" of the whole is probably true, skeptics of computational methods have argued that this fact does not necessarily lead to the implications that computational scholars want of it. One example of this line of criticism is a piece by Jeremy Rosen responding to Mathew Wilkens. In this piece, Rosen challenges the idea that it makes sense to consider the "great unread" on exactly the same analytical ground as works that have been frequently reread, reprinted, imitated, purchased, or

---

8. Matthew L. Jockers, *Macroanalysis: Digital Methods and Literary History*, Topics in the Digital Humanities (Urbana: University of Illinois Press, 2013).
9. Jockers, 7-9.
10. Jockers, 24-27.

cited. Or, more simply, it doesn't make sense to pay equal attention to works that have not been influential as to those works that have. For Rosen, it is precisely because popular, influential, and canonical works are not representative that they are informative: they are the works that have exerted the greatest pull on literary culture, and so provide disproportionate insight on the functioning of the literary system. As he puts it,

> Timeless classics or bestselling texts, in this understanding, are not to be studied because of their self-evident greatness, their transcendent immanent qualities, but because for better or worse they have achieved a position of cultural centrality, been widely circulated and held up as models.[11]

To consider all texts equally, even those that were rarely read, would be to portray the literary market as having a kind of egalitarian structure that it simply does not have.

Importantly, the positions of Rosen and Wilkens are not necessarily in tension. Rather, I see them as emphasizing different parts of a larger literary-historical process. Of course, Rosen is correct to note that it would be absurd to give all texts an equal weight when considering their social impact and influence. Literary history is extremely lopsided: some books have large circulation or frequent reprintings, whereas some are only ever printed and read in small quantities. But one still might want to study the process by which this happens, the cumulative, selective process that causes some texts to widely circulate or get imitated and some texts to remain marginal. In order to do this, one would need to consider the "losers" as a well as the "winners," which involves a big-picture perspective closer to that of Wilkens. In other words, Wilkens' perspective focuses on texts at the moment of their creation or first printing, whereas Rosen looks at texts in retrospective, from the point of view of how they were actually received

---

11. Jeremy Rosen, "Combining Close and Distant, or, the Utility of Genre Analysis: A Response to Matthew Wilkens's "Contemporary Fiction by the Numbers," *Post45* (December 03, 2011). https://post45.org/2011/12/combining-close-and-distant-or-the-utility-of-genre-analysis-a-response-to-matthew-wilkenss-contemporary-fiction-by-the-numbers/

and circulated. Some distant reading research has tried to consider the process as a whole. In fact, this was the original conceit of Moretti's article, "The Slaughterhouse of Literature," which considered the then-contemporary competitors to now-canonical authors.[12]

In the past decade, however, computational methods have been more closely associated with what might be called a "production side" approach, which considers texts themselves at the moment of first printing, independent of how they were circulated or later reprinted. In a recent critique of this approach, Katherine Bode argues that it is limited by way that it models texts: typically, as individual entities limited in temporal and geographical scope to their year and place of first publication. Such a perspective, which imagines that texts are "singular and stable," necessarily requires neglecting a great deal of the historical context of texts as physical entities produced and circulated in a particular time and place.[13] Additionally, this perspective often neglects all forms of evidence—quantitative or not—besides those derived from textual or linguistic signals. The result is a "dematerialized and depopulated understanding of literature" that, insofar as it does not consider contextual evidence, is much closer to close reading than its proponents portray it to be.[14] For Bode, a better approach would incorporate insights from bibliography, textual criticism, and book history, disciplines in which a single text is conceived of as many distinct events—individual printings—which might vary from each other in important ways. These disciplines also consider contextual historical evidence as a matter of course, rather than limiting their analyses to textual evidence alone.

---

12. See Moretti, 2000.

13. Katherine Bode, "The Equivalence of 'Close' and 'Distant' Reading; or, Toward a New Object for Data-Rich Literary History," *Modern Language Quarterly* 78, Vol 1 (2017): 77-106.

14. Bode, 92.

I think Bode is correct to see limitations in the fact that computational literary studies has so far been mostly confined to textual evidence alone. In particular, the original ambition of distant reading—to study the selection process by which some texts get circulated and some die out—necessarily requires making reference to non-textual data. At minimum, it requires data on how often individual texts were circulated, purchased, or reprinted. In this regard, computational literary studies can learn much from work that has already been done in book history. That field's foundational interest in reconstructing books' historical existence as actual, physical objects lent a particular urgency to the question of whether or not contemporary canons were representative of the past, and considering the data—often quantitative—that would allow scholars to evaluate that representativeness. This attitude is summarized nicely in the work of William St. Clair, whose historical study *The Reading Nation in the Romantic Period* provides a model of book history that is attentive to quantitative data, such as book prices or numbers on circulation or print rates.[15] Like many of the computational literary scholars surveyed above, St. Clair expresses dissatisfaction with literary history that presents the printed writings of the past as a "parade of great names described from a commentator's box set high above the marching column."[16] For St. Clair, this mode of historicizing takes those texts "judged to be 'canonical' in a wide sense" as emblematic examples "believed to catch the essence, or some of the essence, of the historical situation from which they emanated."[17] But St. Clair is also dissatisfied with what he called the "parliament" of texts approach, in which the printed texts of a particular historical context are presented as "debating and negotiating with one another in a kind of open parliament

15. William St. Clair, *The Reading Nation in the Romantic Period* (Cambridge University Press, 2004).
16. St. Clair, 2.
17. St. Clair, 2.

with all the members participating and listening."[18] Both the "parade" and "parliament" modes require that the "historian makes his or her selection of texts to be included."[19] For St. Clair, this methodology is inadequate for understanding the history of printed culture as a "complex developing system" composed of "independent but interacting agents."[20] Historicizing reading as a complex system requires supplementing textual evidence with "empirical research, quantification, consolidation, and scrutiny of primary information."[21]

For St. Clair, quantification is immanent to his particular object of study. It is particularly warranted because of the concrete qualities of historical book markets, rather than equally necessary for understanding textual culture in all times and all places. His conception of print culture as a complex system presupposes a world in which a large number of texts are produced, circulated, and read. And for St. Clair, such a world did not begin to emerge until after the widespread adoption of print, and did not kick off in earnest until the British eighteenth century. As he notes, this was a century in which not only was "more reading occurring… than ever before," but there was also a "change in reading practice."[22] Readers increasingly abandoned the ancient practice of "intensive reading," or reading a few texts repeatedly (such as the Bible), for the more modern practice of "extensive reading," or of reading a large number of different texts a single time.[23] This change was helped along by economic and technological changes. During the manuscript age, most "additions to the stock of books" were further copies of existing texts; the age of print, by contrast, was characterized by the rapid "production of more texts rather than

18. St. Clair, 3.
19. St. Clair, 3.
20. St. Clair, 9.
21. St. Clair, 10.
22. St. Clair, 11.
23. St. Clair, 11.

of more copies of existing texts."[24] And it was only after the expansion of the private market for newly printed works in the eighteenth century that it becomes possible to speak meaningfully of a "reading public."

It is notable that, before computational literary studies, book history was really the only discipline in English departments that used quantitative methods. Books as physical objects are intuitively quantifiable: it is completely ordinary to count pages, copies, sales, and readers, whereas counting words can feel strange and unintuitive to those without experience in quantitative text analysis. In a methodological survey on the uses of quantification in book history, Alexis Weedon notes that quantifying books is a natural goal for anybody in the business of producing them: "Because text production — in the past and now — frequently aimed at multiplying and spreading its product as much as possible, and because those texts commonly became subject to markets and market forces, historical records of books and the book trade sometimes take the form of lists of quantities."[25] Many such lists of quantities were compiled by actors in the book industry, such as publishers, industry analysts, and booksellers, rather than scholars. These include statistics on total titles published compiled by trade journals like *The Bookseller* or *Publishers Weekly*; or records of print runs contained in publisher archives.

Of course, such sources are not without problems. As Weedon puts it, due to "[p]rimitive administration, bookkeeping, and reporting procedures," these records are sometimes more "pseudo-statistics" than they are reliable descriptions of the past.[26] Nevertheless, their mere existence says something important about the history of print culture. Publishers and booksellers

24. St. Clair, 22.

25. Alexis Weedon, "The Uses of Quantification," in *A Companion to the History of the Book, 2nd Edition*, ed. Simon Eliot and Jonathan Rose (Wiley-Blackwell, 2019): 31.

26. Weedon, 38.

were counting books long before literary scholars began crunching the numbers on literary trends. They did so because the "problem of abundance" is as old as the print market itself. Already by the eighteenth century, it was the case that more unique books were in circulation than any individual reader would read in their lifetime. Additionally, these books are not particularly substitutable with each other — as St. Clair puts it, texts and books are "highly differentiated and only mutually substitutable within…wide limits."[27] Readers have many different titles to choose from and one title cannot be easily substituted for a different title. In such a context, publishers and booksellers, whose livelihoods depend on their ability to reliably make sure that the number of books printed closely resembles the number of books sold, have often been very concerned with counting books. Publishing professionals, then, turn to quantification for much the same reason as book historians and as computational literary scholars: to find better ways of managing the problem of abundance. Quantitative or statistical methods are appealing for those in publishing because they provide ways of quickly summarizing trends that exceed the representational capacities of human minds.

For most of the history of print, however, publishing professionals have only had access to very rudimentary quantitative and statistical methods. They were able to count books and estimate returns on investment, but they did not have access to statistical software, let alone to the machine learning methods used in modern computational text analysis. As such, they also had to rely on non- or semi-quantitative strategies for managing abundance. In this dissertation, I argue that understanding these strategies opens up new directions for collaboration between computational literary studies and book and publishing history. But in order to understand those

27. St. Clair, 31.

strategies, we must first consider how publishers themselves thought about the problem of abundance.

0.2. The Historical Problem: Literature in a Crowded Field

Understanding how the problem of abundance structures the literary market requires understanding what is unusual about the publishing industry. Many of the ways that publishing professionals have come to manage abundance stem from the particular social and material constraints associated with printing and selling books, which are a very unusual class of commodity. Books are economically peculiar because of their heterogeneity: there are many different titles for sale at any one time, and every single one of them is distinct. When consumers walk into the beauty aisle of a supermarket, they might be faced with a few dozen competing brands of shampoo. By contrast, if they walk into a bookstore, they might be faced with several thousands of unique novels, even within a specific genre—to say nothing of the many thousands of other novels in that genre they could find if they searched through specialty bookstores, used bookstores, and libraries. And these many thousands of distinct individual books are not especially substitutable for each other. If a supermarket doesn't have your favorite brand of shampoo, a generic brand will often work as an acceptable substitute, depending on how particular you are about your hair. But if the bookstore doesn't have the particular book you are looking for, you can't easily substitute an alternative—only the novels authored by J.K. Rowling can tell you how the *Harry Potter* series will end. This quality of books was nicely summarized by industry analyst Leonard Shatzkin in a 1982 book about the problems ailing the industry: "[n]o other consumer industry produces 20,000 different, relatively low-priced products each year, each with its own personality, requiring individual recognition in the market."[28] Although product differentiation is a spectrum, books are on the far end of this spectrum. For this reason, publishers and other actors in the book market have historically had to come up with unique strategies for producing and selling them profitably.

From the publisher's perspective, the heterogeneity and abundance of books presents a vexing problem, because it makes it very difficult to predict which books will actually sell. Any business that wishes to remain in business for very long must make sure that, in general, it sells enough of its products to finance its continued operation into the future. And although all capitalist production involves this element of risk—the shampoo manufacturer also does not know, for sure, whether their products will sell—this risk is particularly acute in publishing. As literary sociologist Robert Escarpit put it in a 1966 study of world publishing commissioned by UNESCO,

> …publication involves considerable risks. No-one publishing a book can foresee exactly how much attention potential readers will give it. He may make guesses or even forecasts, but he cannot outline in advance the path to be followed by a book which has been put into circulation, he cannot determine the stages and limits of its distribution; in a word, he cannot establish a programme for it.[29]

Book sales present considerable uncertainty. As Escarpit notes elsewhere in that same book, there is a huge variation in the average number of copies sold of different titles.[30] A few titles sell hundreds of thousands or even millions of copies, while most titles sell only a few hundred or, if they are lucky, a few thousand. This level of variation is so extreme that it is almost meaningless to speak of an "average" number of copies of a particular title. Unlike functional

28. Leonard Shatzkin, *In Cold Type: Overcoming the Book Crisis* (New York: Houghton Mifflin, 1982): 3.

29. Robert Escarpit, *The Book Revolution* (London: George G. Harrap and Co; Paris: United Nations Educational, Scientific, and Cultural Organization, 1966): 121.

30. Escarpit, 78.

books like dictionaries and textbooks, which have relatively fixed demand, a "literary book" represents a "pure 'adventure'" for the publisher. Its "sales figures have very little to do with printing figures." In general, a "young author's first novel" would be fortunate to find even "a few hundred readers."[31]

Given these conditions, publishing has historically been a business characterized in large part by chance, or, at the very least, by the perception that things are run by chance. As Shatzkin puts it, "the experience of every publisher convinces him that this is a 'gambling' business." This sense of uncertainty, the feeling that the market fundamentally cannot be predicted, is the shibboleth of the publishing professional. It's the structure of feeling that animates the world of those who make their living from selling books. Much of this sense of uncertainty stems from the sheer magnitude of the enterprise, as Shatzkin describes:

> Tens of thousands of titles published into an unstructured distribution system, millions of unpredictable negotiations between sales reps and booksellers, the happenstance of a book getting into a particular store, together with unforeseeable outside influences—reviews, attention by celebrities, quirks of public taste—lead to results in which chance is the dominant factor.[32]

With so many contingent moving parts, the literary marketplace can feel almost sublime in its size and complexity, in the sense that it is beyond the cognitive capacities of any one individual to create a mental representation of its totality.

Unsurprisingly, then, publishers have often experimented with strategies and institutions for managing the uncertainty of the literary market. As Escarpit observes, publishers have sometimes been able to convince readers to purchase books by subscription, which guarantees the size of the market "by the fact that the volume is paid for in advance."[33] Similarly, book

---

31. Escarpit, 78.
32. Shatzkin, 14.
33. Escarpit, 121.

clubs reduce uncertainty by asking readers to commit to selecting only from a smaller, restricted set of books. But one of the most familiar strategies for reducing the uncertainty of a heterogeneous literary market is to simply make it less heterogeneous, by publishing books that are similar to books that have sold well in the past. One might do this, for example, by publishing books within a targeted and specific genre. Many of the classic theorizations of "genre fiction" literary genres like detective, romance, and science fiction describe them as attempts to rationalize the production and consumption of novels. Janice Radway, whose 1984 *Reading the Romance* is still a standard reference on the subject, describes genre fiction in these terms. As Radway notes, popular genres sometimes began as explicit attempts by publishers to mimic the success of past bestsellers, as when, in the nineteenth century, the Beadle brothers began publishing domestic novels and Westerns after the success of Ann Stephens' novel *Maleska*.[34] This strategy was further developed and standardized by mass market paperback publishers in the middle of the twentieth century, which often specialized in one or two highly-specific popular genres. An early example of this was American Mercury Books. Although Mercury began as a general mass market publisher, by 1940, its editors had decided to focus on mysteries "in the interest of establishing better control over their market."[35] Similarly, the famous romance publishers Harlequin Enterprises and Silhouette Books based their publishing practice on the idea that it was possible to achieve "highly predictable" sales, by publishing similar types of books to a stable demographic of readers with relatively-established tastes.[36]

---

34. Janice Radway, *Reading the Romance: Women, Patriarchy, and Popular Literature* (Chapel Hill: University of North Carolina Press, 1984): 23.
35. Radway, 26.
36. Radway, 23.

Another way of managing the uncertainty of abundance is to publish work from authors who have been successful in the past. Although books in general are highly distinct, books by the same author give readers a sense of familiarity, somewhat similar to brand-name recognition in other industries. Indeed, the most famous, regularly-bestselling novelists are sometimes called "brand-name authors" for this reason. Citing an interview with somebody who works in publishing, sociologist John Thompson notes that the selling power of these authors make them easy to work with for the marketing department:

> The easiest books to deal with are those by brand-name authors, the John Grishams and Michael Crichtons and Stephen Kings — these books are 'pure numbers, we put out 60,000 of this here, we'e going to put out 35,000 there and that's it, you don't have to talk about it'... The author's name, the cover and the historical record of the sales of the author's previous books in different channels are all you need to know.[37]

These kinds of authors have fan followings and celebrity appeal that make them reliable options for publishers. Though their sales aren't guaranteed—even a celebrity novelist can write a flop—they at least reduce some of the uncertainty in the expected number of books that they will sell.

Nevertheless, it is easy to overstate the degree to which even the biggest corporate publishers have overcome the uncertainty inherent in publishing. Genre fiction and books by brand-name authors, while important, are not the only novels that get sold. Some novels attempt to modify a genre for readers outside of its typical audience, or simply don't fit neatly into any existing generic mold. And publishers can't exclusively publish novels by established authors, who might go out of fashion, or die of old age—they still need to take risks on new and unknown writers from time to time. For authors who don't fit into these rationalized, predictable

37. John Thompson, "Big Books," in *Merchants of Culture: The Publishing Business in the Twenty-First Century*, 2nd ed. (New York: Penguin Group, 2012).

paradigms, publishers have typically had to rely on decidedly non-rationalized methods for choosing what to publish, and in what quantities.

Consider, for example, the "slush pile," that evocative industry term for the pile (or, in the twenty-first century, electronic data) of unsolicited manuscripts received by a publisher or, more often, an agent. Although publishers and agents often solicit manuscripts directly, this is not always possible, and throughout the twentieth century, many novelists who later became successful got their start by sending an unsolicited manuscript. Such manuscripts present a challenge for the publisher or agent, who may have nothing but the text itself on which to base their decision. In such a context, they generally have no choice but to simply read them all. The challenge of this undertaking is implicit in the nickname "slush pile," which evokes a big mess of gross, undifferentiated matter. There are of course many ways of simplifying the process, or of dividing up the labor. Reading a few paragraphs may be enough to exclude a work based on problems with style alone. And established editors often delegate the task of wading through the slush to those on the bottom of the labor hierarchy, such as junior editors or an agent's assistants. This work can be a slog, and mind-numbingly boring. As one freelance editor put it in an article for *Publishers Weekly* in 1986, "No one who has not sat down, day after day, and shuffled through the stuff that comes in unsolicited can imagine how unspeakably godawful most of it is."[38] Sorting through the slush is hard work, and time-consuming, so editors may prefer manuscripts directly sent to them by an agent, who have already read and vetted the text.

The slush pile is managed in the only way that anybody has figured out how to manage it: by reading it. I think it is worth dwelling for a moment on the implications of this point. In order for a book to even be considered for publication, somebody, somewhere, has to read it. Though

38. Ann Finlayson, "Why We Call It Slush," *Publishers Weekly*, May 16, 1986: 84.

this may seem obvious, it provides an interesting counterpoint to some of the recent rhetoric about computational methods in digital literary studies. Despite protestations that it is not possible to "read it all," it is in fact true that everything in print has been read by someone.[39] Of course, it is not all read by one person—the task of "reading it all" is distributed across all the different actors in the publishing and bookselling industries, whose aggregate behavior we call the literary market. In the case of unsolicited manuscripts, this labor is divided up (unevenly) between editors, editors' assistants, agents, agent's assistants, and freelance editors. Depending on the experience, taste, and goals of the people who first read that manuscript, it might ultimately get passed along to other social actors, before finally moving to the process of publication. When alternative criteria like an author's past success are not available for deciding whether to publish something, some version of this process of collective evaluation is what decides a book's fate.[40]

Even after publication, collective evaluation remains an important part of how the literary marketplace as a whole manages abundance. From the perspective of actual readers, it is probably the most important method for managing abundance: people read books which have been recommended to them by other people whose opinions they respect or whose taste has overlapped with their own in the past. They might hear about them from a friend, get a recommendation from a librarian or bookstore clerk, or, today, see a reference to it on their social media feed. For this reason, publishers put a lot of emphasis on strategies that help

39. This may not be true of self-published eBooks, which can feasibly be "published" without ever having been read by anyone other than the author. For this and other reasons detailed below, the historical scope of this dissertation ends in 2009, around the time that eBook sales numbers were becoming competitive with print sales numbers.

40. Thompson suggests that this process—which he calls the "web of collective belief"—is ultimately the most important determinant in whether a book gets published. See Thompson, Chapter 5, "Big Books."

kickstart collective or word-of-mouth interest in a title. In an interview with *Publishers Weekly* in 1982, Ann McKeown, publicity director of Avon, says that "word-of-mouth is what makes books sell."[41] Word-of-mouth is notoriously unpredictable and hard to control, but there are different methods that publishers employ in order to try to spur its organic growth. In the case of new titles, McKeown suggests that book reviews are important, because they are "the best way… to get word-of-mouth started about a new original."[42] Book reviews are a very personal and labor-intensive method of managing the problem of abundance. They divide the labor of "reading it all" amongst a large number of paid, professional reviewers, who take the time to read—or at least skim—new books and provide their opinions about them. Although it's not necessarily the case that book reviews affect sales in any mechanistic way, at the very least, a book review provides publicity for new books. Given that so many books are published each year that readers won't even hear about most of them, book reviews may draw attention to books that readers might otherwise be unaware of.

As I hope this discussion has made clear, the problem of abundance is not just a methodological difficulty that we have to deal with as literary scholars. It is, rather, a fundamental aspect of modern literary markets, and it structures the behavior of everybody who participates in those markets: publishers, booksellers, authors, and readers alike. If there is one thing that characterizes literary culture, it's that there is anybody who wants to get their proverbial voice heard is faced with an enormous number of competitors. This has been true since at least the eighteenth century, and the competition has only gotten fiercer in the centuries since.

---

41. Carolyn T. Anthony, "Are Reviewers Fair To Trade Paperbacks?" *Publishers Weekly*, October 15, 1982: 34.

42. Anthony, 34.

In this dissertation, I confront the problem of abundance both as a methodological challenge and as a matter of historical fact. Doing so requires not just managing abundance but also theorizing abundance's management. As such, my methods throughout the dissertation are diverse. They comprise quantitative methods applied to bestseller and book review data, historical analysis of the *Publishers Weekly* archive, and close readings of particular novels. In each chapter, I consider a different strategy or institution which manages abundance: bestseller lists, brand-name authors, and book reviews. I describe the topic and intervention of each chapter in greater detail below.

0.3. Chapter Summaries

The chapters of the dissertation are each organized around a single strategy or institution that manages abundance. The first chapter, titled "The *New York Times* Bestseller List and the Rise of the 'Mass Market Hardcover,'" considers the history of what is probably the most famous weekly bestseller list: the hardcover fiction list of the *New York Times*. Bestseller lists are a very simple method for managing abundance; they work by simply publicizing information on titles that are currently popular. In this chapter, I historicize the bestseller list, describe how its contents have changed over time, and consider its significance for the larger terrain of fiction publishing. I argue that, in the twentieth century, the *New York Times* bestseller list became increasingly important for decision-making in the publishing industry. Chain bookstores often provided large discounts based on whether a title had been featured on the list, and publishers gave favorable treatment to previous bestselling authors, who were seen as more likely than other authors to become bestsellers again. After giving some historical context, I turn to the dataset used for the chapter, which matches *New York Times* hardcover bestsellers with titles in

HathiTrust. I show that traditionally middlebrow genres like historical fiction and domestic novels were initially popular in the 1940s and 1950s, but fell out of prominence by the 1980s. They were replaced by the many genres of "genre fiction," especially crime fiction and thrillers. I consider how this trend might reflect broader changes in publishing, such as the fact that these "mass market" genres were increasingly moving into hardcover in this period.

The second chapter is titled "Literary Brand Names and the Political Economy of Authorship." In this chapter, I historicize the institution of brand-name authorship, or authors whose name comes to stand in for a larger literary franchise, independently of the actual role that the author personally played in the production of texts in that franchise. This arrangement includes authors like V.C. Andrews or Tom Clancy, whose series were finished by ghostwriters after their deaths, or authors like James Patterson, who co-write novels with less-established collaborators, lending their brand-recognition to the project. As I show, arrangements of this type became increasingly commonplace among bestsellers after about 1980. I argue that brand-name authors are valuable to publishers because their name recognition gives a level of consumer recognition and repeat consumption that has historically been difficult to achieve in publishing. Then, I give some historical background for this phenomenon, focusing on the history of how copyright law defines authorship. As I explain, the "author" in copyright terms is closer to the owner or manager of a text than it is to its literal writer, and although authors are often also writers, legally they needn't be. Finally, I consider three different emblematic cases of publishing arrangements that make use of a division between writing and authorship. I consider the case of "book packagers," or independent editorial firms that hire writers and assign authorship to a corporate organization. I perform a close reading of *The Agent*, commissioned by book packager Bill Adler and ghostwritten by David R. Slavitt, which explicitly thematizes how authorship changes when the author is a manager rather than a writer. Then, I discuss the case of novels ghostwritten for larger media franchises, such as the many *Star Wars* novelizations that were popular in the 1970s and 1980s. Finally, I discuss the case of author V.C. Andrews. After her death, Andrews' estate—in collaboration with her publisher—commissioned the writer Andrew Neiderman to continue writing in her name. Although Neiderman was originally hired only to finish novels that were either drafted or outlined by V.C. Andrews herself, the continued success of the Andrews brand prompted the estate and the publisher to commission him to ghostwrite entirely new novels as well. The V.C. Andrews case is the purest case of author as brand-name, as the authorship of the later books was assigned to a corporate legal entity: the V.C. Andrews trust. I conclude the chapter with a brief discussion of new directions in corporate authorship.

In the third and final chapter, I consider one of the oldest institutions for managing abundance: book reviews. Titled "Dimensions of Reception: A Cluster Analysis of the *Book Review Index*, 1965-2000," this chapter considers a dataset on over three million book reviews that were indexed by the *Book Review Index,* a reference publication commonly held by research libraries. Book reviews represent a very personal and labor-intensive method of managing abundance, delegating the task of sifting through new books to paid, professional reviewers. In this chapter, I argue that the structure of the book review market shows an important side-effect of abundance: that the literary market is segmented into a variety of distinct subcultures and reading communities, reflecting distinct standards of taste and hierarchies of value. Since it is impossible to read everything, people generally read books that are also of interest to people whose opinions they respect, or whose taste has overlapped with their own in the past. I demonstrate this phenomenon by making use of a suite of clustering algorithms commonly used in machine learning. As I show, these clustering algorithms automatically divide the data into groups of journals, or clusters, with similar taste, based on whether they often reviewed the same books and authors in the period under consideration. These clusters reflect intuitive social divisions in the literary market, including divisions based on profession, genre, geography, market size, and even race and gender. After identifying prominent clusters in the data, I inspect the most frequently reviewed authors within each cluster and compare how different authors performed in different clusters. Finally, I close the chapter with a discussion of how this method of multidimensional data analysis differs from existing work on reception and prestige in quantitative literary studies. Rather than modeling “prestige” as a single-dimensional value that varies in magnitude but not in kind, I argue that it makes more sense to model prestige as a many-dimensional value, with dimensions that may or may not be correlated. In other words, there are many “prestiges,” reflecting the standards and tastes of particular literary communities.

Chapter 1. The *New York Times* Bestseller List and the Rise of the "Mass Market Hardcover"

In this chapter, I provide an analysis of one of the conceptually-simplest institutions for managing abundance: published lists of bestsellers. I focus in particular on the hardcover fiction list of the *New York Times*, which is the among the longest-lasting and most influential of the many bestseller lists published by a mainstream periodical. Bestseller lists manage abundance in a very straightforward way: by surfacing information to readers on books that are currently popular. When a reader browses the weekly bestsellers published by a newspaper like the *New York Times*, they are faced with a small and easily-digestible sample of books that are currently on the market. Today, that number is typically fifteen, though the *Times* has varied the number of weekly bestsellers that it publicizes over the life of its list. For the reader curious about the latest literary trends, that sample has the added benefit of being indirectly curated, in the sense that a title can only reach the list by means of the aggregate assent of the many thousands of people spread across the country who choose to exchange their money for a copy. Although we may question the degree to which purchasing a book necessarily represents an endorsement—and, as we will see, even measuring aggregate purchases is not always a straightforward matter—it nevertheless seems unproblematic to observe that bestseller lists provide a sort of filtration or curatorial service for prospective book-buyers. In fact, it is for this reason that bestseller lists have commonly drawn scorn from commentators on the literary market, who have often accused it of homogenizing the very market that it claims to represent. As novelist Pearl S. Buck, herself the author of multiple bestsellers, put it in a speech to the American Booksellers' Association in 1938: "The best-seller list is not a thermometer. It is an iron mold clamped upon the public mind."[1] The bestseller list, in this account, is like a self-fulfilling prophecy, which alters the world in the act of describing it.

I make two interventions in this chapter. First, I analyze bestseller lists in the twentieth-century United States as they emerged, grew in popularity, and standardized in form and methodology, focusing on the case of the *New York Times* list. Although many city newspapers published lists of local bestsellers in the first half of the twentieth century, it was not until the *New York Times* expanded its coverage to nationwide sales in 1945 that it really becomes possible to speak of a national bestseller list in a mainstream publication (as opposed to an industry publication like *Publishers Weekly*, which started its bestseller list in 1912). The *Times* list quickly established itself as *the* national bestseller list, to such a degree that it became the standard reference point for whether a title could be described as a bestseller. Secondly, I analyze the contents of the *New York Times* bestseller list itself, asking what we can learn about American fiction in the latter half of the twentieth century by considering how the "typical" novel on the *Times* list changed over the period. To this end, I compile a dataset on every novel that has ever appeared on the *Times* hardcover fiction list, paired with matching volumes in the HathiTrust Digital Library. The dataset begins in 1946, the first full year that the *Times* publishes a national bestseller list, and ends in 2009, when the HathiTrust dataset I use ends. My major finding in analyzing this data is that the list experienced a genre inversion between about 1960 and 1980: whereas historical epics and domestic fiction were the most common genres of bestseller before 1960, after 1980, they were replaced by detective and mystery novels and the many subgenres of what are called "thrillers." As I explain, one explanation for this inversion has to do with the evolution of book formats in the publishing industry: since this is specifically

1. "Pearl Buck Scores Best-Seller Lists," *The New York Times*, April 28, 1938: 21.

a hardcover bestseller list, this evolution reflects the fact that genres traditionally limited to mass market paperback format increasingly received hardcover printings in this period.

If we think about the "problem of abundance" in terms of the stakes that I laid out in the introduction, then we can see that bestseller lists provide an entry point for theorizing abundance from both an epistemological and a historical angle. Epistemologically, they are interesting as a source of summary information on popular books: they give book historians and quantitative literary scholars a tractable sample of popular books, which might tell us something interesting about the vagaries of popular taste over some span of history. This is how they have traditionally been used by scholars, especially in quantitative literary studies, as I will show. On the other hand, they are also important as historical phenomena: the fact that newspapers found that there was a reliable audience for information on bestsellers in the first place is an important index of how large and diverse the literary market had become by the latter half of the twentieth century. But theorizing their historical importance also shows their limitations as sources of information. As I discuss in more detail in the body of the chapter, bestseller lists are not a neutral source, but rather reflect editorial decisions about how and what to count. I focus in particular on the format distinction, noting that the fact that this is a hardcover list means that the most popular novelists of the immediate postwar period are excluded, since these novelists made the vast majority of their sales in paperback. This shift is what is referred to by the second half of this chapter's title: the genre inversion that I observe in my quantitative analysis of the list is probably due, in part, to changing industry norms about hardcover and paperback publication, rather than merely changing readerly tastes.

The remainder of the chapter is laid out as follows. In the next section, I provide some historical context on bestseller lists in general, as well as the *New York Times* bestseller list in particular, in the twentieth century. Following that discussion, I provide an introduction to the data and methodology consulted in the quantitative portion of the chapter. Next, I report the central finding of my analysis of the data: the inversion of genres that the *Times* list experiences between about 1960 and 1980. I provide visualizations that capture this inversion from different angles. In the following section, I discuss some of the reasons that we might suspect that this inversion in part reflects specific changes in hardcover publishing, rather than simply reflecting a general shift in the literary tastes or purchasing habits of readers. These reasons include the fact that the top-selling crime novelists prior to 1960 would not have been counted, due to their disproportionate paperback sales. I also note that, in this period and especially in the 1980s, publishers frequently noted that hardcover norms were rapidly changing. I close with a discussion of how studying the publishing industry can inform future work in quantitative literary studies. I suggest that, since scholars using quantitative methods are often interested in studying literary-historical trends that are large in geographic or temporal scope, their findings will likely be influenced by institutional changes in how books are printed and distributed.

1.1. The *New York Times* in the Universe of Bestseller Lists

Histories of bestseller lists often attribute their invention—or at least popularization—to *The Bookman*, which started publishing lists of bestselling books in 1896. As Alice Payne Hackett, author of a well-known midcentury history of bestsellers, notes, the first lists published by *The Bookman* were "uncorrelated," in the sense that the periodical simply republished lists of bestsellers as they were reported by "a dozen or so" major bookstores throughout the country, as well as a few in Canada. There was no assumption that these bookstores would report the same title and there was no attempt to aggregate them into a single list of overall bestsellers. As they grew in popularity, however, they were expanded to cover more and more stores and, eventually, combined into a single overall list of six bestsellers. From there, Hackett suggests, bestseller lists proliferated in different venues, such that, by the early twentieth century, the term "bestseller" had permanently entered the lexicon.[2]

In the first half of the twentieth century, though it was commonplace to talk about the latest bestsellers or this or that bestseller in particular, there was not a lot of centralized, standardized information on bestsellers. *The Bookman* and later *Publishers Weekly* published national lists, but the lists published by local newspapers were often only local in scope. Though in casual conversation it was and remains commonplace to refer to individual titles as "bestsellers" without inquiring too closely as to what they honorific actually entails, the early history of bestsellers draws attention to why such monikers can be misleading: there is not one single bestseller list, but rather many bestseller lists, reflecting the tastes of the particular sample of readers under consideration and the particular method of counting used to assemble each particular list. As Hackett observes: "Many newspapers publish their local bestseller lists; the trade papers and book review sections print national lists based on countrywide canvass; even individual bookstores publicize their own best sellers."[3] Given that book sales may vary wildly in different regions or even from bookstore to bookstore, the haphazard nature of these competing bestseller lists means that they may report very different titles. As Laura Miller puts it in a general history of bestseller lists, "best-sellers for a Borders are likely to be different from the best-sellers for an independent women's bookstore."[4] Obviously, local and specialized lists

2. Alice P. Hackett, "The 'Best Seller' Lists--and How they Grew," *New York Times,* October 6, 1946: BR 22.

3. Hackett, BR 22.

4. Laura J. Miller, "The Best-Seller List as Marketing Tool and Historical Fiction," *Book History* 3, Volume 1 (January 2000): 291.

will reflect the idiosyncratic tastes of the particular bookstore patrons whose purchasing habits are being surveyed.

In this early period of bestseller lists, commentators both in the publishing industry and in the world of book reporting sometimes commented on the apparently-irrational quality of these lists. This was a source of professional frustration for M. Lincoln Schuster (co-founder of Simon and Schuster), who in a 1932 interview went so far to propose that the publishing industry ought to form an independent trade committee to "codify some systematic procedure" for the preparation of the many bestseller lists that were in circulation at the time. Without some codification, Schuster noted, "book-sellers and book-buyers are bewildered," unable to make informed decisions. Schuster suggested that the then-contemporary norm of having many different unstandardized, competing, and unregulated bestseller lists resulted in many "abuses" and risked making the term "best-seller" "utterly meaningless." With many different lists, including some released by publishers themselves, any book could, in theory, be called a "bestseller" by someone, making it difficult to ascertain which books were truly selling widely and in large numbers. A neutral, standardized, and systematic bestseller list, by contrast, could give a sense of aggregate national trends rather than geographically-local ones.[5]

The actual history of bestseller lists would go in a somewhat different direction than the one proposed by Schuster. There would be no independent trade commission for the production and publication of bestseller lists, though the bestseller lists maintained by *Publishers Weekly* do fulfill something of that function. Instead, bestseller lists remained a private affair, often published by newspapers. In the years immediately following Schuster's interview, however, one paper in particular did make a new effort to expand their bestseller coverage to be truly national

5. "Schuster Assails 'Best Seller' Lists," *New York Times*, July 19, 1932: 15.

in scope, and to adopt a relatively-systematic procedure for collecting data on bestsellers: the *New York Times*.

Although the *New York Times* published its first bestseller list in 1931, before Schuster's interview, the *Times* list at this early moment was restricted in scope to book sales in New York City itself. Later that year, it expanded coverage to eight cities, and by the early 1940s, fourteen cities. But these early lists were still city lists, rather than national aggregates. The *Times* only began publishing its nationwide list in 1942, when it added a new chart that ranked both fiction and nonfiction titles based on the number of distinct cities for which that title was a local bestseller. In other words, it was in part a measure of a title's geographic reach, rather than raw sales numbers, since a title could in theory sell twice as many copies in city A as in city B, but still have the same rank in each city (if, for example, more books overall were sold in city A than city B). It was not until the September 9, 1945 issue that the *Times* officially scrapped city-specific lists for a single national list, which purported to be based on "reports from leading booksellers in twenty-two cities."[6] Presumably, this new list better reflected how individual titles were selling across the nation as a whole.

If the continuously changing methodology of the *Times* list in this period seems arbitrary to modern sensibilities, this is in part because it is easy to forget just how much of a logistical challenge it is to compile sales numbers for book titles across the entire United States. Today, services like the Nielsen Company's BookScan, which automatically tracks sales based on point-of-sale information, make it possible to quickly and reliably acquire data on the sales of individual titles. But no centralized sources existed before computerization. In a predigital world,

6. Miller, "The Bestselling List," 290. The summary of the *Times* methodology in the following paragraphs is based primarily on this piece.

how does one actually find out which books are selling? The only source of information on sales was whatever information was buried in the heads and documents of clerks at thousands of bookstores across the country. So, there was no better method to aggregate this information than to simply ask those bookstore clerks which titles were selling well. And that is more or less what the *Times* did, for most of its history. Each week, the *Times* would survey a sample of booksellers—including both wholesalers and bookstores selling directly to consumers—by sending them a form that asked respondents to report sales for all the titles that the times was tracking. The *Times* would apparently try to keep its sample of stores representative, and it expanded the number of stores that it queried at different moments in its history. In 1977, for example, it expanded its sample from about 250 bookstores to about 1400, and in the 1990s, it had increased to more than 3000.[7]

Likely because the *Times* list was the first mainstream (as opposed to industry) publication to publish regular national bestsellers, it quickly developed a reputation as the bestseller list to watch. According to Laura Miller, it held this honor as early as the 1950s.[8] However, it's true economic power would come once inclusion on the *Times* list began to be used as a decision-making factor by publishers and booksellers. By the 1970s, chain bookstores in mall outlets and superstores were offering promotional discounts specifically to books that made it onto this list, and often would dedicate a special section at the front of the store to the latest *Times* bestsellers. Similarly in the 1970s, publishers sometimes included clauses in authors' contracts that would award the author thousands of additional dollars if the book made it

7. Miller, "The Best-Seller List," 290-291.
8. Miller, "The Best-Seller List," 290.

onto the list. So, the list has a self-fulfilling quality to it: sales of a novel might be increased if it becomes a bestseller, since bookstores would charge less for a book by virtue of that fact.

The reason for the importance of the *Times* list to the industry is not difficult to understand. As the most widely read bestseller list, inclusion on the list means that a book will get much more publicity than it might otherwise. Although the actual causal effect that the list has on sales is difficult to measure, it is worth noting that one of the few economic studies of the *Times* bestseller list found that the most likely channel by which it affects sales is in publicity. The list helps sales by bringing attention to books that might otherwise have gone unnoticed by prospective readers. The study in question, published in 2007, found that inclusion on the *Times* bestseller list had the greatest effect on sales of books for debut authors.[9] In other words, making the list will give you the biggest "bump" if you are somebody who is not already famous. Since everybody already knows that the next Danielle Steel or James Patterson novel will be a bestseller, it is not very remarkable when it turns out, in fact, to be a bestseller. By contrast, when a debut novelist makes the list, everybody takes notice; it's more newsworthy when the first *Fifty Shades of Grey* book becomes a bestseller than when the last *Harry Potter* book does.

Given its importance, it is perhaps unsurprising that the methodology of the *Times* list has frequently come under scrutiny. The *Times* has even been sued by authors alleging that their books were improperly omitted. In 1983, for instance, William Blatty—author of *The Exorcist*—sued the *Times*. His lawsuit alleged that the *Times* had incorrectly excluded his latest novel *Legion*, a sequel to *The Exorcist*, from the bestseller list. According to Blatty's lawyers, *Legion* had sold enough copies to warrant a spot, so its absence was due to negligence or fraud, for

9. Alan T. Sorensen, "Bestseller Lists and Product Variety," *The Journal of Industrial Economics*, Volume LV, No. 4 (December 2007): 715-738.

which Blatty was entitled to compensation. The *Times* countered that the list is not mathematically objective but is instead editorial content protected by the First Amendment. The court ruled in favor of the *New York Times*. Although the methods used by the *Times* are not arbitrary, this court case does serve as a reminder that their definition of bestsellers is not the final say on the matter.[10]

The *Times* list, then, is relevant to the project of this dissertation from two angles at once. On the one hand, it is an attempt to manage abundance at a conceptual level: it gives its readers a "snapshot" of what the literary market looks like on a given week. This is the way that bestseller lists have sometimes been used in quantitative literary studies, as I discuss in greater detail later in the chapter. On the other hand, as this discussion has shown, the *Times* list is also important as a historical symptom of abundance. It doesn't just measure the market, but also intervenes in it, by drawing attention to some works rather than others. For publishers and authors, this intervention can be extremely valuable, in that it is an effective way for a title to stand out from the crowd. Getting included on the list is a valuable kind of publicity. In the following section, I consider what kinds of books have tended to receive this publicity over the whole life of the list.

1.2. Historical Bestsellers, Bestselling Histories

In the previous section, I considered the bestseller list as an historical phenomenon. In the following section, I consider it as an historical source. What kinds of books have tended to become bestsellers, as reported by the *New York Times*? Have the kinds of books that have appeared on the list changed over time? If, as I argued above, bestseller lists are important

10. Miller, "The Best-Seller List," 297-8.

because their contents shape the behavior of publishers, authors, and booksellers, then it follows naturally than one would want to know what kinds of books have historically appeared on the list. Authors who hit the bestseller list gain additional publicity and may improve their bargaining position with respect to publishers when negotiating a contract for future books. As we saw, bookstores in the 1970s and 1980s sometimes gave special discounts for books that hit the *New York Times* bestseller list specifically. The list seems to exert a sort of gravitational pull on the market as a whole, drawing the attention of readers and publishers to the titles that appear on its pages, even if its causal effect on the sales numbers for a specific title remains ambiguous.

To this end, the chapter consults a dataset I compiled that matches titles that made it onto the *New York Times* bestseller list at least once with fiction titles available in the NovelTM Dataset derived from the HathiTrust Digital Library and compiled by Ted Underwood *et al*.[11] This dataset enables me to summarize some of the general trends in the books that appeared on the list over its whole history. The dataset contains entries for each week of the list between 1946, the first full year that the *Times* list is a national list, and 2009, when coverage in the NovelTM dataset ends. I was able to find at least one volume for 2,438 of the ~4,660 works of fiction that made it onto the list for at least one week in that period, or about 52.3% of the total. Coverage for the key years of my analysis, between 1946 and 2009, is shown in **Figure 1.1** below. Coverage is more complete between about 1946 and 1960, when almost all titles have a matching HathiTrust volume, and least complete around 1980 and 1990 and 2005 and 2009. Note that there is variation in the number of unique titles to hit the list over its history. In part,

11. Jordan Pruett, "NYT Hardcover Fiction Bestsellers," *Post45 Data Collective*, 2022. For the NovelTM dataset, see Ted Underwood, Patrick Kimutis, and Jessica Witte, "NovelTM Datasets for English-Language Fiction, 1700-2009," *Journal of Cultural Analytics*, Vol. 5, Issue 2 (May 28, 2020).

this is due to changes in the *Times*' own methodology; for a period of time in the 1960s and early 1970s, the *Times* published only ten bestsellers rather than the usual fifteen. But there is also a genuine spike around the turn of the twentieth century; bestsellers in 2000 spend less time on the list than bestsellers in 1950, giving space for a broader variety of books.

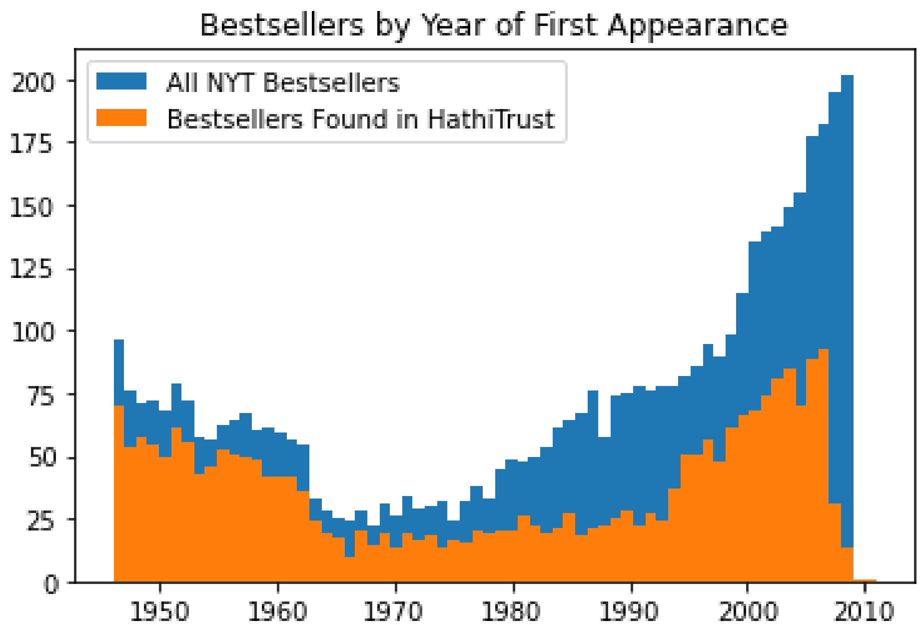

Figure 1.1. Bestsellers in HathiTrust by year of first appearance.

Since more than 2,400 books is already too large a number to be tractable for manual inspection, I adopt methods from computational text analysis designed for summarizing the contents of large corpora. I make use of two methods in particular. First, I use topic modeling by means of latent Dirichlet allocation, an unsupervised method for discovering the latent thematic structure of a collection of documents. Topic modeling represents each document in a corpus as a distribution of topics or themes, defined as words that tend to be used together in the same context. Second, by matching titles with their corresponding entries in the Library of Congress, I chart the rise and fall in prominence of different genres on the bestseller list. In order to test whether the genre labels used by the Library of Congress are meaningful at the level of style, I make use of text classification with logistic regression. I broadly find that the genre categories are coherent at the level of style, particularly the genres which show the most dramatic changes in prominence over the whole life of the list, and to which I therefore devote the most attention. I provide additional details on each method in the body of my discussion below.

The primary finding that I explore in this section is that the genre composition of the *Times* bestseller list experienced an inversion between 1946 and 2009. Before about 1960, the most common genres on the bestseller list were domestic and historical fiction, which roughly correspond to the two great themes of middlebrow realism: social change and family life. Between 1960 and 1980, and especially after 1980, these genres were decisively supplanted by genres new to the bestseller scene: mystery novels and what the Library of Congress calls "suspense" novels, a name that describes thrillers and their many related subgenres ("techno-thrillers," "spy thrillers," etc.). This inversion is visible both in the topics I uncover in the topic modeling portion of the analysis and in the genre labels provided by the Library of Congress.

But in interpreting the meaning of this shift, I am careful not to imply that the *Times* bestseller list is a transparent window in the history of popular taste, or even into the history of book-purchasing trends. As the discussion of its methodology in the previous section implies, the *Times* list is hardly a comprehensive or neutral source, particularly in the decades before computerization. It is possible that some or all of the trends I observe in this dataset are due to changes in the methodology used by the *Times* list, rather than changes in what Americans were actually buying or reading. For instance, since the *Times* list increased the size of its sample multiple times in the period that I study, it is possible that mystery novels and thrillers were more common before 1960, but simply were not being properly counted. In Section 1.3, I consider one way that the methodology of the list likely affects the observed trend, relating to the fact that this list specifically tracks hardcover novels, rather than novels in general.

First, let's turn to the topic model trained on the data. Topic modeling algorithms automatically generate some number of "topics" composed of words that tend to co-occur in the same document. Then, each document is assigned a value for each generated topic reflecting the portion of that document's words that are devoted to that topic. For this topic model, I adapted the method used by Mathew Jockers and David Mimno.[12] Before training the model, I separated individual novels into chunks of 2000 words, limited the corpus vocabulary to the top 20% of words by document frequency, and kept only nouns, which tends to produce a more "thematic" output. Nouns give a sense of what a novel is about, rather than the style in which a novel is narrated.

Although we generally refer to the word clusters produced by LDA as "topics," in the case of fiction they can often be more accurately described as "discourses." Concretely, they are simply words that tend to appear together in the same collection. For example, Topic 3 of 100 from this topic model includes the following words:

12. Matthew Jockers and David Mimno, "Significant Themes in 19th-Century Literature," *Poetics* 41: 6 (December 2013): 750-769.

Topic 3

top words	top novel by topic proportion
prison, death, cell, prisoners, years, guard, prisoner, guards, murder, jail, name, crime, police, day, life, days	*The Fixer* (1966), by Bernard Malamud, about a Jewish man wrongfully imprisoned in Tsarist Russia.

Table 1.1. Top words and sample novel for Topic 3

It's easy enough to see that this topic was generated from scenes of prison life and discussions of criminal justice. But although many of the words most associated with this topic are explicitly about prison, some of them instead seem to reflect the kind of discourse that generally accompanies a fictional prison scene. *Day* and *days*, for instance, may reflect the fact that imprisoned characters often narrate the number of days that they have been imprisoned. If one travels even further down the list of words associated with the topic, one sees other words that similarly seem included for narrative rather than strictly "topical" reasons: *escape*, *execution*, and *suicide*.

A topic model provides a straightforward method for visualizing the rise and fall of themes and discourses in a collection of documents over a long period of time. Since all documents are assigned a proportion for each topic, we can view the topic breakdown for all documents for a given period. For instance, we can get a sense of the "typical" document in a given year by simply averaging the topic distributions for all documents associated with that year. We can think of this average as the topic proportion we would expect for an imaginary novel that used all topics at exactly the average for the period. I retrieved the five highest topics by topic proportion for the "average" novel for each five-year span of the corpus, where each novel is grouped based on the year that it first appeared on the *Times* list (rather than publication date). **Table 1.2** below shows these topics for three different five-year spans of the corpus, organized into a table. I've included topics only from three periods in order to keep the table a manageable size.

How does the "average" bestseller differ in these three periods, spread across the full timeline of the corpus? In the first period, the "average" bestseller is dominated by topics 22, 69, and 27. Topic 22 seems to describe scenes of urban middle-class sociality, composed of words like *darling, evening, table, dinner, girl, conversation, party, lunch,* and *afternoon*. Looking somewhat further down the list of top words, one sees words that reflect a concern with the subtleties of interpersonal behavior, befitting realist scenes of this type: *feeling, eyes, expression, tone, silence,* and *glance*. An examination of the novels in this period with the highest proportion of their discourse devoted to this topic adds some support for this interpretation. In the 1940s, there is A. A. Milne's *Chloe Marr,* a novel about a young woman making her way through London society, seen through the eyes of the friends and suitors who encounter her. Another novel with a high proportion of discourse devoted to this topic is Henry Green's *Nothing,* another novel about London's upper-middle class, told almost entirely in dialogue (many of the novels in Topic 22 are by British authors). Topic 27 has words drawn from historical fictions, especially from scenes of interaction between historical politicians and military leaders. Top words include *manner, master, attention, men, company, gentleman, country, reason, business,* and *service*. Top novels include *Proud Destiny*, Lion Feuchtwanger's novel about Benjamin Franklin's trip to France, as well as a number of Thomas B. Costain's historical novels.

period	topic	example words	example novel
1946-1950	22	darling evening table dinner girl conversation party lunch afternoon	*Chloe Marr*, by A.A. Milne
	69	people things lot course house place kind day	*The Butterfly*, by James Cain
	27	master attention men company gentleman country service	*Proud Destiny*, by Lion Feuchtwanger
	18	eyes face children house voice	*Pavilion of Women*, by Pearl S. Buck
	43	hand eyes face head hands mouth hair fingers	*Knock On Any Door*, by Willard Motley
1970-1975	69	people things lot course house place kind day	*And Other Stories*, by John O'Hara
	89	press meeting committee people country campaign	*Preserve and Protect*, by Allen Drury
	25	life mind course day fact truth sense reason death	*The French Lieutenant's Woman*, by John Fowles
	99	police car door telephone policeman street name	*The Salzburg Connection*, by Helen MacInnes
	67	kid hell lot guys kind kids stuff right couple shit ass	*Semi-Tough*, by Dan Jenkins
2000-2005	68	life years day night children kids things	*Bittersweet*, by Danielle Steel
	46	phone house car head night cell minutes call kitchen couple job coffee	*Mr. Perfec*t, by Linda Howard
	55	eyes voice face head hand smile look things hands	*No Safe Place*, by Richard North Patterson
	43	hand eyes face head hands mouth hair fingers	*The Fiery Cross*, by Diana Gabaldon
	79	police case murder detective office crime	*Ashes to Ashes*, by Tami Hoag

Table 1.2. Top five topics in three different five-year periods in the bestseller corpus.

These topics mostly fade from prominence as the century continues. Two of the topics that become newly prominent have something to do with crime or detective fiction: topics 99 (*police, car, door, telephone, policeman, street, name*) and 79 (*police, case, murder, detective, office, crime*). Looking at some of the novels that devote a large portion of their discourse to these topics, it seems that they might be more properly described as thrillers. Helen MacInnes' novel *The Salzburg Connection* is a thriller about a British spy on a mission to retrieve a box in Austria containing evidence that would implicate Nazi collaborators. Tami Hoag's *Ashes to Ashes*, meanwhile, is a crime thriller about a serial killer known as "The Cremator," who burns the bodies of their victims. Topic 46, ambiguous at the level of diction, is comprised of many novels in the genre known as "romantic suspense," which typically features narratives structured around both personal relationships and criminal threats. Examples of novels in the corpus with a high value for this topic include Suzanne Brockmann's *All Through the Night*, Linda Howard's *Mr. Perfect*, and Catherine Coulter's *Riptide*. Coulter's *Riptide*, which follows a couple that both work for the FBI, gives a sense of the genre's interests. One can see a way of attributing its top words to the genre's joint interest in romance and thrills: it includes scenes of domestic life (*kitchen, couple, house*) and words that, with some imagination, might plausibly be associated with suspense (*minutes, cell, car, coffee*).

The prominence of the thriller and romantic suspense genres among these topics helps shed light on some of the other topics that take center-stage in the later period, which are organized around terms that make reference to specific details of the human body. This includes Topic 55 (*eyes, voice, face, head, hand, smile, look, things, hands*) and Topic 43 (*hand, eyes, face, head, hands, mouth, hair, fingers*). The top novels for these topics are Richard North Patterson's political thriller *No Safe Place* and Diana Gabaldon's *The Fiery Cross* (fifth in the

Outlander series), respectively. These topics have less to do with narrative content, in the sense of the events, and more to do with the types of objects that the narrator focuses on in the course of narration. Consider a scene very early in *No Safe Place*, in which an anti-abortion extremist murders three employees at a women's health clinic:

> The receptionist's throat moved in a convulsive swallow, choking her words. "What do you want?"
>
> Sean took out his gun. "To stop you," he answered, and pulled the trigger.
>
> It was a kind of magic. As her head snapped back, a hole opened in her forehead. There was a soft concussive sound, like a melon hitting cement, almost lost in the red-haired girl's cry.
>
> Sean stared in stupefaction as the woman died in front of him, blood trickling from her forehead. Only when she hit the carpet did Sean turn to the girl.[13]

Notice the way that references to body parts drive the narration. The receptionist's fear is registered in movements of her throat; the cry comes from the "red-haired" girl. Even the killer's own affect is described by means of his exterior appearance: he "stared in stupefaction." In terms of narrative style, this is a an almost entirely exterior world. The narrator guides the reader's attention, but gives little insight as to characters' interior states, as though pointing a camera. It almost reads like a screenplay.

In terms of the topical makeup of *New York Times* bestsellers, then, there is a fairly dramatic difference between novels at the beginning and end of the period. By the turn of the century, the bestsellers have dropped their interest in dinner parties and upper-class conversation; the bestseller list no longer features novels of manners imported from Britain. Similarly, the historical novels of authors like Thomas Chastain have gone out of style. The genres that make it onto the list have changed: historical fiction and domestic realism are out, crime thrillers and romantic suspense are in. This shift is also visible the genre labels that have been assigned to

---

13. Richard North Patterson, *No Safe Place* (New York: Knopf, 1998).

these novels by cataloguers. Consider **Figure 1.2**, constructed by matching *Times* bestsellers with a corresponding entry in the Library of Congress catalogue and examining their assigned “Form/Genre” heading. I was able to acquire genre information for 4,060 of the 6,806 books on the Times list, 1,500 of which are among the 2,440 for which I also have text data from HathiTrust. The figure depicts the proportion of novels for each five-year period that were tagged with suspense (the Library of Congress’ term for “thriller”), mystery (including “detective and mystery”), historical, or domestic. Note that novels can have more than one tag.

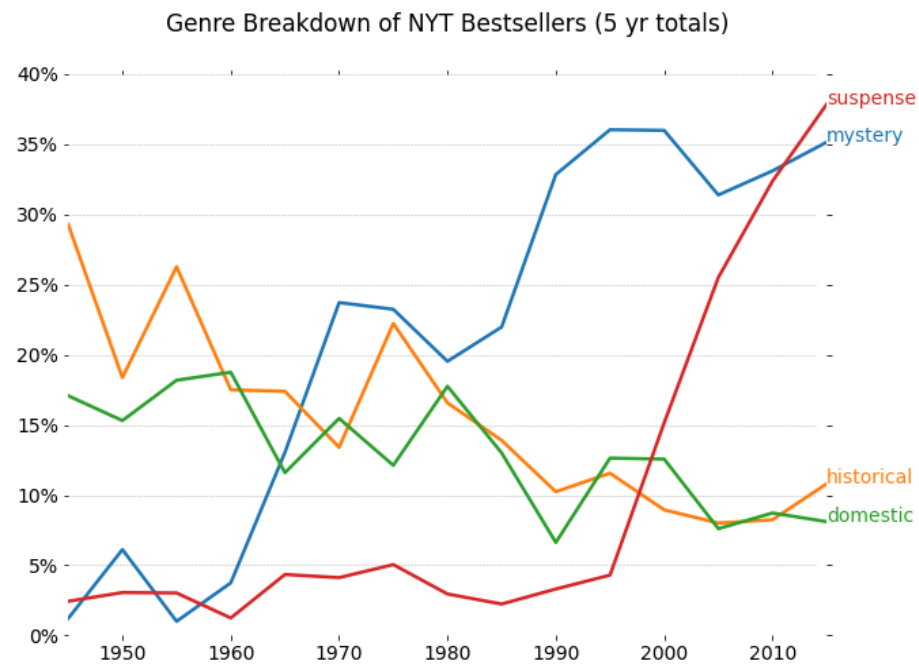

Figure 1.2. Genre breakdown of NYT bestsellers (five-year totals).

The advantage of the Library of Congress data is that it includes information on many of the novels that were on the *Times* list but for which I was not able to find a corresponding volume in the NovelTM dataset. As such, the fact that it tells much the same story suggests that the reversal of the generic makeup of the average *Times* bestseller is not just due to quirks of the HathiTrust collection. If anything, it paints an even more dramatic picture than the topic model.

Domestic fiction declines in popularity, from a high of about 18% in 1960 to less than 10% by the end of the period. Similarly, although almost 30% of bestsellers between before 1950 are described by librarians as "historical fiction," that proportion has declined to around 10% by 2015. Meanwhile, books described as "mystery" or "suspense" follow an even more dramatic trajectory in the opposite direction. By the end of the period, both "mystery" and "suspense" appear as genre descriptions for more than one in three of all bestselling novels for which the Library of Congress lists a genre description.

The Library of Congress data is just in the form of labels; these trend lines are not based on a quantitative analysis of the actual text. But computational analysis of the style of the novels tagged with these labels produces a similar story as the one we saw with the topic model. Following the procedure of previous computational humanities work using text classification, I trained two logistic regression models to distinguish between genres based on word frequencies.[14] I trained two models: one to distinguish "historical" from "mystery," and one to distinguish "historical" from "suspense." The models can distinguish their categories with 96% and 93% accuracy, respectively. Anything better than 50% is better than a random guess, so these are very distinct categories of fiction. Some of the words that the model finds most useful for identifying each genre are listed below. Note that this method evaluates all words, as distinct from the topic model, which only considered nouns.

---

14. Ted Underwood, "The Life Cycles of Genres," *Journal of Cultural Analytics,* Vol. 2, Issue 2 (May 24, 2016). https://doi.org/10.22148/16.005

historical		mystery		suspense		
loyalty	*war*	*murder*	*arrest*	*quick*	*anxiety*	*kick*
state	*civil*	*clue*	*asked*	*rushing*	*instinct*	*slumped*
dignity	*carriage*	*scene*	*remember*	*screaming*	*suspicion*	*pain*
argued	*navy*	*detective*	*admitted*	*echoed*	*alert*	*crush*

Table 1.3. Top words by coefficient for genre classifiers.

The terms associated with "historical" and "mystery" have a clear relationship to content. Historical novels are concerned with *war*, as well as one's *loyalty* to a *state*. This set includes novels like Lella Warren's *Foundation Stone*, chronicling a family of Alabama settlers from the 1820s and through the civil war; and Herman Wouk's *The Winds of War*, about the Second World War. Detective novels are about, well, detectives, as well as clues and murders. They depict scenes in which people are *asked* to *remember* important information. Examples include Susan Isaacs' *Magic Hour* and Walter Mosley's *Black Betty*. But the terms associated with "suspense" novels are evocative of mood or style as much as content. These novels depict a world run on *instinct*, with characters who are always on *alert*. The genre also favors words used to describe speed and the human body. Words like *kick*, *slumped*, *pain*, and *crush* are words you might encounter in a scene of sudden violence. The genre also has a faster pace, evident on the page itself: thrillers have, on average, fewer words per page than historical novels (~370 vs ~427).

These results add nuance to a recent piece by Andrew Piper and Eva Portelance, which studies the textual different between different segments of the contemporary literary market, including prizewinners, bestsellers, and genre fiction.[15] They find that, in terms of diction, the most dramatic differences between bestselling and prizewinning fiction have to do with descriptions of time. Bestsellers describe time urgently, at the scale of *minutes* or *seconds*. Prizewinners, meanwhile, operate at a slower pace and in a more retrospective mode; they are more likely to use words like *summer*, *winter*, and *childhood*. Further down their published list of words associated with bestsellers, however, are a number of words that are clearly associated with specific popular genres, especially detective fiction. For example, they report that bestsellers more frequently use words like *investig, polic, arrest, lawyer, gun, bullet, murder, statement,* and *victim* (they report word stems rather than full words).

It seems plausible, then, that their characteristic bestseller is a relatively recent phenomenon. To explore this possibility, I took the 200 words that Piper and Portelance reported as being most closely associated with bestsellers and prizewinners, respectively, ordered by the reported effect size of the association. Then I calculated the annual proportion of all words in the *Times* corpus which are one of these words. **Figure 1.3** shows resulting proportions plotted over time. If this measure is any indication, then the characteristic bestseller reported by Piper and Portelance did not begin to emerge until the early 1990s. In fact, if we turn back the clock to before 1960, the picture gets flipped on its head: words that Piper and Portelance report as being associated with prizewinners actually occupy more space in the bestseller lexicon than words they report as being associated with bestsellers. The 200 words within this trend includes retrospective or nostalgic words that they report as characteristic of the contemporary style of prizewinning novels, such as *childhood* and *youth*, words for family relationships such as *mother*

15. Andrew Piper and Eva Portelance, “How Cultural Capital Works: Prizewinning Novels, Bestsellers, and the Time of Reading,” *Post45* (May 10, 2016).

and *uncl*, and words describing natural environments, such as *tree, mountain, fog, horizon,* and *cloud*. In this analysis, the period between 1960 and 1990 appears as a kind of transitional period, where the contemporary bestseller style coexists with the prizewinner style. The break point in 1990 seems important, as that is also the year that "suspense" fiction becomes more common among bestsellers, according to the Library of Congress data.

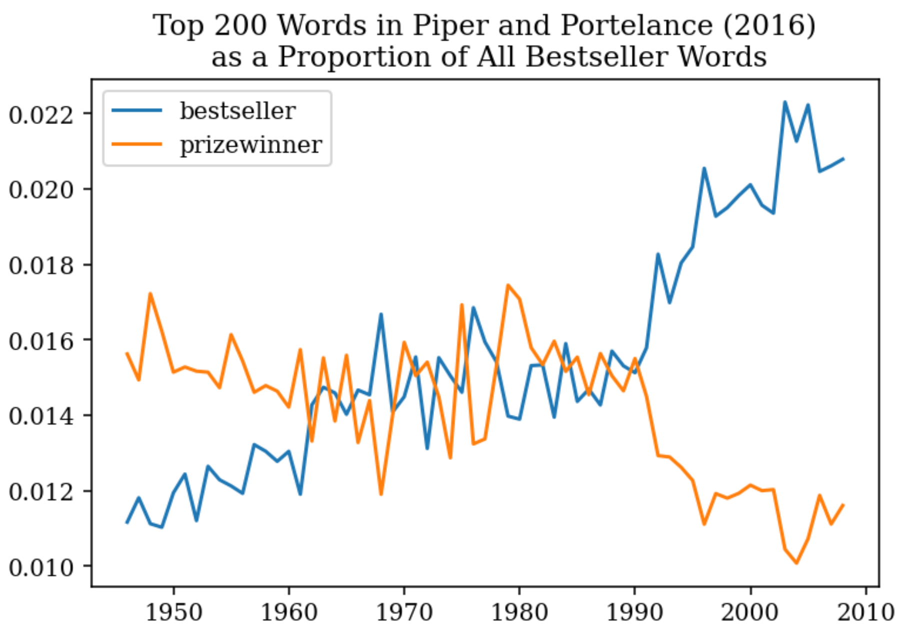

Figure 1.3. Top 200 words in Piper and Portelance (2016) as a proportion of all bestseller words.

The inversion of bestseller genres that I have been tracking up to this point also resembles a result from a recent piece by James F. English, although the dates do not line up quite as directly.[16] In this project, English assembles a corpus including novels that appeared on the *Publishers Weekly* annual bestseller list and novels that were shortlisted for a major literary

16. James F. English, "Now, Not Now: Counting Time in Contemporary Fiction Studies," *Modern Language Quarterly* 77: 3 (September 2016): 395-418.

prize. Then, he tags whether the novel was set twenty or more years in the past (at the time of its writing), within twenty years of the present, or twenty or more years in the future. Aggregating over the whole corpus, English finds an inversion similar to the one that I find here, except played in a different key: bestsellers became less likely to be set in the past, and more likely to be set in the present, after about 1980, whereas prizewinners moved in the opposite direction, becoming more likely to be set in the past. Although my method in the chapter does not directly consider the time of a novel's setting, I do similarly find that explicitly "historical" fiction became less likely to be a bestseller as the twentieth century approached its close. And although I don't have quantitative confirmation of this, anecdotally, the suspense and crime novels that become hegemonic after 1990 are more likely to be set in the present.

Let us summarize what we have observed so far: between 1946 and 2009, the style and genre of the typical *New York Times* bestseller shifted. In the earlier part of the period, historical novels and domestic fiction were especially prominent on the list. This began to change between 1960 and 1980, when these genres started to be replaced by mystery or detective fiction, and especially after 1990, when suspense or thriller novels became the decidedly dominant genre. Although previous work showed that contemporary bestselling novels are characterized by fast-paced language that depicts time at the scale of minutes or seconds, the results of the above analysis suggest that this is a relatively recent phenomenon. Bestsellers before 1960 were actually more likely to depict time at the slow, contemplative scale previous research has shown is, today, more characteristic of novels that win literary prizes. [17] The fast-paced language associated with bestsellers is particularly true of genres like suspense or thriller novels, which are more likely to be bestsellers today than they were in the immediate postwar period.

17. Piper and Portelance, 2016.

What to make of this? In the following section, I consider how to interpret this phenomenon in light of publishing history. My explanation hinges on the history of genres and formats. In particular, I focus on how the trend observed in my analysis—and, possibly, the results of English's project referenced above—might be in part a symptom of changing norms around book formats near the end of the twentieth century. As I discuss below, publishers in the 1970s, and especially after the 1980s, were increasingly willing to publish traditionally mass-market genres like crime fiction in hardcover. Since the *Times* list is specifically a hardcover list, it will obviously reflect changes in the kinds of books that are published in hardcover.

1.3. The Mass Market Hardcover?

Returning to the topic model for a brief moment, it is worth noting the class-specific character of many of the novels that made it onto the list in the early period of the corpus, between about 1946 and 1960. Topic 22 is made up of many British society novels that explicitly describe characters moving through upper-class life, like *Nobody* and *Chloe Marr*. Similarly, one can discern an element of class taste in many of the novels in this period that have a "historical" label in the Library of Congress. Many of these novels are examples of what Tom Perrin has usefully called the "middlebrow epic."[18] In this category, Perrin includes long realist epics like John Steinbeck's *East of Eden* (1952), Taylor Caldwell's *Dear and Glorious Physician* (1959), and James Michener's *Hawaii* (1959), as well as novels by Pearl S. Buck and Lloyd C. Douglas. All of these authors appear on the bestseller list at least once in the period before 1960.

18. Tom Perrin, "The Second-Greatest Stories Ever Told," in *The Aesthetics of Middlebrow Fiction: Popular US Novels, Modernism, and Form, 1945-75* (New York: Palgrave MacMillan, 2015): 91-108.

The reason why there may be a class component to the contents of the bestseller list is that the *Times* list is specifically a hardcover bestseller list, and hardcover novels were more expensive than paperbacks, especially in the early postwar period. In the decades following World War II, the publishing industry was characterized by a schism between the more expensive hardcover format, generally sold in bookstores, and the cheaper mass market paperback format, which could be found in a variety of outlets. The “paperback revolution” of the 1940s achieved its broad reach because publishers made use of existing magazine distribution infrastructure, selling their cheap books to less wealthy, more geographically diverse readers shopping at drugstores, newsstands, and supermarkets.[19] The market division was reflected on the pages of those magazines that published lists of bestsellers: neither the *New York Times* nor *Publisher’s Weekly* included regular weekly paperback lists until 1976 (though the *Times* did publish irregular monthly paperback lists as early as the 1960s).[20]

By virtue of the higher price of the hardcover books whose sales they publicized, bestseller lists in the early postwar period likely reflect the tastes of comparatively wealthier readers. As a history of bestsellers co-written by one of the compilers of the *Publisher’s Weekly* list puts it, these lists often “represented the conservative taste of those who could afford to buy rather high-priced hardcover novels.”[21] In some cases, the format distinction is not important, since books would get reprinted in paperback after a hardcover release and high sales numbers in one format could translate to high sales in the other. Where the “conservative taste” of the hardcover book-buyer begins to matter is when a sold disproportionately many more copies in

19. For a general history of the paperback revolution, see Kenneth C. Davis, *Two-Bit Culture: The Paperbacking of America* (Boston: Houghton Mifflin, 1984).

20. Davis, 294.

21. Alice Payne Hackett and James Henry Burke, *80 Years of Best Sellers, 1895-1975* (R. R. Bowker: New York, 1977), 214.

paperback than in hardcover, or in the case of paperback originals that never received a hardcover printing at all. Importantly, paperback originals were often works of popular genre fiction that could achieve sales numbers exceeding those of hardcover bestsellers. Erle Stanley Gardner's *Perry Mason* novels, for instance, would not have even been considered by the *Times* despite the fact that they routinely hit sales numbers that would have qualified them in hardcover. In the middle of the twentieth century, a hardcover novel could hit an annual bestseller list with sales of over 100,000 copies. Ex-publisher Michael Korda, for instance, cites Nabokov's *Lolita* as hitting the *Publisher's Weekly* annual list in 1958 with hardcover sales of 150,000.[22] Erle Stanley Gardner's paperback books, by contrast, would often have first printings in excess of 750,000 or even 1 million copies.[23] Gardner does not appear on the *Times* list even once, though *Lolita* does. Another author of crime fiction from this period who is conspicuously absent is Mickey Spillane. Spillane only made it onto the *Times* list a single time, with his 1952 novel *Kiss Me, Deadly*. It topped at barely number eleven on the list, even though it sold millions of copies in paperback.[24] Some portion of the relatively low numbers of "crime" and "suspense" bestsellers prior to 1960 is due, then, to the fact that the top-selling authors in these genres were limited to paperback, and therefore were not being counted on hardcover lists.

Notably, in the 1970s and 1980s, when crime fiction was becoming more prominent on the *Times* list, the boundaries between hardcover and mass market formats in publishing were

---

22. Michael Korda, *Making the List: a Cultural History of the American Bestseller, 1900-1999* (New York: Barnes and Noble, 2001), 104.

23. Davis, 125.

24. Janet Staiger reports that Spillane's first six novels—*Kiss Me*, *Deadly* was his sixth—averaged 2.5 to 3 million copies sold each. If those were hardcover sales, every single one of those novels would have made the list.

Janet Staiger, "*Kiss Me Deadly*: Cold War Threats from Spillane to Aldrich, New York to Los Angeles, and the Mafia to the H-Bomb," in *New Directions in American Reception Study*, ed. Philip Goldstein and James L. Machor (Oxford: Oxford University Press, 2008): 281.

blurring. Traditionally, the phrase “mass market paperback” referred to both a book format and a type of publisher. Mass market publishers like Avon and Bantam did not publish in hardcover for approximately the first thirty years of their operation; in fact, for years they did not publish original books at all but simply acted as merchandisers for reprints of books released by traditional hardcover publishers.[25] In the 1970s and especially in the 1980s, however, it became increasingly common for the same company to control imprints in both formats. In some cases, mass market publishers started publishing hardcover books on their own, leveraging their marketing experience into new terrain. Both Bantam and Warner Books began their first hardcover line in 1980.[26] Hardcover houses were also themselves becoming increasingly involved in the mass market. As Phyllis Grann of Putnam explained in 1986, hardcover houses were learning “how to apply mass merchandising techniques to hardcover books” and were “penetrating markets once closed to hardcovers,” such as supermarkets.[27]

As these boundaries blurred, paperback-only authors gradually found opportunities for hardcover publication. Many of these authors will be familiar as the brand-name authors of genre fiction who still dominate bestseller lists today. Danielle Steel, for instance, made her debut in hardcover in 1981 with *The Ring*, a romantic saga that spans from Nazi Germany to then-contemporary America. Her publisher, Dell, was originally opposed to the idea, convinced that her readers would not accept hardcover prices.[28] Other publishers were more forward-thinking about urging their authors to change formats. Bantam moved Clive Cussler into hardcover with

---

25. Davis, 294.

26. Jennifer Chrichton, “Warner Books: a Love Affair with the Bestseller,” *Publisher’s Weekly* (Riverton, NJ), July 13, 1984: 18.

27. Carolyn Anthony, “New Players in the Mass Market,” *Publisher’s Weekly*, June 27, 1986: 30.

28. Lorenzo Benet, *The Lives of Danielle Steel: The Unauthorized Biography of America’s #1 Best-Selling Author* (New York: St. Martin’s Press, 1995), 201.

his 1981 *Night Probe!*, his fifth novel featuring adventurer and shipwreck-salvager Dirk Pitt. Both books were *New York Times* hardcover bestsellers.

These successes were fairly modest compared to other hardcover bestsellers in the 1980s, many of which broke first-printing records. In fiction, the big story was Stephen King. Unlike Steel and Cussler, King had been published in hardcover since his first novel *Carrie* (1974). But his hardcover sales hadn't compared to his paperback sales; *The Shining* (1977) sold 50,000 copies in hardcover but more than two million in paperback. In the 1980s, his hardcover sales started to catch up with paperback: 285,000 for *Firestarter* (1980), 350,000 for *Cujo* (1981), and 657,000 for *Pet Sematary* (1983). In 1984, *The Talisman*, co-written with Peter Straub, set records with a first printing of 600,000 copies, only to be surpassed by *It* (1986) and *The Tommyknockers* (1987), which both passed one million.[29] This represented a shift in the structure of the book market, as hardcover novels began to be purchased by new readers. As Ralph Vinccinanza, one of King's agents, put it in an interview with *Publisher's Weekly*, hardcover publishing historically "catered to the intelligentsia… but Stephen changed all that."[30] King was only the most high-profile example of a broader shift. Linda Grey, then vice-president of Bantam, noted in a 1986 interview about the rise in hardcover sales that it was already becoming common wisdom among publishers that "the hardcover, some hardcovers anyway, has crossed over to the onetime mass market reader."[31] One retrospective article describe the 1980s at the "decade of the mass market hardcover."[32]

---

29. Bill Goldstein, "King of Horror," *Publisher's Weekly*, January 24, 1991: 9.
30. Goldstein, 9.
31.Anthony, 31.
32. Elizabeth Mehren, "The Decade of the Mass-Market Hardcover," *Los Angeles Times*, December 31, 1989. https://www.latimes.com/archives/la-xpm-1989-12-31-bk-139-story.html

By the 1990s, it had become more ordinary for authors of mass market fiction to move into hardcover. During this period, some of the authors of genre fiction that appeared prominently in the quantitative section of this chapter were making their own transitions into hardcover. Notably, many of the authors of romantic suspense (a subgenre that contributes to the spike in "suspense" bestsellers in the 1990s) on the bestseller list got their start writing mass market paperback romances. This includes authors like Sandra Brown, Tami Hoag, Linda Howard, and Iris Johansen.[33] Recall that both Tami Hoag and Linda Howard appeared prominently among crime and suspense topics in the topic model.

Why did the scale of hardcover publishing change in the 1980s? Publishers sometimes attributed the shift in part to changes that were occurring on the distribution side of the industry. Hardcover books were being marketed to new audiences and were being sold in new venues. In particular, the rise of chain bookstores like Dalton and Waldenbooks brought bookstores to regions of the country that had never had a bookstore.[34] Richard Snyder, then-president of Simon and Schuster, said to a publishing reporter in 1980 that

> ...what is really changing the face of publishing in America is not the conglomerates but the giant book chains… it's the fact that the chains serve a different community of book readers from any that the book business has ever had before… The minute you get into the suburbs, where ninety per cent of the chain stores are located, you serve the customers, mainly women, the way you would serve them in a drugstore or a supermarket.[35]

Drugstores and supermarkets: the traditional outlets of mass market paperbacks. For Snyder, chain stores were pushing the hardcover market to resemble that of mass market paperback, by

---

33. "What Is This Thing Called Love?" *Publishers Weekly*, November 9, 1998: 43.

34. For a history of book chains, see Laura Miller, *Reluctant Capitalists: Bookselling and the Culture of Consumption* (Chicago: University of Chicago Press, 2007).

35. Thomas Whiteside, "Onward and Upward with the Arts—The Blockbuster Complex, II," *New Yorker*, October 6, 1980: 136, 138, quoted in Janice Radway, *Reading the Romance: Women, Patriarchy, and Popular Culture* (University of North Carolina Press, 1984): 37-38.

bringing hardcover books to mass market readers. The chain stores increased the geographic accessibility of hardcover books. B. Dalton and Waldenbook opened hundreds of new stores, often in suburban shopping malls.[36]

Hardcover novels were also becoming more affordable relative to mass market paperbacks. Whereas in 1960 readers could purchase eight or more paperbacks for the price of a single hardcover novel, the ratio was closer to four-to-one by the early 1990s. The price gap could be even smaller in the case of bestsellers, since it was common practice at chain stores to heavily discount novels that hit the *Times* bestseller list. During the mid-1990s, for instance, both Borders and Barnes and Noble would typically discount all hardcovers by 10% and *New York Times* bestsellers by 30%.[37]

These factors combined to weaken the historical division between mass market and hardcover publishing. By 1990, the terrain of book distribution in the United States looked very different. For some publishers, the expanded scale of the hardcover book market felt unprecedented. As the vice president of one publishing house put it, “At the beginning of the 1980s, you could not imagine first printings of 500,000… The fact that 1.5 million people walk into a bookstore and pay $20--and up--for a hardcover book is mind-boggling.”[38] Ian Ballantine,

36. Lewis A. Coser, Charles Kadushin, and Walter W. Powell, *Books: the Culture and Commerce of Publishing* (New York: Basic Books, 1982): 349.

37. Laura Miller, *Reluctant Capitalists*, 147. Discounting was already a common theme in trade journals by the 1980s. See, for instance, “Hot Hardcovers for Summer,” ed. Joann David and William Goldstein, *Publisher’s Weekly*, April 24, 1987:36, which attributes “rising hardcover sales” in part to the “shrinking difference between hardcover and paperback prices” driven by “discounted prices” and “rising disposable income.” Phylis Grann of Putnam likewise emphasized discounting, noting that a discounted $20 hardcover might sell for $15, a price tag that is “not so far from the $5 paperback,” and so “people see a better bargain in the hardcover.” Anthony, 30.

38. Elizabeth Mehren, “The Decade of the Mass-Market Hardcover,” *Los Angeles Times*, December 31, 1989.

founder of Bantam Books, saw the shift as a victory for the mass market paperback publishers, whose marketing and distribution strategies had become industry norms. Though paperback publishers had begun as reprinters subordinate to the New York publishers, they had changed the way that the industry does business. Or, as he gleefully put it to Charles Scribner in the 1980s: "The mass market publishers control the world!"[39]

1.4. Conclusion: Seeing Like a Publisher

The changing norms in publishing around book formats in the final decades of the twentieth century put a different spin on the results of the previous section, as well as the results of other projects that use bestseller lists as a point of reference. Though something clearly changed about the style and genre of *New York Times* bestselling novels between about 1960 and 1980, that change is probably due in part to a migration of formats: traditionally-mass market authors and genres were moving into hardcover. If we had access to a list of paperback bestsellers between 1940 and 1960, it is likely that the relative proportions of "suspense" and "mystery" novels on that list would be higher than on the *Times* list. The shift was not a literary one, or at least, was not exclusively literary: crime thrillers like those of Mickey Spillane were popular and influential the 1940s and 1950s, but due to the deep divide between hardcover and paperback formats at that time, they do not appear on published bestseller lists. As that divide softened, and crime thrillers moved into hardcover, their popularity was increasingly registered by the *Times* bestseller list.

39. Joseph Barbato, "A Cheerful 40th Birthday for Bantam," *Publisher's Weekly,* October 18, 1985: 28.

But I suspect that this is the kind of historical change that is at risk of going unobserved in quantitative literary studies. For example, James F. English's aforementioned project on the setting of bestselling and prizewinning novels does not consider the fact that the *Publishers Weekly* annual bestseller list likely undercounts paperbacks before the 1970s, if it counts them at all. How would our interpretation of his finding change if we considered the distinction between hardcovers and paperbacks? For one thing, the trend would look considerably more like an artifact of composition: it's not that the top-selling novels for Americans as a whole were becoming more focused on the present, but rather than the top-selling novels among hardcover customers were moving towards the present. There are different reasons why this might have been the case. It might have been that the taste of relatively wealthier readers had changed: hardcover readers decided they wanted to read about crime, rather than biographies of historical figures or the social lives of the British upper class. Or it might have been that traditionally mass-market readers became more willing to purchase hardcover books, perhaps because they were becoming more affordable, or because their status as objects of pretension had softened. Distinguishing between these explanations—or other possible explanations—is beyond the scope of this chapter. But even formulating these hypotheses in the first place requires understanding the relationship between book formats and their respective readerships. In other words, it requires seeing the literary field how a publisher sees it.

For publishers, the format distinction is a fundamental aspect of their trade, particularly in the historical period considered in this dissertation. Because of the particularities of book prices and distribution chains, the format in which a book is printed has an important effect on which potential readers even have a chance of encountering it in the first place. Readers who purchase new hardcover releases in urban bookstores have different purchasing habits than those who buy paperbacks in a rural drugstore. These differences in purchasing habits are a fundamental consideration when deciding what to publish, and how to publish it.

Quantitative literary history can learn much, I suspect, by learning to see literature as a publisher sees it. For a publisher, literature is fundamentally big: thousands of new titles are released each year, purchased by millions of readers. The challenge of ensuring that those titles get matched up with those readers in sufficient proportions gives publishers a powerful motivation to try and understand—and, as much as possible, predict—which readers will purchase which books. Sometimes, that means chasing the divides in the market, such as divides along class, income, and geography. Other times, it is as simple as trying to repeat the past: just publish the next novel by whoever was on the *New York Times* bestseller list last year or the year before that. In quantitative literary studies, we are often interested in studying literary-historical trends that are visible over decades or even centuries. At that scale, observed trends are likely to be affected by the institutional and demographic structure of the book market. Making sense of our quantitative results may require studying the institutions that publishers and booksellers have set up to organize that market.

Chapter 2. Literary Brand Names and the Political Economy of Authorship

If you spend enough time staring at the spreadsheets of *New York Times* bestsellers considered in the previous chapter, you'll begin to notice a curious phenomenon. As the list approaches the end of the twentieth century and moves into the twenty-first, dual-authored novels become increasingly common. What began as a relatively unusual practice yielding only a handful of bestsellers per year morphs into an institutional norm. In the 1950s and 1960s, dual-authored novels were infrequent. Many of these are co-authored by family members or spouses such as Benedict and Nancy Freedman's *Mrs. Mike*, or Marrijane and Joseph Hayes *Bon Voyage*. In the 1990s, the publishing industry launches bestsellers by dozens of authorial pairs per year, and many of these pairs hit the list multiple times. They are most prominent in series fiction, especially long running series in a well-known genre, as in Tim LaHaye and Jerry B. Jenkins'

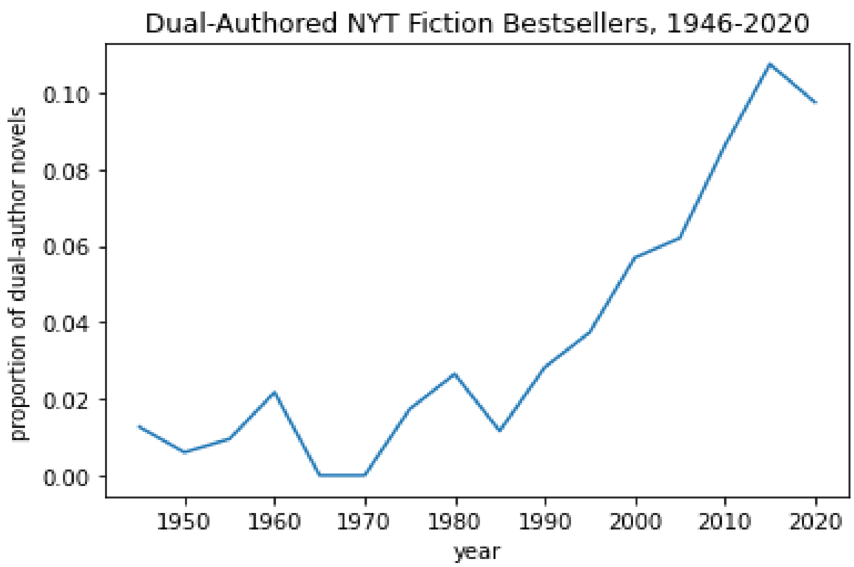

Figure 2.1. Dual-Authored NYT Fiction Bestsellers, 1946-2020. The vertical axis represents the proportion of all bestsellers with two or more authors listed.

Left Behind series, a cycle of Christian post-apocalyptic novels, or James Patterson and Maxine Paetro's *Women's Murder Club* series, which features an ensemble of female leads solving crimes in San Francisco. In some cases, the industry invites a celebrity from another industry to co-write a novel with a professional writer, as in the case of several historical novels co-authored by ex-Speaker of the House Newt Gingrich and William R. Forstchen. A spike in the prevalence of dual-authored bestsellers begins around 1990 and continues into the new millennium. Today, as many as one in ten *New York Times* bestsellers are listed as having multiple authors.

The rise in prominence of this collaborative, nontraditional method of literary production would seem to provide an unusual type of confirmation for the historicist turn in authorship studies as it appears in book history, literary criticism, and the sociology of literature. As Christine Haynes observes in a 2005 article summarizing the state of authorship studies, the field has been increasingly dominated since the 1980s by an abundance of historical studies that investigated particular and contingent social arrangements in which textual production has taken place in modernity. [1] Though we have long known that the image of the isolated, Romantic genius has always been at least partially a fiction, this historicist strand of scholarship showed the particular historical reasons why this was the case. As Haynes puts it,

> in many times and places, authorship has often been more collective and interactive than individualistic and inspired. Whether in the context of a gentlemanly coterie, a theater performance, a patronage system, a writing partnership, a nineteenth-century fiction 'factory,' or even a modernist position in the literary field, authors acted not alone but in conjunction with fellow writers, benefactors, actors, printers, booksellers, editors, literary agents, and publishers.[2]

1. Christine Haynes, "Reassessing 'Genius' in Studies of Authorship: The State of the Discipline," *Book History*, Vol. 8 (2005).
2. Haynes, 314.

Historicist explorations of collective authorship have become commonplace; by now, literary scholars are well aware that historical actors have adopted a wide variety of practices for composing texts. But despite this, Haynes observes, many aspects of the Romantic model of authorship remain quietly taken-for-granted by literary critics and book historians. The problem extends beyond theory all the way to methodology. It remains common for scholars to, for example, engage in "biographical studies" of a text's author when attempting to make sense of its meaning and history. Though we are well-aware that what Foucault called the "author function" is in many ways a historical and institutional fiction, it is difficult to avoid reifying it indirectly, by appealing to evidence that depends on the author function for its intelligibility.

Haynes suggests that one way that scholars of authorship can finally shake themselves free of this methodological baggage is to "to understand cultural conceptions of authorship within broader discourses of political economy."[3] Citing Michael Newbury's work, Haynes argues that this requires considering authorship as an actual social practice, "embodied by particular forms of labor." Scholars can expand the scope of investigation beyond merely historical "ideas about authorship" to also include the actual "practices of writing" that were typical in a given time and place. Such a shift of focus necessitates a shift in objects as well: scholars should "return the manuscripts, proofs, and letters of authors, as well as the business records of publishers and the paratextual materials surrounding texts," not to determine the intention of authors but their "everyday habits," as well as the habits of the other actors with whom they worked.[4] Haynes' injunction is salutary insofar as it proposes a genuinely historical

3. Haynes, 315.
4. Haynes, 315.

materialist methodology for authorship studies. Authors no doubt look like different creatures when we consider their habits alongside their ideas about themselves.

In this chapter, I take up Haynes' injunction to consider authorship alongside political economy in the context of this rise in prominence of multiple-author bestsellers. The crux of my argument is straightforward. In the late twentieth century, the American publishing industry became increasingly accommodative of nontraditional authorial arrangements that introduced some division of labor into the composition of texts. These included dual-authored texts, ghostwritten novels, and literary franchises that exceed the work of individual authors and which may be owned by a larger corporation, independently of those authors. Though these arrangements have not and likely will not become the majority method of composition in any simple quantitative sense, they have become increasingly influential. The *sine qua non* of the authorial arrangements I consider in this chapter is that there exists some division of labor between two different aspects of textual composition that we are used to treating as united: conception, or the invention of new plots, characters, narrative situations, or other elements of story; and composition, or the rendering of that story into a particular textual form, or a discourse. As I will show, only the former of these two aspects of composition is necessary to being considered an author, legally and institutionally. All authors invent new ideas, but only some authors are writers.

Why is this distinction important? The stakes of my argument in this chapter have to do with the broader discussion of how the publishing industry and the larger literary market manage the problem of abundance. When it comes to the presentation of books in the market—on the shelves of bookstores, in the advertising pages of newspapers, or, today, in the search results of Amazon—the author's name is a major aspect of how the book is presented to prospective readers. Besides the title, the next biggest piece of text on a jacket's cover is almost always the name of the author (in some cases, the author's name is even bigger than the title). This name serves a sort of branding function: it organizes reception of the work, prompting readers to mentally liken it to other works by the author with which they are already familiar. Traditionally, this branding function marks something about the style of the book we are about to encounter, since we expect that an author's works will all be relatively similar to each other. But as I will show, the branding function of an author's name can sometimes come to stand in for a larger literary franchise, regardless of whether the named author actually wrote the book. In some cases, the named author may not even still be living.

This argument will be clarified with a specific example. Consider the case of James Patterson, whose dual-authored books frequently appear on the *Times* bestseller list. Though Patterson is one of the most commercially successful authors of all time, he has insisted in interviews that the actual composition of texts—the construction of an original sequence of words in an original order—is of secondary importance to the construction of an original story. As Patterson once put it, "Above all my brand stands for story. I became successful when I stopped writing sentences and started writing stories. Editors think it's about style. It's not. It's all story."[5] Patterson's account of his success has a double-meaning. On the one hand, he is proposing a hierarchy of aesthetic priority: attractive sentences count for nothing without a compelling story. But the statement is also literally true, to a point: Patterson achieved new heights of success when he stopped writing sentences and began outsourcing (some of) that labor to co-writers, focusing instead on the art of storytelling. Beginning with *Miracle on the 17th*

5. John Deighton, "Marketing James Patterson," *Harvard Business Review*, Case Study, 2004: 5.

Green in 1996, co-written with Peter De Jonge, Patterson has perfected a system of literary production in which he collaborates with a lesser-known writer, lending that writer his storytelling ability and literary brand. He is very open about this fact, that he literally does not write many of his sentences: it is “not an issue,” he has explained, since the real purpose of putting his “name on the cover” of books is to give their prospective buyers the “assurance of a good read.”[6] He explains his propensity for collaborative writing as less a matter of resting on his laurels than it is a matter of playing to his strengths, of adopting a division of labor commensurate with the skills of each individual participant. As he explains in a different interview, “In my case I’ve always been a good storyteller. I’m very good at plot and characterization but there are better stylists.”[7] In addition, he simply generates stories at a pace that he cannot personally keep up with, with outlines of dozens or hundreds of potential narratives resting on his desk at a time. So, he focuses on the characters, the structure of the narrative, and the plot twists, leaving the prose to one of those “better stylists.”

As the manager not only of his own writing labor but the writing labor of others, James Patterson might be described less as an author or even novelist as we have traditionally imagined that term but somebody closer to the director of a film or the creative director of a videogame. His role in the creation of new literary commodities is both creative and managerial: he conceives of new products and product lines and also directs co-authors to produce their texts in a particular way. Legally, Patterson remains the author, in the sense of the owner of texts’ intellectual property. American copyright law has over a century of precedent regarding the

---

6. Deighton, 2.

7. Simon Edge, “James Patterson - the bestseller who doesn’t write his own books,” *Express*, February 26, 2013. https://www.express.co.uk/entertainment/books/380231/James-Patterson-the-best-seller-who-doesn-t-write-his-own-books

"work for hire doctrine," in which authorship rights over a commissioned work of art are accorded to the one who hires the labor rather than the one who supplies it.[8] The work for hire doctrine provides the legal basis for much of the writing that takes place in the industries of artistic, cultural, and intellectual production; it affords culture-industry firms and their owners the assurance that any cultural commodities conceived or produced by their employees at work will remain the property of those firms and owners. In addition, the work-for-hire doctrine helps secure the division of labor used in the actual production of cultural commodities like films and television shows, in which even a task like writing is often performed by small collaborative teams rather than a single, isolated author.

The structure of the chapter is as follows. In the first section, I review the theoretical literature in literary studies and legal history on the concept and institution of the "author." I observe that the distinction between writing and authorship in modern copyright law actually depends on the idea of the Romantic author for its intelligibility. The author, like the Romantic genius, is the conceiver of the idea of the work, and may or may not be the same as the person who puts that work's idea into a specific form. In the following section, I consider how authorship fits into the Marxian theory of value. I note that, for Marx, texts do not have "value" in the technical sense once they have been created, since they do not require additional expenditures of living labor to be reproduced. I conclude that, in Marx's categories, authors are closer to owners of a financial asset than they are to wage laborers, and the income derived from

8. For a history of the work-for-hire doctrine, see Catherine Fisk, *Working Knowledge: Employee Innovation and the Rise of Corporate Intellectual Property, 1800-1930* (Chapel Hill: University of North Carolina Press, 2009).

For an elaboration of this history in the specific case of professionalized, salaried writers like screenwriters, see *Writing for Hire: Unions, Hollywood, and Madison Avenue* (Cambridge: Harvard University Press, 2016).

authorship is closer to a rent than a wage. As such, understanding the political economy of authorship requires attentiveness to how ownership of this unusual kind of financial asset is fought over, distributed, and secured.

The remainder of the chapter is dedicated to close analyses of particular authorial arrangements that shed light on how the asset of authorship is distributed in late-twentieth century and early-twenty-first century American literary production. I consider three authorial arrangements in particular, all of which introduce some explicit distinction between the writer and the author: commissioned works in which a writer is hired to execute the idea of another person; multimedia franchises that exceed the work of any individual author, which may comprise intellectual property owned by a corporate author; and popular book series that get completed by ghostwriters after an author's death. This last case is the clearest example of a situation in which authorship functions like a financial asset. I focus on the case of popular novelist V.C. Andrews, author of the novel *Flowers in the Attic*, as well as several bestselling sequels. After her death, her estate continued publishing novels in her name, prompting the IRS to assess an estate tax on the capitalized trademark value of her name. I close with a discussion of the recent phenomenon of "literary development companies," small companies that generate new narrative franchises and then hire professional writers to develop those properties into particular mass media forms, including novels, films, and television shows. I discuss the direction that these companies may take authorship in the future and argue that, although they will likely become more common, they are unlikely to ever entirely eclipse the classic arrangement in which author and writer are the same person.

2.1. What Is An Author?

Let us begin by revisiting that classic theme of poststructuralist literary theory: the question of what constitutes an author. For both Roland Barthes and Michel Foucault, the author was first and foremost a historical phenomenon. In his classic article on the contemporary author-figure, Foucault works to establish the historicity of this persona, pointing to its historical contingency, the legal and institutional regime upon which it depends, and the seemingly-arbitrary conceptual boundaries between the types of texts that seem to demand an attribution in order to be understood.[9] Foucault is explicit that modern authorship emerges alongside print, as a method of controlling the reproduction of politically dangerous or transgressive material. Authorship was originally characterized by what Foucault calls "penal appropriation," in that its main function was to make it possible for individual persons to be held responsible for texts, particularly when those texts were politically transgressive. From here, it gets elaborated in its modern form alongside a "circuit of property values," the modern regime of copyright that emerges only toward the "end of the eighteenth and beginning of the nineteenth century."[10] Furthermore, he notes that even in Europe, the genres of texts to which it has seemed necessary to assign an author has continually shifted: at one time or another, "narratives, stories, epics, tragedies, comedies" were "accepted, put into circulation, and valorized without any question about the identity of their author."[11] We might note, as well, that this second feature of authorship is closely related to the first: stories seem to require an author more once they become commodities, objects of sale and ownership, rather than freely-distributed common knowledge.

---

9. Michel Foucault, "What is an Author?" in *Textual Strategies: Perspectives in Post-Structuralist Criticism*, ed. Josue V. Harari (Ithaca: Cornell University Press, 1980).

10. Foucault, 148.

11. Foucault, 149.

Barthes, meanwhile, explicitly associates the “author” with capitalist modernity. “The author is a modern figure,” he announces, “a product of our society, insofar as… it discovered the prestige of the individual, or, as it is more nobly put, the ‘human person.’”[12] Continuing, he contends that “It is thus logical that in literature it should be this positivism, the epitome and culmination of capitalist ideology, which has attached the greatest importance to the ‘person’ of the author.”[13] Where Foucault adopts a largely analytical posture, Barthes is polemical, describing the reign of the author as “tyrannical.” His “death of the author” is both a historical proclamation and a performative utterance, an attempt to kill the author, to dislodge the cultural authority that authors wield over the texts they are said to have authored. In some places in this short, manifesto-like piece Barthes almost sounds like he is calling for revolutionary violence; literary criticism meets Sacco and Vanzetti. Without the author, Barthes announces, writing can be seen as an “anti-theological activity, an activity that is truly revolutionary since to refuse to fix meaning is, in the end, to refuse God and his hypostases—reason, science, law.”[14] The removal of the author further resituates the agency of interpretation to readers: “we know that to give writing its future, it is necessary to overthrow the myth: the birth of the reader must be at the cost of the death of the Author.”[15]

Despite the emphasis of both writers on the historical specificity of the author-function to capitalist modernity, neither text addresses the author from the point of view of production, as a person who receives income in return for the creation of texts. The rules and customs for attributing authors to texts are closely tied up with the practices and institutions under which

---

12. Roland Barthes, “The Death of the Author,” in *Image, Music, Text* trans. Stephen Heath. (London: Fontana Press, 1977).
13. Barthes, 142-3.
14. Barthes, 147.
15. Barthes, 148.

particular people actually compose those texts. In an intellectual history that elaborates upon the Foucauldian story, Martha Woodmansee argues that the Romantic conception of authorship as aligned with individual inspiration emerged alongside a new kind of literary professional, a generation of full-time composers of prose attempting to earn income off of market sales rather than rich patrons. For Woodmansee, legal and aesthetic theorists of authorship in the eighteenth century emphasized inspiration over work or craftsmanship. As she puts it:

> …minimized the element of craftsmanship (in some instances they simply discarded it) in favor of the element of inspiration, and they internalized the source of that inspiration. That is, inspiration came to be regarded as emanating not from outside or above, but from within the writer himself. 'Inspiration' came to be explicated in terms of original genius, with the consequence that the inspired work was made peculiarly and distinctively the product—and the property—of the writer.[16]

The insistence that a text cannot be understood independently of the inspiration of its author corresponds to a particular social and legal position for professional writers, as self-employed artisans who answer to nobody except their own internal muse, and whose self-employment depends in part on the legal title they hold over their creations. This model of textual composition, in which legally-independent writer-owners compose original texts and then sell the rights to reproduce that text to a publisher, lurks in the background of the concept of the "author," even when it doesn't very accurately describe the origins of a particular text. In fact, we sometimes neglect to even use the word "author" when describing a person who composes texts at the direction of somebody else, as in the case of ghostwriters, writers of advertising copy, technical writers who compose instructional manuals, and so on. The typical terms for such people are not "ghost authors" or "technical authors."

---

16. Martha Woodmansee, "The Genius and the Copyright: Economic and Legal Conditions of the Emergence of the 'Author,'" *Eighteenth-Century Studies* 17, no. 4, Special Issue: "The Printed Word in the Eighteenth Century" (Summer, 1984): 425-448

This distinction has an explicit legal basis: in the case of employed writers who compose texts in the course of their employment, American copyright law explicitly designates the "author" of those texts to be the employer rather than the employee, regardless of the actual division of labor that went into the composition of the text. As Catherine Fisk explains, this "work for hire doctrine" was elaborated in American jurisprudence in the final decades of the nineteenth century and first decade of the twentieth century.[17] Before 1860, American copyright law generally assigned ownership to the actual human person who composed a text, regardless of whether that person composed it in their capacity as an employee of another. Fisk generally refers to this model as "Romantic authorship," and it is closely related to the sense of "authorship" imagined by critics of the concept like Barthes and Foucault. By the turn of the century, however, this legal rule became reversed: courts assigned authorship to employers by default, interpreting the forfeiture of authorship rights as implicit in employment contracts. As Fisk notes, the first case that explicitly held that a corporation owned a copyright in the creations of its employees actually appealed to the Romantic conception of authorship in order to make its argument. In a case concerning the design of an advertising woodcut that had been designed by the corporation's president but painted by its employee, the court ruled that the president's role as designer entitled him to copyright protection. As the court explained, "The sculptor seldom touches the marble from which his statues are carved. The fact that the brush which embodied [the president's] idea was held by another artist rather than by himself cannot be important in considering a question of this character."[18] In other words, authorship flows from conception rather than execution, idea rather than material, content rather than form. This attitude was made

17. Catherine Fisk, "Authors at Work: The Origins of the Work-for-Hire Doctrine," *Yale Journal of Law & the Humanities* 15: 1 (2003): 1-70.

18. Fisk, "Authors at Work," 56.

national law by The Copyright Act of 1909, which included explicit provisions for artworks made for hire.

In a similar vein, Peter Jaszi suggests that the work for hire doctrine is not just a legal fiction but is the "logical (if perverse) working out of the underlying assumption that the essence of 'authorship' lies in original, inspired creative genius."[19] In other words, since the ultimate ground of Romantic authorship is inspiration, that is, the act of conceiving an original idea, it is not particularly important whether authors do any writing themselves or hire somebody else to do writing for them. Supervising or directing the creation of a text is explicitly a kind of authorship. As Jaszi explains, legal decisions on this topic maintain that the key distinction for authorship is the "motivating factor" of the employer who instigates the project (by hiring someone to complete it), rather than the labor of a specific individual. [20] The employer's "motivating factor" is the inheritor of Wordsworth's inspiration.[21] Jaszi's account complicates that of Fisk. Responding to Barthes, Fisk writes that "[r]umors of the author's demise were… premature. The author isn't dead; he just got a job."[22] More precisely, we ought to say that it was the writer who got a job. The author, by contrast, became the owner of an authoring corporation, one which may employ writers.

In the early twentieth century, the "work for hire" model of authorship was widespread in those fields of cultural production characterized by corporate production: film, pulp magazine fiction, radio plays, and television, or all those fields that C. Wright Mills would call "the

---

19. Peter Jaszi, "On the Author Effect: Contemporary Copyright and Collective Creativity," *Cardazo Arts & Entertainment Law Journal* 10 (1991-1992): 34.

20. Jaszi, 34.

21. Note also that this conception of authorship is more obvious in the case of film: we call a director an *auteur* if the film is seen as the expression of that director's personal vision, regardless of the actual division of labor that went into its production.

22. Fisk, "Authors at Work," 1.

cultural apparatus."[23] For Mills, the kinds of for-hire cultural production typical of state and corporate bureaucracies represented a decisive break from the classically-modern figure of the independent cultural producer. As he tells the history, "most of our inherited images of 'the intellectual' and of 'the artist'" correspond to a period of cultural history in which "the cultural workman becomes an entrepreneur" selling a commodity to "the bourgeois public."[24] This institutional situation formed the basis for a vision of artistic and intellectual work that still persists, that of "the inherently and necessarily free man, and the cherished and heroic notion of the advance-guard." Put together, "[t]his notion, one might say, is 'the myth' of the intellectual." But although this vision of artistic independence is a "myth," it is a myth grounded in a particular social reality, characterized by mostly self-employed artists, intellectuals, and writers.[25]

Mills contends that this model of cultural production was being increasingly supplanted by that of the "cultural apparatus." Crucial to Mills' sense of the shift is the subordination of artists to the instructions of their boss, which transforms both the self-understanding and practice of those artists: "art, science and learning are brought into subordinate relation to the dominant institutions of capitalist economy and nationalist state," creating a situation in which "[t]he cultural workman has little control over the means of distribution of which he becomes a part."[26] In this new milieu, there are essentially two paths for cultural work: you can become either a "commercial hack" or a "commercial star." By "hack," Mills seems to simply mean an employee, or anybody who produces anonymously for the cultural apparatus. The "star," by

---

23. C. Wright Mills, "The Cultural Apparatus," in *Power, Politics, and People: The Collected Essays of C. Wright Mills*, ed. Irving Louis Horowitz (Oxford: Oxford University Press, 1963).
24. Mills, 411-12.
25. Mills, 411-12.
26. Mills, 418-419.

contrast, is “a person whose productions are so much in demand that, to some extent at least, he is able to use distributors as his adjuncts.” The structure of the cultural marketplace tends to polarize cultural workers into one of these two categories and “kill off” any hope of becoming “a worthy and independent craftsman.” Cultural production becomes an all-or-nothing competition: “[o]ne is a smash-hit or one is among the failures; one is a best-seller or one is among the hacks and the failures; one is either absolutely tops or one is just nothing at all.”[27]

It is not entirely clear why Mills sees the forms of authorship characteristic of the cultural apparatus as intrinsically opposed to the older bourgeois model of the cultural workman as “entrepreneur.” Following Jaszi, we could see authorship within the cultural apparatus as continuous with the older bourgeois model of authorship, insofar as they both take conception as the ultimate grounds for authorship as a legal and social right. From this perspective, the cultural apparatus appears less as a radical break from the bourgeois model than the continued extension of the division of labor. This point of view helps draw out the fact that even some of the most corporatist forms of cultural production often depend on the figure of the author both as an ideology and as a practical social relationship. The Walt Disney Corporation is a classic example. From its inception, animation was a highly collaborative form of cultural production that combined the labor of artists, writers, and technicians. But Walt Disney took the lion’s share of income and credit derived from corporate productions. In fact, this was a source of tension within the studio in its early history. Disney’s supporters saw no problem with a method of cultural production in which only the proprietor or director takes credit; one art historian even analogized the situation to the classical workshop system whereby craftsmen worked as the anonymous assistants for the master of the workshop. But many of the cartoonists themselves

27. Mills, 418-419.

were dissatisfied with their anonymity. Securing credit rights was a major demand of the 1941 Screen Cartoonists' Guild strike at the studio.[28]

Legally, then, it should be clear that the boundary between the classical bourgeois or Romantic model of authorship and the modern corporate model of authorship is fuzzier than one might imagine. Insofar as the basis of authorship rights have been traditionally grounded in a desire to credit the conceivers of original ideas, the corporate author and the author who works alone are not so different. Legally, at least, authorship flows from the mind rather than from the pen. With this established, we are now in a position to examine authorship from a political-economic point of view.

2.2. Authorship and the Theory of Value

In recent years, a burgeoning discourse within Marxian cultural theory has attempted to account for cultural production from the point of view of Marx's critique of value. A prominent voice in this discourse is Dave Beech, whose book *Art and Value* challenges much of the common wisdom about the production of art in capitalist societies.[29] As Beech notes, contemporary discourse about the absorption of art by capitalism—whether condemnatory or triumphalist in tone—fails to address fundamentally economic questions about whether art is capitalist at the level of production. Focusing on the case of the visual arts, Beech notes that artistic production lacks many traits that are considered characteristic of capitalist production:

---

28. For a history and discussion of the strike, see Michael Denning and Holly Allen, "'Who's Afraid of Big Bad Walt?' Disney's Radical Cartoonists," in Michael Denning, *The Cultural Front: The Laboring of American Culture in the Twentieth Century* (New York: Verso, 2010): 403-422.

29. Dave Beech, *Art and Value: Art's Economic Exceptionalism in Classical, Neoclassical and Marxist Economics* (Chicago: Haymarket Books, 2016).

artists are generally self-employed rather than wage laborers; art does not typically experience productivity-enhancing technical change; and, most importantly, social and cultural norms regarding the uniqueness of individual works of art prevent the operation of what Marx called "the law of value."

This last element is a core aspect of Beech's argument and is worth dwelling upon more closely. As Beech notes, for both Marx and the classical political economists whose work he critiqued, works of art are economically unusual because they are fundamentally unique. Many commodities—such as Marx's linen coats—tend to have their prices regulated by their cost of reproduction by virtue of the fact that competing producers can offer similar commodities for a competing price. If one shopkeeper's linen coat is exorbitantly priced, it is easy enough for me to walk down the street and purchase another one for a price that I find more reasonable. But this price competition is only possible if I am indifferent to the different qualities and origins of the two coats; that is, they both must supply the particular use-values that I am interested in, such as warmth, style, comfort, and so on. Furthermore, I must be indifferent to its physical origins, to the practical process by which it was produced and brought to market. In practice, this is generally the case; in fact, I probably have no idea whose labor was supplied to create the coat, the conditions under which that person labored, and so on. This is a key aspect of what Marx famously calls "abstract labor." Mass-produced commodities obscure the concrete and specific conditions of their production; we treat them as if the labor required for their reproduction were all the same.

But as Beech points out, this aspect of commodities is generally absent from the arts; or, at least, from works of visual or plastic arts, which are organized around powerful social and institutional norms regarding their uniqueness. Unlike a linen coat, the value and significance of a work of art bears an intrinsic connection to the specific human person who created it, or who is held to have created it. No matter how convincing or beautiful a copy of the *Mona Lisa* may be, it will never sell for as much as the original. Original paintings and sculptures sell for millions of dollars because they are originals, that is, because they are held to be the unique creations of famous and respected artists. Importantly, this quality is a social rather than a technical limitation. There is no purely technical reason why an individual work of art cannot be reproduced capitalistically, according to machine reproduction, a technical division of labor, wage labor, and so on. It is perfectly possible to mass produce reproductions of Van Gogh's paintings or copies of Cindy Sherman's photographs; many museums will sell reproductions of this type at the gift shop. But these copies will command a considerably lower price than the original, which by definition cannot be reproduced. This high price is sustained by the collective belief—propped up by tradition and institutional authority—that establishes the work's originality, authenticity, and importance. Economically, original works are one of the few cases of a true monopoly: since their supply is absolutely fixed, their price is not influenced in any way by their cost of production and is instead determined only by the amount of money that buyers are willing to pay. This fact also means that the prices of works of art bear little to no relationship to their cost of production, which was a key aspect of Marx's theory of value.[30] Works of art that cost comparable amounts in terms of materials and labor might have totally divergent selling costs, based on nothing other than the reputation and prestige of the artists.

---

30. Though Marx does not theorize the political economy of art directly, he does sometimes specify that works of art represent important exceptions to the law of value as he understands it. Consider, for instance: "Leaving aside actual artistic works, which are excluded from our subject by the very nature of the case…"

Karl Marx, *Capital*, Vol 3, trans. David Fernbach (London: Penguin Books, 1991): 893. Subsequent citations refer to this edition unless specified.

Beech's book is an important contribution to contemporary Marxist theory about art. In particular, it deserves credit for helping to revive interest among Marxist cultural studies in the Marxian theory of value. But much of his theory of the "exceptionalism" of art is restricted to the fairly specific case of the visual and plastic arts, especially those art-forms that are considered important by museums and collectors. His intervention begs the question of how we are to make sense of those cultural industries in which social norms about uniqueness do not apply, such as publishing and film. A book, for instance, is clearly a mass-produced commodity, one that may seem economically more similar to a linen coat than to an original painting. The production of books is unambiguously capitalist: it takes place in factories and is performed by printers and bookbinders who are paid in wages; advances in printing technology have steadily reduced the amount of printer labor required to produce an individual book; and the possibility of competition between booksellers creates a powerful incentive for individual book producers to keep their costs low in order to compete on price. At the same time, it does seem like some of the norms of aesthetic originality remain important for the economy of book production. Readers develop powerful attachments to individual authors, authors whose works are fundamentally distinct from each other. We might imagine that a reader would find it harder to substitute, say, a Danielle Steel novel for a Thomas Pynchon novel than that same reader would find it to substitute one brand of soap for another. Or, to use a more concrete example, sales of *The Cuckoo's Calling* skyrocketed after it was revealed that "Robert Galbraith" is a pseudonym of J.K. Rowling.[31] In the spirit of Beech's critique, what would it mean to revisit the culture industry in light of the Marxian theory of value?

31. See Maev Kennedy, "Lawyer who uncovered JK Rowling's Robert Galbraith Alter Ego fined £1,000," *The Guardian*, December 31, 2013.

Here it is useful to consider the distinction between writing and authorship from an economic perspective; in particular, to introduce distinction between, on the one hand, the composition of texts, and on the other hand, the appropriation of the income derived from the sale of physical copies of those texts, whether books or streams of electronic bits. Importantly, texts and books are not the same thing. Textual production actually has two distinct moments: composition, in which a new assemblage of words is created; and copying, or the actual production and distribution of the physical books or digital files that bear those words. After the process of composition and editing are completed, the writer is no longer needed in any way for the production of the actual commodities that will be sold for money (though they might, of course, still be very important in terms of marketing and publicity). The particular assemblage of words that make up their finished draft can be sent to the printer or copied digitally, all without the participation of the writer. In other words, a finished text exists in the world entirely independently of its author. Whether the publisher decides upon a print run of ten thousand copies or ten million, the author has worked the same amount. By contrast, the production of additional physical books requires additional expenses of living labor: labor must be performed by lumberjacks, paper-pulpers, printers, bookbinders, warehouse staff, delivery drivers, bookstore personnel, and so on. There is a radical disjuncture between the composition of texts and the production and sale of the particular saleable commodities on which those texts are printed, whether codices, pamphlets, or PDF files. Authors are irrelevant to this latter process: after a text has been composed, it is available for copying by anybody with the technical capacity and legal right to do so. The author is dead, as Barthes would have it; indeed, even the author's literal death presents no obstacle to the actual production of copies.

But authors, of course, receive income from the sale of copies of texts they have authored, despite the fact that they themselves are socially and technically irrelevant to the actual process of producing copies of those texts. Why is this? The reason is a legal one: authors often hold a special kind of property title, a copyright, that gives them a legal monopoly over the commercial copying of texts that they are held to have authored. Importantly, ownership of this property title is distinct from the question of how the text itself was brought into being. There are many different types of social relations in which textual composition may take place. A writer may engage in their writing practice as an anonymous hack writer, receiving payment by the word or for completed manuscripts. They may work as a comparatively high-status professional, receiving a guaranteed salary in return for continued compositional labor. They may write simply for the honor or respect of being a writer, living off income from grants or unrelated employment. Or, they may write with the express intention of acquiring new copyrights that they can then lease to publishers in return for income. None of these relations makes one iota of difference in understanding the nature of the income derived from a copyright (though they are essential for understanding how ownership of that copyright is struggled over and determined). No matter how somebody came about owning a copyright, the income that copyright allows the owner to extract from the sale of new copies of an already-composed text is always a property income, something akin to the rental income that the owner of a parcel of land can charge for the use of that land. This is true regardless of the relationship that the owner of that copyright bears to its author. Once a text exists, then the question of who created it and under what conditions they did so is largely irrelevant to the question of who is able to derive income from it. It is true that the owner of this monopoly right is often also the same person as the originator of the text. But this is only because of particular regimes of copyright law that have historically given a privileged place to authors as conceivers of new ideas. This tight association between originator and proprietor needn't be the case, and indeed, often is not: copyrights may be assigned jointly, may be transferred in whole or in part, may be inherited, or may be assigned to nobody at all, as in the case of texts in the public domain.

This last case is particularly instructive and is worth dwelling upon more closely. In the situation of texts entirely in the public domain, the capitalist publisher doesn't have to spend a penny on composition. He confronts the text as an instance of what Marx called a "free gift of nature," something simply available for use in production without an expense of labor, like the air we breathe or the knowledge of the Pythagorean theorem.[32] Publishing companies might still profitably sell new editions of public domain works, whether copies of the King James Bible or the Penguin edition of *Capital*, Volume 1. They must still pay for paper, printing, distribution, storage, advertising, and so on, but they do not have to pay a licensing fee to an owner of an intellectual property right, much as a construction company need not pay for the use of scientific knowledge by its engineers in the absence of an enforceable patent. Importantly, free gifts of nature have no "value" in the technical, Marxian sense of that word, since by definition "value" is an attribute of commodities produced by human labor. Marx is explicit that anything made use of by capitalists for production that does not require an additional expense of labor has no labor

32. For the phrase "free gift of nature," see Karl Marx, *Capital*, Vol 3, ed. Friedrich Engels, trans. Institute of Marxism-Leninism (International Publishers: New York, NY [n.d]): 548. https://www.marxists.org/archive/marx/works/download/pdf/Capital-Volume-III.pdf.

Note that not all translators choose this phrasing. David Fernbach instead translates the same passage as "free natural power of capital." See Karl Marx, *Capital* Vol 3, ed. Friedrich Engels, trans. David Fernbach (Penguin Books: New York, NY: 1991): 879.

The concept of a "free gift of nature" has a long and tenuous history in Marxian theory. For an excellent elaboration of how the concept fits into Marx's own framework, see Paul Burkett, "Nature's 'Free Gifts' and the Ecological Significance of Value," *Capital and Class*, Vol 23 (1999): 89-110.

value: "Natural agents as such cost nothing. They cannot, therefore, add any value to the product…"[33] In a similar way, freely-appropriable social knowledge costs capital nothing, as it requires no additional expense of labor: "Once discovered, the law of the deflection of a magnetic needle in the field of an electric current, or the law of the magnetization of iron by electricity, cost absolutely nothing."[34] This costless quality is no barrier to its appropriation by capital: "Science, generally speaking, costs the capitalist nothing, a fact that by no means prevents him from exploiting it."[35] Capitalists readily make use of scientific discoveries that cost nothing after their discovery.

Public-domain texts, then, have no value. But in Marx's framework, something can have a price without having a value.[36] This is the case for anything that is not produced by means of capitalist production but that nevertheless bears a price tag; the classic example is uncultivated land monopolized by a property right.[37] Though there is no human labor "objectified" in such land, as Marx puts it, the property title held by its owner nevertheless allows that person to extract a monetary payment for its use. This sum of money Marx refers to as a "rent." Though rents are most characteristic of landowners, Marx suggests that rents emerge any time "capital comes up against an alien power that it can overcome only partly or not at all, a power which restricts its investment in particular spheres of production."[38]

---

33. Karl Marx, *Theories of Surplus Value*, Part 2 (Progress Publishers: Moscow, 1968). Cited in Burkett, 93.

34. Karl Marx, *Capital: A Critique of Political Economy*, Vol 1, trans. Ben Fowkes (New York: Penguin Books, 1990): 508. Subsequent citations refer to this volume unless specified.

35. Kennedy.

36. : "...a thing can, formally speaking, have a price without having a value." Marx, *Capital* Vol 1, 197.

37. "...as for instance uncultivated land, which is without value because no human labour is objectified in it." Marx, *Capital* Vol 1, 197.

38. Marx, *Capital* Vol 3, 896.

This discussion of valueless "gifts of nature" that are nevertheless rented out for a monetary price helps to clarify a well-known remark about Milton that Marx made in his unpublished manuscripts of 1863-4, later included as an appendix to *Capital* Vol. 1. Since the passage has been so frequently cited and debated by Marxist literary critics, it is worth quoting in full. Reflecting on the distinction between productive and unproductive labor in political economy, Marx writes:

> For instance, Milton, who wrote *Paradise Lost*, was an unproductive worker. On the other hand, a writer who turns out work for his publisher in factory style is a productive worker. Milton produced *Paradise Lost* as a silkworm produces silk, as the activation of his own nature. He later sold his product for £5 and thus became a merchant. But the literary proletarian of Leipzig who produces books, such as compendia on political economy, at the behest of his publisher is pretty nearly a productive worker since his production is taken over by capital and only occurs in order to increase it.[39]

Literary critics have often dismissed this passage, frustrated either with the bizarreness of the comparison or with the apparent idealism of Marx's description of literary production as an "activation" of a writer's own "nature."[40] But discussions of the idealism of the silkworm comparison gloss over what seems to me a much simpler and more analytically-powerful interpretation of this passage: that, in economic terms, Milton's text is akin to the free gifts of nature that Marx discusses elsewhere, something with no labor value but to which somebody could still attach a price tag (and in doing so become a merchant). This interpretation is supported by the fact that in Marx's framework, the productivity of silkworms falls pretty clearly into the category of nature's free gifts. Capital pays for the silkworms themselves, the human labor necessary for breeding and caring, and the human labor necessary for harvesting and

---

39. Marx, *Capital* Vol 1, 1044.

40. For a discussion of the reception of this passage by literary critics, as well as a defense of it, see Kathryn Crim, "Marx, *Silk Poems*, and the Pretext of Qualities," *Representations* 151 (2020): 96-126.

processing that silk. But it does not pay for the capacity of silkworms to produce silk. Like other natural conditions, capital appropriates this quality without expending labor: "So also physical forces, like steam, water, &c., when appropriated to productive processes, cost nothing."[41] If silkworms were legal subjects who participated in market transactions, they too might, like Milton, sell their silk and become merchants.[42] The point is simply that capital encounters these qualities—silk productivity, on the one hand; the text of *Paradise Lost*, on the other—as something already available for use in production without an additional expenditure of labor.

If texts, once they exist, are in some important sense "outside" of value, this implies that understanding the author-function requires theorizing capitalist authorship as something distinct from labor as Marx uses that term. Though the production of physical books might be reasonably describable in value-theoretic terms, texts themselves do not have value in the technical Marxist sense. Nevertheless, authors and texts still participate in the capitalist economy, and a copyright might enable its owner to capture value produced elsewhere. In what follows, I examine three cases of authorship arrangements in which the separation between the author-function and the actual work of writing are crucial for making sense of the social relations between the various actors involved in the production and distribution of the text. As I show, in all cases, the "author" is clearly somebody who inhabits an ownership or managerial function. First, I consider "book packagers," or independent editorial companies that contract with writers (and other workers,

41. Marx, *Capital* Vol 1, 508.

42. Note that this interpretation also avoids any accusations of Marx's failure to appreciate the social quality of literature, since Marx also describes capital as freely appropriating "natural forces of social labor" such as the division of labor. The productivity gains to be had from a division of labor are "natural" in the sense that they are an automatic or given quality of the division of labor itself; but it remains the case that *particular* divisions of labor are socially and historically contingent. For a discussion of "natural forces of social labor" in Marx, see Burkett, 93-4.

like artists) to produce works-for-hire, which are then sent to publishers for final publication. Here, the author is a manager: someone who designs a literary strategy and hires a writing employee to bring it to fruition. Next, I consider media franchises for which ghostwriters are hired to write individual works, such as novels in the *Star Wars* universe. Here, the author is to corporation that owns the intellectual property rights of the commissioned texts. Finally, I consider the case of V.C. Andrews, whose name became a sort of franchise brand after her death. Andrews is the clearest case of authorship-as-ownership: after her death, her estate inherited her author name as an intangible asset and even had to pay an estate tax on its capitalized value.

2.3. The Author as Book Packager

In the early 1980s, agent and publishing professional Bill Adler had an idea for a book. The premise was simple: a mystery novel that, rather than revealing the identity of the murderer at the end of the book, instead held a contest among readers asking them to try and solve the mystery themselves. Such an inversion of the role of writer and reader is already implicit in the puzzle-like, highly-participatory quality of the genre itself, in which narrative pleasure is derived in part from playing mental detective alongside the protagonist. But Adler's idea would take this implicit game-like structure and make it explicit, pairing it with a publicized competition and prize money for the reader who could most accurately solve the mystery therein. Adler contacted author Thomas Chastain—a writer of mysteries who would later become president of the Mystery Writers of America—and pitched him the idea, asking him to write the book. The result of their collaboration was the 1983 novel *Who Killed The Robins Family?*, published by William

Morrow.[43] Each chapter of this meandering whodunit describes the murder or disappearance of a different member of the titular Robins, a wealthy eight-person family that owns and operates a cosmetics business. Their murders take place in variously-implausible locales, many of which are selected in homage to canonical stories of the genre: the Orient express, a gothic mansion, a cruise ship. Marketing for the novel emphasized the prize money most of all: $10,000 to the reader or readers who could produce the most accurate answers to the eight murders or disappearances. Adler received over 25,000 submissions, with the prize money eventually going to a group of friends in Denver—members of an informal group of mystery lovers—who solved the mystery collectively. The marketing gimmick worked; the cost in prize money was paid off by sales many times over, with the book topping the *New York Times* bestseller list in January 1984 and taking a lower position on the list for months after.[44]

Prior to his death in 2014, Bill Adler was one of the more famous examples of what is called a "book packager" or "book producer" in the publishing industry, an independent literary producer that coordinates the labor of multiple creative workers, which may include writers, illustrators, researchers, and even editors, in the production of a finished manuscript that is then sold to a different company for its actual publication. Historically, book packagers have been used most often for nonfiction books, especially commodity books that require any kind of complicated expert labor above and beyond that of writing, such as photography or research. As the American Book Producers Association, a trade association for book producers founded in 1980, explains on its website,

43. Bill Adler and Thomas Chastain, *Who Killed the Robins Family?* (New York: William Morrow, 1983).

44. Weekly bestseller lists are available in the data used in Chapter 1. See Jordan Pruett, "NYT Hardcover Fiction Bestsellers," *Post45 Data Collective*, 2022. https://doi.org/10.18737/CNJV1733p4520220211.

> Book producers, also known as packagers, are akin to independent producers in film and television… They are some of the most valuable players on the creative side of the publishing industry today. They have earned this distinction for their mastery of one type of book in particular—namely, the 'complicated' book, which is just about any book that involves more than a straightforward, single-author text. Complicated books include all manner of highly illustrated, elaborately designed, or multi-authored titles, such as how-to books, coffee table books, reference books, textbooks, cookbooks, and more.[45]

The complexity of the "complicated book" lies in the logistics of its production: a cookbook, for example, requires input from a number of different professionals in order to be ready for publication, who may include photographers, visual designers, chefs or culinary experts, writers, editors, and researchers. A single-author novel, by contrast, may be conceptually or stylistically complex, but in general its composition is logistically simple; one person writes it largely alone, though perhaps in collaboration with an editor or research assistant.

For this reason, book packagers have less often been involved in the production of narrative fiction, with some notable exceptions, such as illustrated children's literature. In this regard, Adler's career in publishing was a precocious one, at the vanguard of a more general mainstreaming of co-authored or ghostwritten bestselling novels. One of his early successes was *Kids' Letters to President Kennedy*, a 1962 compilation of letters written to the president by American children.[46] The cover lists no author for the book, but describes the letters as having been "selected by" Bill Adler. Adler, then, is closer to a producer in film than either an author or an editor; he is an administrator who brings projects into being. In his career, we can see the influence on literary production that was exerted by the film and media industries during the

45. American Book Producers Association, "About Us." https://abpaonline.org/about-us/

46.Allen Appel, "Remembering Bill Adler Sr.," *Publishers Weekly*, April 4, 2014. https://www.publishersweekly.com/pw/by-topic/columns-and-blogs/soapbox/article/61732-requiem-for-a-book-packager.html early 1980s, industries with which publishers had become increasingly entangled in the previous decade.[47]

Given this institutional entanglement, it is unsurprising that many of Adler's creations adopt a crediting scheme borrowed from film and television: the cover of *Robins Family*, for instance, informs us that it was "created by Bill Adler and written by Thomas Chastain." In television, the "created by" credit has a specific technical and legal meaning. It refers to the writer who retains "separated rights" over the content of a new series, or certain non-televisual rights of reproduction, which may include dramatic rights, publication rights, or merchandising rights.[48] As the Writers Guild of America West explains in a guide to separated rights for their members, writers generally receive a "created by" credit and its attendant rights for one of two reasons: by receiving a "story by" or "written by" credit for the pilot episode or by writing the "format" of a new series, which is a kind of detailed outline or premise for the series.[49] In other words, the "created by" attribution is designed to formalize credit for the labor of creating the overall premise, idea, concept, or framework of a television franchise, as well as secure for the

47. For a discussion of the impact that film had on publishing during this period, see Thomas Whiteside, *The Blockbuster Complex: Conglomerates, Show Business, and Book Publishing* (Middletown, CT: Wesleyan University Press, 1981).

48. See Writers Guild of America West, "Separation of Rights," in *Television Credits Manual* [n.d.]: 26-30. https://www.wga.org/uploadedfiles/credits/manuals/tvcredits_manual20.pdf

49. Writers Guild of America, *Understanding Separated Rights* (2000): 35. https://www.wga.org/uploadedfiles/know_your_rights/SeparatedRights.pdf

The full definition of a "format" is a written description of the "central framework within which the central running characters will operate and which framework is intended to be repeated in each episode; the setting, theme, premise, or general storyline of the proposed serial or episodic series; and the central running characters which are distinct and identifiable, including detailed characterizations and the interplay of such characters. It may also include one or more suggested story lines for individual episodes." *Understanding Separated Rights*, 27.

writer the privilege to reproduce the franchise's fictional world in another medium. It makes it possible to explicitly mark a distinction between writing as conception and writing as textual composition: a writer needn't necessarily write any actual lines of dialogue in order to be marked as the creator of the character who speaks that dialogue. In literary history, we are used to collapsing conception and composition into a single concept, that of "authorship." But in film and television, these functions are frequently completed by different people entirely and there exists a fully-elaborated institutional procedure for marking the various duties and rights held by those that fulfill each function.

Although we cannot ever actually know for certain the degree to which the division of labor between Adler and his co-writers resembled the one used in television, at minimum it is clear that he generally first came up with an idea and only later contracted a writer to execute that idea in textual form. The "created by"" credit is thus at least used analogously to the way it is used in television. This division of labor helps explain why so many of Adler's creations have a highly schematic or abstract structure—like *Robins Family*, they have a hypertrophied sense of "premise," as though their entire plot has been mapped out on graph paper in advance. Consider the 1987 novel *The Agent*, created by Adler and written by novelist David R. Slavitt, which is remarkable both for its highly conceptual structure and for the fact that it explicitly thematizes the shifting institutional terrain of authorship in the publishing industry of the 1980s.[50] A conspiracy novel about the publishing world, *The Agent* tells the story of Leonard Castle, one of New York's most powerful literary agents, as he paranoically roams the city's streets avoiding a man who may or may not have been hired to assassinate him. The novel has a modular, disaggregated structure. Late one night, after hearing a series of loud pops that he interprets as

50. Bill Adler and David R. Slavitt, *The Agent* (New York: Doubleday, 1986).

gunshots, Castle hops into a cab and sets off, on the run from an unknown assailant. While fleeing, he desperately tries to "solve" the case of the attempted murder that he has just narrowly avoided, thinking back through his most controversial and emotionally-charged book deals in a mental search for who might want him dead. Each chapter details one such remembered episode: a trip to the secluded Wyoming ranch of Bart Blackpool, an almost parodically-masculine movie star; negotiating the memoirs of Amy Breckenridge, a senator's wife whose tell-all would threaten her husband's career; a project with Mihail Szabo, an Eastern European director who became publicly hated after it was revealed he attempted to begin a sexual relationship with an underage girl; and several more. The novel affords interpretations as a *roman a clef*: Szabo is clearly inspired by Roman Polanski, and Harlan Farnsworth, a disgraced ex-president, is closely reminiscent of Richard Nixon. At the end of each episode, the story jumps back to the narrative present as Castle attempts to distill insights into potential motives for murder from the passage we have just read.

As a work of popular fiction, the novel was less successful than *Robins Family*, never becoming a bestseller itself; this comparative lack of success may have been due in part to its meandering, disorganized structure, or maybe for its over-indulgent enthusiasm for the life of the literary agent, a topic perhaps less interesting to readers who do not themselves work in the publishing industry. For our purposes, it is most interesting as a historical document, a sort of meta-reflection on the interpenetration of managerial and creative labor in the publishing industry of the 1980s. What role does the literary agent play in this book? While he never himself writes a line of the text that appears in the books he manages, he is closely involved in both the generation of new project ideas and the managerial work of ensuring that his authors actually complete the book projects they have been hired to produce. Most of the vignettes that make up the narrative meat of the novel revolve around Castle trying to convince this or that potential writer to take on a new book project, projects generally conceived of by Castle or own of his associates in the publishing industry. For example, after deducing that actor Bart Blackbool is a committed naturalist, he convinces him to co-author a coffee-table book of nature photographs with an environmentalist theme, to be released in time for Christmas. As a character, Castle makes things happen; he is somebody with a knack for convincing other people to complete projects. As he muses later in the book, he "was only an agent, a facilitator, a helper along of people and projects."[51]

His position as facilitator introduces a crisis of agency: he is only a facilitator, not the direct creator of the narratives that he commissions. This crisis of agency comes to a head in the book's conclusion, when it is revealed that the person who has attempted to shoot Castle was Dickie Byrd, a well-known talk show host. Castle had convinced Byrd to allow a ghostwriter to write his memoir—and, more importantly, to allow that ghostwriter to go public with details about Byrd's homosexuality. Byrd's hesitation towards coming out stemmed from his career and from his insistence that the shock of the revelation might kill his mother. In the end, Castle was able to convince him, in part by insisting that the world had changed and in part by offering him a $1 million advance. But Byrd's fears turned out to have been well-placed; shortly after the book's publication, his mother is found dead of a heart attack in her apartment, with the book sitting in her lap. Furious, Byrd decides to kill the ghostwriter, Pete McCracken, then settles on killing Castle instead. His reasoning for the decision turns on the question of where to place agency for the creation of the book. McCracken may have written it, but it was Castle who was ultimately culpable: "McCracken might have been the Doberman, but Castle had been the

51. Adler and Slavitt, 232.

handler, the one to unleash the animal and set it on Byrd's trail."[52] Though Castle is able to convince Byrd to spare his life in the novel's climax, a coda reveals that Byrd beat to death the man who first revealed his secret to Castle and the ghostwriter, before taking his own life. Overall, the ending spirals apart rather than resolves. In part, this is a function of the scattered and disorganized structure of the novel, perhaps a side-effect of its co-authored composition. But the lack of satisfactory resolution is also symptomatic of its theme and its protagonist. Culpability is dubious for the literary agent, who generates texts but does not directly produce them. Their responsibility for the effects of those texts becomes ambiguous. If, as we saw in our earlier discussion of Foucault, authorship emerged historically as a way of regulating textual culpability, then it makes a certain amount of sense that the unraveling of authorship in this book should present a crisis of agency.

2.4. Franchise Fiction

In the world of book packaging, Bill Adler's fictional creations are largely outliers. Book packages only rarely produce novels and narrative fiction, though they are somewhat more prominent if children's literature is included in that category. In part, the dearth of packaged novels is simply a function of the bargaining power of authors. If you, as an author, are perfectly capable of writing a novel that will sell, why would you agree to do so as the paid employee of a book packager, rather than sell the novel to a publisher directly? In the 1980s, book packagers frequently asked for large royalty splits with their writers, often in the range of a full fifty-fifty split. But as book packager Sandra Choron put it in a conversation with *Publisher's Weekly*, a

52. Adler and Slavitt, 238.

split that steep can only be justified to a writer "...when a book is highly illustrated and where there is special editing, fact-checking, and photography… if a packager is getting 50% simply for the idea, then it's just highway robbery."[53] Working alone, a professional writer is simply incapable of producing a manuscript for an illustrated coffee-table book ready to be sent to a publisher. The coffee-table book requires a division of labor by its nature. A novel is a different matter, insofar as the bulk of fiction writing happens alone.

In the case of entirely new novels, an idea by itself does not count for much. But ideas can be worth quite a lot if they are part of an established narrative franchise, one that comes with a readymade audience and pre-existing publicity. Extremely popular franchises have an audience that demands a particular type of story that writers cannot legally go reproduce on their own, which makes it easier for the owners of those franchises to persuade a writer to write the next entry on a work-for-hire, contractual basis. As such, it is perhaps unsurprising that many novels written by writers who do not retain sole authorship rights—and virtually all novels in this category that make it onto the bestseller lists—are part of an existing narrative franchise. In this section, I will consider one type of what might be called "franchise fiction." I consider the case of novels associated with multimedia franchises, which developed from the movie tie-in books of the 1970s and became an institutionalized strategy of many publishers in the 1980s and 1990s. These novels were often commissioned by larger entertainment organizations, which generally retained most or all of the stake in the intellectual property underlying the novel's universe and characters.

53. Charles Salzberg, "Book Packaging and the Big Time," *Publishers Weekly*, August 9, 1985: 34.

Though Hollywood and the publishing industry have been involved with each other for virtually the entire lifetime of the former, for most of the history of the American film industry, narratives often flowed downstream from publishing to film. Early film borrowed liberally from literary source material for a large number of screenplays. To give a sense of the norms, according to one estimate, even by the 1930s, the British film industry used literary source material rather than original screenplays for about half of all films.[54] Although comparable statistics do not exist for Hollywood to my knowledge, this number shows that literary adaptations were commonplace at the time. As is well-known, writers like William Faulkner frequently dabbled in screenwriting to pay the bills. During this time, it was much less common for adaptation to move in the opposite direction, from film to the novel. When they were interested in publishing a novelization of a Hollywood screenplay, publishers often found that they could easily persuade movie studios to sell them publication rights on the cheap. One industry journalist reported that, before the 1960s and especially the 1970s, publishers could buy book rights to a film for an average of only $2,500, or even less. In some cases, studios simply gave book rights away for free, as they were more interested in the value of the free publicity than they were in the price of the adaptation right itself.[55] As former Bantam editor Nancy Hardin put it in 1975, "[in] those days, all the studio hoped to get out of the tie-in was the promotional value of having its film's title and ad campaign displayed in one more place before

54. Gerben Bakker, *Entertainment Industrialized: The Emergence of the International Film Industry, 1890-1940* (Cambridge: Cambridge University Press, 2008): 310.

55. Eileen Lottman, "Novelizations: Are the Plums Drying Up?" *Publisher's Weekly*, October 10, 1977: 31.

the public's eye."[56] Tie-in books functioned more as advertising for a film than as profitable commodities in their own right.

This dynamic shifted slowly over the entire postwar period, but changed particularly quickly in the 1970s, when a number of paperback adaptations of Hollywood blockbusters became bestsellers in the mass market. An early case of this phenomenon was Erich Segal's 1970 novel *Love Story*, adapted by Segal from his screenplay of the same name at the request of Paramount. Though its intended function was publicity, Segal's novelization was already a bestseller by May of 1970, even before the film was released that December. The successful release of the film only accelerated the pace of sales of the novel. By early 1971, it had over one million hardbacks in print, with more than 4.3 million copies in paperback.[57] The novel's success set off a spate of attempts at imitating the formula in the early 1970s. Some of them, such as Herman Raucher's *Summer of '42*, were successful; others, such as Marc Norman's *Oklahoma Crude*, were not able to replicate the success.[58]

In accordance with industry norms in Hollywood during the time, most contracts between screenwriters and movie studios explicitly granted novelization rights to the screenwriter.[59] This was the case with Segal, who wrote both the novel and screenplay version of his narrative. In other cases, however, a writer might work for-hire to generate a novelization for a narrative franchise over which they personally hold no rights. This was the case for one of the most successful movie spin-off novels of the 1970s: the novelization for *Star Wars*. Released at the end of 1976, a little less than a year before the release of the film, *Star Wars* became a minor

56. Nancy Hardin, "For Paperback Houses, There's No Business Like Show Tie-In Business," *Publisher's Weekly*, February 17, 1975: 47.

57. Jack Doyle, "The Love Story Saga, 1970-1977," *PopHistoryDig.com*, June 29, 2011.

58. Hardin, 48.

59. Hardin, 49.

paperback success before it ever debuted in theaters; six months after the release of the film it had sold over three million copies in paperback.[60] The original edition credits George Lucas himself as the author and displays his name on the cover. But the novel was actually composed by a ghostwriter, Alan Dean Foster, who would later go on to write a number of novels of his own set in the *Star Wars* universe. Foster was brought on for a two-book deal: the novelization of *Star Wars*, as well as a sequel, designed to supply narrative material for a low-budget film sequel if the first film ended up being unpopular. Foster's original contract was a true work-for-hire arrangement; he was to receive royalties only on the sequel, with no ownership stake in the first title. However, George Lucas decided to give him a one-half percent royalty stake in the novelization at the last minute.[61]

Foster's brief contribution to the first episode in one of the most lucrative entertainment franchises of all time is interesting for how it draws our attention to the shifting institutional norms around authorship in the 1970s. In practical terms, Foster worked very much as an employee; he had little creative autonomy of his own and was instead asked to work based off of a rough draft of the screenplay as well as concept art for the film. During interviews, Foster has frequently described work-for-hire writing as requiring a kind of ego discipline, in which the writer suspends his or her own aesthetic judgements. Describing the challenge of writing a novelization for a *Transformers* film of which he was not particularly fond, Foster explained that "[i]t's like you're hired to paint somebody's house, and you think it should be sandstone but they want it green and pink, and though you may go puke in the corner, you go puke in the corner and

---

60. Robert Dahlin, "Is There a Robot in Your Future? Yes, There Is. At least in a Harvest/HBJ Spring Paperback," *Publisher's Weekly*, December 5, 1977: 31.

61. Ryan Britt, "How George Lucas Made a Young, Anonymous Author Rich," *Inverse - Entertainment*, May 4, 2017. https://www.inverse.com/article/31165-may-the-4th-star-wars-day-alan-dean-foster-george-lucas

come back and paint it."[62] This ego discipline is framed as a professional skill, akin to the discipline required of any other profession performed at the direction of an employer, such as house-painting. In another interview, Foster similarly analogized his work with George Lucas to the work performed by a contractor. Asked if he found it difficult to see Lucas receive full credit for his novelization, he replied, "Not at all. It was George's story idea. Not having my name on the cover didn't bother me in the least. It would be akin to a contractor demanding to have his name on a Frank Lloyd Wright house."[63] He has also compared his model of work to the model of artistic patronage characteristic of Renaissance artists like Michaelangelo.[64]

Note the emphasis Foster places on the fact that the idea for the story came from George Lucas himself. The hierarchical division between conception and execution, content and style, has recurred throughout this chapter; we saw similar justifications in sites as diverse as James Patterson's statements about his work and early court decisions formalizing the procedures for assigning copyrights in artistic works produced for hire. Both legally and practically, the rights and privileges of authorship generally flow from a claim of original conception. Apparently, one needn't necessarily participate directly in the production of particular instantiations of an idea in concrete textual form in order to claim authorship over that idea. In the case of Foster and Lucas, the hierarchy of owner-originator and worker-producer neatly resembles an employment relationship. As we will see, however, another side effect of the association of authorship with ideas rather than particular texts is that it is not necessary for ownership of those ideas to be held by an actual, natural person. The author-function can operate effectively independent of an actual

62. Rob Queen, "Interview with Alan Dead Foster," *SFF World*, November 22, 2007.

63. Lionel Bently and Laura Biron, "The Author Strikes Back: Mutating Authorship in the Expanded Universe," in *Law and Creativity in the Age of the Entertainment Franchise*, ed. Kathy Bowrey and Michael Handler (Cambridge: Cambridge University Press, 2014): 44.

64. Queen.

author and may in fact outlast its author's natural life. In such a case, it becomes a brand, along with all the legal implications thereof—even, remarkably, the fact that it can be inherited as property.

2.5. The Undeath of the Author

In publishing, it has become common to refer to bestselling authors of genre fiction as "brand name" or "name brand" authors. The classic examples are the ultra-prolific authors who rode the fiction boom of the 1980s and 1990s: Stephen King, James Patterson, Tom Clancy, Danielle Steel, John Grisham. The honorific has a double-meaning. On the one hand, it evokes a sense of greater economic value: these are the "name brand" commodities, as opposed to their "generic" competitors, the low-quality imitators churned out by mass market publishers. These authors often publish in hardcover and the name on their works is a kind of guarantee of quality, as Patterson's quote at the beginning of this chapter indicates. But the phrase also refers to the fact that their name has itself become a kind of "brand," the title of a corporate entity that exceeds the individual author. This is clearest in the cases of authors whose creations have spawned derivative works and works in other media, such as Tom Clancy. The videogame series *Rainbow Six*, for instance, bears the full name *Tom Clancy's Rainbow Six*. The name "Tom Clancy" here is quite literally a brand, signaling its participation in a larger franchise, not unlike the name "Disney." This brand is less the name of a concrete, biological human person than it is a trademark, a legal fiction that connects the economic activities of disparate agents. Indeed, the death of Tom Clancy in 2013 has presented no obstacle to the continued production of "Tom Clancy" branded products.

The reason why the brand-recognition associated with a particular author name is valuable to publishers is very intuitive. Advertising expenses represent a big up-front cost: it can be very hard to get the name of a new author stuck in the minds of prospective readers, but once it's there, there is always the possibility that readers will recognize it in the future. Then, the publicity almost comes free; mega-popular authors like James Patterson have a "built-in" audience of previous readers who are happy to at least consider the latest book with his name on it. Indeed, publishers are in an unusual position among capitalist producers in that, since their brands are attached to actual human persons, they can die. Imagine the inconvenience that would befall the fast-food chain Wendy's if it ever happened that the titular Wendy died, and, for some inexplicable reason, it became impossible to produce new Wendy's-branded products in her absence. Publishers typically don't have the luxury of working with brand-names that are entirely fictional. Although their products might be fiction, the name on the cover is somebody real.

Or, at least, that's how it usually goes. Since authorial brand names are in principle separable from the person who they describe—you don't need Tom Clancy physically present to put his name on a book—publishers have sometimes had success continuing publication of an author-branded line of fiction after the literal death of the author whose name originally supplied the brand. This separability is nowhere clearer than in the case of novelist Cleo Virginia Andrews, better known by her pen name V.C. Andrews. Born in 1923, Andrews turned to writing fiction later in life, with her 1979 Gothic novel *Flowers in the Attic*. A runaway bestseller, *Flowers* is the first of what would become a five-novel series about the "Dollangangers," which follows the lives of four children—Chris, Cathy, Cory, and Carrie—as they struggle to deal with the traumas caused by their family's incestuous past. The first book begins with the death of the children's father, Christopher Sr., in a car accident, which forces her mother Corrine to take the children to live with their cruel grandmother at the manor Foxworth hall. There, the children are forced to live imprisoned in the attic, hidden from view by their grandfather, Malcolm, who is unaware they exist. Corrine reveals that the children's father was actually her half-uncle and that therefore the children are the product of incest; Malcolm disinherited the couple after learning of their relationship. Corrine promises that the children need only stay hidden for a few days while she regains her father's respect, but this promise proves empty. Days turn to months, and finally years, as the children grow sick and malnourished. After Cory dies, supposedly of pneumonia, the remaining children decide to escape. During their escape, they overhear a conversation between servants that suggests Cory was actually poisoned and that their grandfather has been dead for more than a year. Chris' discovery that Corrine's inheritance was conditional on her having no children from her previous marriage leads them to deduce that she was likely the one who poisoned Cory.

The novel was controversial upon release for its frank discussion of incest and sexual violence, including a scene in which Cathy is raped by her brother Chris, and has been banned by school libraries multiple times since its release. But despite its controversial material—or perhaps because of it—the novel achieved a level of extraordinary success unusual for a debut author. According to its publisher, Pocket Books, it sold 1.75 million copies in just the first ninety days after its release.[65] Andrews herself became a literary celebrity almost overnight. Over the next seven years, she published three additional Dollanganger novels and two in a new series, all of them bestsellers, and worked as a consultant for the first film adaptation of *Flowers*, released in 1987.

---

65. Advertisement, *Publisher's Weekly*, February 8, 1980: 32.

But at the end of 1986, Andrews died of breast cancer. Her abrupt death, less than a decade into her career as a published author, put an end to two ongoing popular series of novels with an enormous fan following. This presented a problem for Simon and Schuster, the conglomerate that owned Pocket Books. According to company president Jack Romanos, only fifty novelists in the history of American literature had ever achieved sales numbers comparable to those secured by Andrews. As he put it in a staff memorandum shortly after her death, "[s]he was one of those rare authors who brought millions of readers into the stores within weeks of each new release."[66] Unsurprisingly, then, Romanos was interested in finding a way to keep Andrews' two ongoing series alive. Anita Diamant, Andrews' agent, suggested a solution: the series could be continued by ghostwriter Andrew Neiderman, an obscure author of horror novels who Diamant also represented. At first, the Andrews estate was opposed to the idea, as they were worried about the financial risks; Neiderman's novels being rejected by fans might hurt future sales of the books that were originally authored by Andrews herself. But when Neiderman submitted an outline and several pages of ghostwritten text for a proposed prequel to *Flowers in the Attic*, the estate was impressed with imitation of Andrews' style and cautiously agreed to commission a full manuscript.

The full manuscript was published as *Garden of Shadows* in October of 1987, less than a year after Andrews' death. Any financial misgivings still harbored by the estate evaporated in the first few weeks after its publication. The novel quickly made it onto the number one spot of the weekly paperback bestseller list compiled by *Publisher's Weekly* and maintained that place for

66. Quoted in *Estate of Andrews v. United States*. The memorandum and statement of facts from this case provide most of the material for the following two paragraphs.

several weeks thereafter.[67] Impressed with the success of this first, experimental venture, the estate worked with Neiderman on a contract for an additional three novels, including a sequel to Andrews' final published novel *Dark Angel*. Both the publisher and the Andrews estate were not particularly forthright about the origins of these early continuations of the Andrews franchise. The cover to *Garden* displays the name V.C. Andrews in large letters (now branded by a tiny stamp signaling it is a trademark) and a blurb on the inner flap describes the novel as part of the same Andrews' series that has "captivated millions." The only indication of the involvement of a ghostwriter appears above the publication information in small text that a reader more interested in the story itself could easily miss: "Following the death of Virginia Andrews, the Andrews family worked with a carefully selected writer to organize and complete Virginia Andrews' stories and to create additional stories, of which this is one, inspired by her storytelling genius." The book tells us nothing, however, about the identity of this "carefully selected writer." In this first experiment with posthumous composition, Neiderman's name is nowhere present.[68]

At the level of style, *Garden of Shadows* is a convincing-enough imitation of Andrews' prose. Like Andrews, Neiderman puts language into the voice of his protagonist that sounds almost dated, as though it were itself an imitation of nineteenth-century realism. But with the benefit of hindsight, it is hard not to read some degree of subtle (or perhaps symptomatic) self-referentiality into Neiderman's rendering of the Dollanganger origin story. There is something almost too faithful in the degree to which the novel's plot beats resemble—or even entirely

---

67. See, for example, "*Publisher's Weekly* Paperback Bestsellers," *Publisher's Weekly*, January 8, 1988: 80.

68. Though I have been unable to locate any documentary evidence to this effect, one wonders if the publishers were concerned about Neiderman's gender: would Andrews' fans—mostly women—be skeptical of a male author taking over the persona and voices of Andrews' female protagonists?

reproduce—those in the source material. Olivia (the grandmother from *Flowers*) leaves her family home to start a life with Malcolm, her rich and handsome husband, at his family home, Foxworth Hall. As in *Flowers,* the Hall turns out to be a dark and frightening place; Olivia is raped by her husband at the end of only the third chapter. The middle third of the book even reproduces the "illegitimate child hidden in the attic" plot, after Malcolm fathers a child with Alicia, the young new wife of his own father Garland. On the one hand, this plot construction clearly serves the typically-Gothic function of establishing that, in this family, history always repeats itself. But it also reads as an attempt to "legitimize" a new fictional family line. As the ownership of the Dollanganger franchise has been thrown into question by the untimely death of its matriarch, the inheritor from outside the family is thrown into a position of needing to legitimize his tenuous claim to the inheritance. The new novel earns its place by imitation, quickly establishing its willingness to tread the same narrative ground: cruel and abusive parents; frank depictions of sexuality, even rape; vengeful relatives fighting over inheritance; incestuous partnerships and revelations about uncomfortable overlaps in the family tree.

In terms of production, the Andrews-Neiderman story is not particularly remarkable. The use of ghostwriters to continue the works of deceased popular writers has a storied history in the twentieth century, including such well-known franchises as *Nancy Drew* and *The Hardy Boys*. What is more remarkable is the degree to which the estate, the publishing industry, and the legal system were prepared to treat the author-attributions appearing on the front cover of books published after her death as a gesture pointing to a legal and economic relationship rather than a personal or artistic one. The name "V.C. Andrews" no longer referred to a natural person; it had become a name for an entertainment franchise. In 1994, seven years after Andrews' death, this fact was made explicit in a district court in a lawsuit between the Andrews estate and the United

States government. In *Estate of Andrews v. United States,* the Andrews estate sued the government in an attempt to secure a refund for almost $1 million worth of estate tax plus interest that had been paid to the IRS. When Andrews died, the IRS had assessed the name "V.C. Andrews" as an intangible asset of the estate similar to a copyright, and which had a value of more than $1.2 million at the time of Andrews' death. The estate countered that her name by itself could not be valued at more than $150,000. In the end, the court settled on a middle figure of $703,500. Notably, neither side denied the validity of treating the deceased's name as an intangible asset; they disagreed only on the specific procedure by which such an asset could be assigned a plausible dollar value. The court opinion makes clear that everybody involved always knew that the privilege of publishing novels under the name "V.C. Andrews" was a valuable right: "The record... is clear that the predominant reason for concealing Niederman's identity was that Andrews was the preeminent author in this unique genre and that books written by a well-known author have an enhanced chance of commercial success." Sequels were more likely to sell if they bore Andrews' name.

Although the Andrews case is dramatic for its legal explicitness (as well as for the large dollar amounts involved), one lesson we may take from it is that it merely draws our attention to something that has always, in some sense, been true: author-attributions serve a branding function as much as anything else. Plastering the name of the author in large letters across the book cover acts as a kind of quality-assurance guarantee; it makes an implicit promise to readers that the contents of this codex will resemble those of previous works published under the same brand. In a recent article, legal scholar Laura A. Heymann makes just such an argument.[69]

---

69. Laura A. Heymann, "The Birth of the Authornym: Authorship, Pseudonymity, and Trademark Law," *Notre Dame Law Review,* Vol 80, Iss. 4 (2005): 1377-1450.

Suggesting that what she calls “authornyms” should be explicitly covered by trademark law, she notes that even apparently “true” author-attributions involve presentation decisions on behalf of publishers and authors: is it “Joane Rowling” or “J.K. Rowling”? “Virginia Andrews” or “V.C. Andrews”? Such decisions are not altogether different from the choice made by other corporations over whether to market soap under one brand name as opposed to a different one. Heymann’s argument seems particularly convincing in the case of authors who work under multiple names. When Stephen King published books under the name “Richard Bachman” between 1977 and 1982, it was in part a marketing decision: King and his publishers were worried that audiences would not accept more than one book per year by the same author.[70] Similarly, novelist John Banville writes crime fiction under the name Benjamin Black. In this case, both Banville and his publishers are open about the true identity of the author. The purpose of the name is simply to organize his works into different market segments: literary fiction on the one hand and crime fiction on the other.

In this way, author names fulfill a function that in other industries would be fulfilled by a brand or trademark: they provide an anchor point around which consumers can organize repeated consumption, which in turn gives producers a way of rationally predicting how many units of their product they might sell. As publishing industry analyst Leonard Shatzkin puts it,

> There is virtually no trademark identification in trade book publishing. A consumer who has enjoyed a Sara Lee cake or had a good experience with a GE washer is well disposed to try another Sara Lee product or GE appliance. The consumer who has been up all night reading an exciting novel will not look for ‘another Random House book’… though he may search out other books by the same author.[71]

70. Stephen King, “Frequently Asked Questions,” *stephenking.com*. Retrieved June 6, 2022.

71. Leonard Shatzkin, *In Cold Type: Overcoming the Book Crisis* (New York: Houghton Mifflin, 1982): 3.

This is why authorial brand names provide one solution to the problem of abundance: it allows publishers to take advantage of the name recognition that other industries can exploit by means of their brands and trademarks. But although its function is homologous, its ownership structure is not, since publishers almost never own their authors' names. Even in the case of someone like V.C. Andrews, the name was transferred to her estate, not her publisher. As I hope the preceding analyses have shown, this fact means that class conflict in the political economy of textual production often takes the form of bargaining over attribution rights. In screenwriting, that bargaining has sometimes looked like union contracts which explicitly guarantee attribution to all writers. But in fiction, authors typically work alone and rarely bargain collectively. As such, an author's bargaining position when it comes to new books is often based on nothing other than the celebrity around their name. For unknown authors, this counts for very little. For the famous, however, this can be a powerful bargaining chip, with the most well-known authors able to bargain for advances in the tens of millions of dollars.[72] That bargaining chip can even sometimes become a financial asset bequeathed to one's children.

2.6. Conclusion: Literary Developers

In this chapter, I have considered a number of non-traditional arrangements for organizing the production of texts and the distribution of income and credit for those texts' production. Though my primary aim has been to explore an aspect of the publishing industry that has been neglected by scholars, these newly-prominent authorial arrangements also help us

72. James Patterson set records in 2009 by signing a deal with Hachette for seventeen books over three years, worth $150 million. See Namera Tanjeem, "10 Of The Biggest Book Deals In History," *Bookriot*, February 24, 2020. https://bookriot.com/biggest-book-deals/

retheorize the singular type of authorship taken for granted in literary studies, which I have called "Romantic authorship." In this perspective, "Romantic authorship" is not simply an ideological "bourgeois" fiction but is rather a particular arrangement of textual production and accreditation. Romantic authorship can be reconceptualized as a social arrangement: it is a form of textual production that assigns certain tasks and rights to a single human person. In the Romantic arrangement, one person—who we call the "author" of a text—is responsible for the conception and planning of the content and the composition of the text; in return, they receive a legal monopoly over the reproduction of the text and any income derived from the sale of those reproductions. In this way, Romantic authorship is an ideology about the status of professional writers, a collection of "everyday habits," and a legal arrangement, at the same time.

Many attempts to historicize the collection of ideologies and habits entailed by "Romantic authorship" have described it as a response to capitalist modernity, either as the self-affirming ideology of a nascent class of self-employed, professional writers or a set of social and legal privileges granted to that class. The authorship arrangements that I have considered in this chapter suggest the possibility that the link between Romantic authorship and capitalism has been overstated. Or, more precisely, these cases suggest that the *sine qua non* of "Romantic authorship" lies in conception more than in composition *per se*. Legally and practically, one can be a writer without being an author, or an author without being a writer. The future of literary production may look more and more like the recent past of other creative industries, in which a creative director of some sort, like an *auteur* director of film or a showrunner in television, conceives of the overall premise of a literary work and directs its execution by a larger team, who work as his or her employees. Such an arrangement is less the "death" of the author than a case of the author-function being split up into a more complicated social division of tasks.

This is essentially the model of authorship adopted by a number of recent literary start-ups staffed by professional writers who independently produce novels or other narrative content on a collaborative or for-hire basis. Sometimes called “literary development companies,” these small organizations resemble the older book packager model, except updated for the twenty-first century: often explicitly transmedia in scope, they generate the story “content” that may appear in film, television, or print, as the case may be. Traditional book packagers collaborate almost exclusively with publishing houses and frequently employ ex-publishing professionals such as editors; they often act as adjuncts to traditional publishers and fulfill many of the same tasks in the publishing circuit. They only rarely compose fiction in-house. Literary development companies, by contrast, bill themselves as story producers first and foremost. Notable contemporary examples of such companies include As You Wish Literary Development, co-founded by journalist Stacey Hertz and bestselling novelist Sacha Wunsch; Cake Creative, a self-described “content generator” whose president, Dhonielle Clayton, “originates” each story herself; or Glasstown Entertainment, founded by bestselling author Lauren Oliver and poet Lexa Hillyer, and named for the childhood fantasy world of the Brontë sisters. For the writers they employ, who are often debut authors or writers without an established audience, the appeal of working with a literary development company comes both from the stability and the opportunity for a larger career. Though these writers may not be authors in the classic sense, they inhabit an ethos of professionalism. As Paula Suhr, a writer discovered by Glasstown Entertainment (then known as “Paper Lantern Lit”) puts it: “Ever since middle school, I’ve liked writing, but you grow up not thinking of it as a viable career choice… I learned craft, discipline, and the industry from working with PLL—and that continues to make me a better writer."[73]

The "literary development" model of authorship looks a lot like the model of authorship that has been institutionalized in film and television for more than a half-century: the writers' room. Is the future of literature—or, at least, of commercial literature—a world in which novels are produced in the same way as screenplays or television scripts? It certainly seems safe to assume that this model will become more common. But there are economic reasons to believe that it will never be as hegemonic in the literary economy as it is in the other culture industries. In particular, it is in the nature of texts that they are comparably cheap to produce; unlike a film or a television show, the composition of novels doesn't necessarily require a substantial investment in special machinery or in the labor of other people. A young writer can manufacture a new manuscript with only a laptop and their own sweat and tears. For this reason, literary production will likely continue to be a somewhat economically-contradictory field; a stubbornly-independent literary yeomanry, resistant to wage labor, will likely co-exist with corporate authors who pay salaried or contracted professionals to write by committee

73. Danielle Sacks, "Paper Lantern Lit Matches Undiscovered Writers With Publishers to Create Mass Market Hits," *Fast Company*, October 15, 2012. https://www.fastcompany.com/3001757/paper-lantern-lit-matches-undiscovered-writers-publishers-create-mass-market-hits

Chapter 3. Dimensions of Prestige: A Cluster Analysis of the *Book Review Index*, 1965-2000

> "Who are these 'elect' anyway? Every writer or artist has his 'elect,' intellectuals being in fact fragmented groups and sects of the 'elect'..."
> —Antonio Gramsci[1]

This chapter explores a dataset that supports research on the history of American literary reception between 1965-2000: the digitized contents of the *Book Review Index* from its first edition in 1965 through the year 2000. Owned and published by educational publisher Gale for most of its history, the *Index* aimed to provide a comprehensive reference document tracking metadata on book reviews for new titles published in more than 500 different periodicals and newspapers. The *Index* provides rich portrait of the reception history of literature in this period. For each of over one million unique books, the dataset includes information on the book's title and the author, as well as a reference to the page and publication of every indexed review of that title published in over 900 distinct periodicals. These publications span a diverse range of genres and intended audiences. They include many with substantial circulation numbers, such as industry publications like *Publisher's Weekly* and *Booklist*; major newspapers like *The New York Times* and *The Los Angeles Times*; and mainstream periodicals like *Time*, *The New Yorker*, and *The Economist*. But it also includes reviews published in a wide variety of smaller publications targeted at more niche audiences, including literary magazines like *Kenyon Review*, academic journals in history and sociology, genre-specific popular magazines such as *Analog Science*

1. Antonio Gramsci, *Selections from Cultural Writings*, ed. David Forgacs and Geoffrey Nowell-Smith, trans. William Boelhower (London: Lawrence and Wishart, 1985): 275.

Fiction, and journals published by or for particular ethnic communities, such as the now-defunct journal *Negro Digest*. The *Index* tracks Anglophone journals; most of its journals are American, but it contains some British and Canadian periodicals, as well. The extracted data does not track specific dates of publication for reviews, but it is segmented into four unevenly-sized historical blocks: 1965-1984, 1985-1992, 1993-1997, and 1998-2000.[2]

The dataset itself was originally compiled by Richard So and Dan Sinykin for two unrelated projects on the history of the American publishing industry.[3] Previous work with this data had a thickly-descriptive focus and supplemented entries on individual titles with additional demographic and historical metadata on authors, including information on gender and race. By contrast, in this chapter, I make use of all of the more than three million book reviews of the full dataset. The breadth of this data provides a unique perspective on the institutions of literary reception in the late-twentieth century. It allows one to ask which authors were most frequently reviewed within different segments of the literary field, as well as to compare authors by the different social contexts in which they were received. But unlike in previous projects derived

---

2. Gary C. Tarbert and Barbara Beach, eds., *Book Review Index: A Master Cumulation 1965-1984* (Detroit: Gale Research Company, 1985).

Neil E. Walker and Beverly Baer, eds., *Book Review Index: A Master Cumulation 1985-1992* (Detroit: Gale Research Inc., 1994).

Beverly Baer, ed., *Book Review Index: A Master Cumulation 1993-1997* (Detroit: Gale Research, 1999).

Beverly Baer, ed., *Book Review Index: 1998 Cumulation* (Detroit: Gale Research, 1999).

Charles B. Montney, ed., *Book Review Index: 1999 Cumulation* (Detroit: Gale Group, 2001).

Charles B. Montney, ed., *Book Review Index: 2000 Cumulation* (Detroit: Gale Group, 1999).

3. Richard So, *Redlining Culture: A Data History of Racial Inequality and Postwar Fiction* (New York: Columbia University Press, 2020).

Daniel Sinykin, *The Conglomerate Era* (New York: Columbia University Press, Forthcoming).

See the appendix for a discussion of how the data was prepared, as well as how it was used in these previous projects.

from this data, the scale of the full dataset makes it impractical to match individual authors or titles with additional descriptive information. At the level of individual titles, the data I consider in this chapter is extremely sparse: for each review, it tracks only the book's title, the author's name, and the publication in which the review appeared. The dataset includes no information on the review itself, such as its content, length, or even its disposition towards the title in question. At first glance, this sparsity would seem to present problems for using it for reception history: how can we ask questions about how an individual title was received if we don't even know whether its reviews were positive or negative? But as I will show, what this zoomed-out, macroscopic perspective lacks in detail it makes up for in comprehensiveness. Although the data lacks information on the content of reviews, the very decision to commission and publish a review of a title is itself an implicit act of evaluation; if nothing else, it indicates that the publication's editors believe that information about the book is worthy of their readers' attention. When millions of these acts of implicit valuation are taken in aggregate, they provide an unprecedentedly-detailed map of literary reception in the United States at the end of the twentieth century.

My primary claim is that literary reception in this period cannot be accounted for in one-dimensional terms, of a single hierarchical pyramid of prestige with canonical or "elite" authors at the top and the unwashed masses of popular culture at the bottom. Instead, literary reception at the end of the twentieth century was stratified and balkanized, characterized by isolated communities of reception with their own distinct hierarchies of value. As I will show, review periodicals were often highly specialized, selecting the books they reviewed with an eye for a particular audience. These audiences came in different shapes and sizes. Some periodicals were targeted at a particular kind of literary professional: *Booklist* for librarians and *Publishers Weekly* for publishing professionals. Others targeted a readerly niche defined by genre preference: romance readers, or readers of science fiction. Still others targeted a regional niche, or a religious one. These hierarchies sometimes overlapped with each other: an author might, for example, have been prominent both in science fiction journals and in wide-circulation mainstream magazines. But more often, these communities had very little overlap at all; authors who were stars in one literary subfield might be completely absent from the book reviews of a different one. In order to make the claims of this chapter, I make use of a suite of computational methods for measuring similarity and clustering high-dimensional data.

In one respect, the state of affairs described by this chapter is, by and large, a very intuitive one. It will not come as much of a surprise to learn that *Parnassus* and *Romance Reader* have almost zero overlap in the set of books that they reviewed between 1965 and 2000. But it might be more interesting to learn that the *Los Angeles Times Book Review* has greater review overlap with other regional publications, such as the *West Coast Review of Books*, than it does with other broad-circulation national newspapers, such as the *New York Times Book Review*. When one repeats this process for comparisons between over 500 distinct periodicals, it quickly becomes clear that the history of literary reception in the late twentieth century is more granular and diffuse than it has sometimes been depicted in our literary histories of the period.

In terms of the wider aims of this dissertation, book reviews are another institution for managing the problem of abundance. Whereas bestseller lists and brand-name authors recommend new titles based on past sales, book reviews represent a more personal—and more labor-intensive—method of managing abundance. Book reviewers and the editors who commission their reviews are a cadre of professional recommenders: they survey and evaluate a large number of books, in order to generate relatively succinct and digestible accounts of whether this or that book is worthy of a reader's attention. This is not to say that book reviews are necessary for a title's success, or even terribly important; accessing the causal impact that a good or bad review has on the larger reputation of a book is beyond the scope of this chapter. But one thing that is clear is that the size of the book review market between 1965 and 2000 in part reflects the growth of the publishing industry in that same period. With so many unique books on the market and so many different potential readers to purchase and read them, journals that were specialized in particular readerly niches found more than enough material to keep their book reviewers busy.

The structure of the chapter is as follows. In section one, I survey the existing historical scholarship on the institutions and culture of book reviewing in the United States, with particular focus on the book review market in the second half of the twentieth century. In section two, I provide a detailed description of the data, including summary statistics on the distribution of reviews. In section three, I identify and describe the most prominent "reception communities" in the data, or publications that tend to review the same authors. I compare the composition of the different communities and consider which authors figure most prominently within each. In section four, I discuss how my methodology both builds upon and departs from existing work in computational literary studies on prestige in literary reception and provides an alternative model for theorizing prestige in quantitative terms. Whereas previous research tended to model prestige as a single-dimensional variable, something which can only vary in magnitude and not in kind, I propose a method of modeling prestige in which prestige is a multidimensional variable. In other words, there are many "prestiges," reflecting the particular tastes and hierarchies of value associated with specific reading communities, which may or may not overlap with each other. In a concluding section, I discuss how the reception communities discovered by the clustering methodology relate to the concept of a literary canon.

3.1. The American Book Review Market

In the following section, I survey how book reviews have been theorized in book history and the sociology of literature. Although quantitative studies of book review behavior are rare in literary history proper, book reviews themselves have been important objects for the theorization of literary reception in these adjacent fields. Few institutions evoke key cultural sociology concepts like “gatekeeping” and “taste making” more poignantly than the market for reviews. Insofar as the book review is a genre quite literally designed to mediate the reception of a work of art, it is a favorite object for the theorization of literary prestige. Book reviews figure prominently in the work of Pierre Bourdieu, whose analyses of cultural markets provide the theoretical basis for much recent work on reception history in the sociology of literature. For Bourdieu, art critics—along with other cultural intermediaries such as publishers, teachers, and gallerists—are among the key “producers of the meaning and value” of a work of art. Bourdieu portrays these intermediaries as essential to the production and reproduction of aesthetic value: their “combined efforts produce consumers capable of knowing and recognizing the work of art as such.”[4] Book reviews are a favorite object for identifying the specific tastes of particular aesthetic communities and class-fractions. In one case study, for instance, Bourdieu compares how newspapers located in different regions of the literary field received Françoise Dorin’s play *Le Tourmant*. The fact that critics on different side of the field “can pick out the same properties”

4. Pierre Bourdieu, *The Field of Cultural Production: Essays on Art and Literature* (New York: Columbia University Press, 1993): 37.

but with opposite interpretations (ironic/non-ironic) provides evidence of the polarized structure of the field.[5] Elsewhere, Bourdieu emphasizes the gatekeeping role of critics and other intermediaries. As he puts it, "Entering the field of literature is not so much like going into religion as getting into a select club: the publisher is one of those prestigious sponsors (together with preface-writers and critics) who effusively recommend their candidate."[6] For Bourdieu, the precise profession or job title of cultural intermediaries is less important than their power to consecrate, that is, their power to legitimize an artist by granting him or her a place in the exclusive club. In many cases, artists themselves perform the gatekeeping function. In fact, the power to consecrate other writers—with "a preface, a favorable review, a prize, etc."—is itself the best evidence of a writer's own consecrated status. [7]

The taste making function of the book review also figures prominently in the scholarship on middlebrow culture. Beth Driscoll describes newspaper book reviewers as "gatekeepers... constructing canons and identifying cultural movements, while also teaching people how to read: what to value or despise in a reading experience."[8] Focusing on the reception of *Harry Potter* in reviews of this kind, she notes that the series was lambasted by many reviews in "high-status publications" like *The Guardian* and the *New York Review of Books*, which expressed a "disdain for popularity and a defense of elite literary culture." But more often, however, reviews adopted a "permissive stance," based on a traditional middlebrow logic that values ethical development.[9]

---

5. Bourdieu , 93.

6. Bourdieu , 77.

7. Bourdieu , 42.

8. Beth Driscroll, *The New Literary Middlebrow: Tastemakers and Reading in the Twenty-First Century* (Palgrave Macmillan, 2014): 101.

9. Driscroll, 105.

For a broader discussion of the concept of the middlebrow, see Janice Radway, *A Feeling for Books: The Book-of-the-Month Club, Literary Taste, and Middle-Class Desire* (Chapel Hill: University of North Carolina Press, 2000).

However, evaluation is not the only purpose of the literary genre of the book review. Reviews also serve a related but distinct news function: they transmit basic descriptive information about books on the market, especially newly-released books. In a market where many thousands of books are in print at any one time, this informational function is highly valued by publishers, authors, book buyers, and readers. Although the evaluative function is perhaps what we most often think of when we think of a “review,” the more descriptive and less overtly prescriptive “news” mode of book reviewing was predominant for much of American literary history. As Joan Rubin shows, the majority of book reviews published in the United States in the nineteenth and early twentieth century were primarily conceived as documents for disseminating news about current and recently-published titles. In fact, in the antebellum period, what Rubin calls the “news approach” was “virtually the only mode of presentation in the daily press.” Newspapers frequently published “book notices,” brief and unattributed announcements of new titles, that were sometimes prefabricated by publishers themselves, blurring the line between news and advertisement.[10] Even after genteel magazines like *Harper’s*, *Scribner’s*, and *The Atlantic* helped to popularize and formalize a model of the book review as a genre of mass-circulation literary criticism, the majority of reviews published in newspapers in the late nineteenth and early twentieth century continued to emphasize reporting about literature, rather than evaluating it. Of course, the line between description and evaluation was always blurry; reviewers frequently inserted their own moralistic or aesthetic judgements in this supposedly-descriptive genre.[11] But the news-function of book reviewing remained pervasive even at the turn

10. Joan Shelley Rubin, *The Making of Middlebrow Culture* (Chapel Hill: University of North Carolina Press, 1992): 35-6.

11.Rubin, 40.

of the twentieth century. According to Rubin, when Adolph Ochs created the *Saturday Review of Books and Art* in 1896, he did so in the belief that "books could be treated as news."[12]

In the second half of the twentieth century, publishers, editors, and book reviewers frequently remarked that the institution of book reviewing was caught in tension between these two related functions. In this period, the book review market was in state of transition. The postwar expansion of book production set off by the paperback revolution had considerably increased the size of the American reading public, and with it, the potential size of the audience for book reviews. This expanded audience was both larger and more diverse than its predecessors, in terms of income, education, and geography. Some editors of local newspapers felt that the evaluative, literary-critical style of reviews published in serious New York literary magazines was poorly-suited to the needs and interests of this new audience. At a booksellers' conference in 1961, Stanton Peckham, book editor of *The Denver Post*, argued that reviews ought to be written for "the general public." [13] In that piece, Peckham, observing that over a million people see the book review section of the *Post*, claimed that it would be "the height of absurdity to address them all as people of superior IQs and holders of Ph.D. degrees." He went on to suggest that those people who "understand and appreciate the highfalutin reviewer whose discussion of a new book includes application of Aristotelian principles and Freudian analysis" are "few among a million" and so were not the "people [they were] trying to reach." As Peckham saw it, the "book buying public" wanted "first of all to know what it is offered for its money." Although some evaluation was acceptable, Peckham believed that his audience often had "only a casual interest in the reviewer's opinion, good or bad" and were often interested in a book based

12. Cited in Rubin, 40.

13. Stanton Peckham, "The Reviewer, the City and the Bookstore," *Publishers Weekly*, March 20, 1961: 26-27.

on “subject matter.” With a certain chagrin, he noted as well that dismissing a book as “tripe” may actually sell it, a lesson he learned after one of his bad reviews had him “denounced in every woman’s club in town.”[14]

Meanwhile, reviewers at prestigious literary magazines and other members of the New York literary establishment frequently took the opposite line, arguing that the evaluative function of the book review was more important than ever. At a seminar at the National Book Awards Week in New York titled “Book Reviewing Today,” one New York-based magazine reviewer disparaged newspapers that treated their book sections as mere “shopping guides,” which made no attempt at serious criticism.[15] Theodore Solotaroff, former book review editor of *Commentary* and *Book Week*, expressed frustration at the pressure he sometimes received to emphasize the “publicity” or the “news” aspect of book reviewing. In his opinion, the point of a book review is neither to build publicity for a book nor to report news about it; rather, the “only thing that really matters is a reviewer’s judgement.”[16] Richard Gilman, book review editor of the *New Republic*, expressed a similar opinion. For Gilman, the difference between book reviews and book news was akin to the difference between literature and books. As he put it, “Books are not necessarily literature. The history of the cash register is a book, but not part of our culture. Most books of general interest are not literature. Most novels are not literature.” [17] He further suggested that the book review media ought to reflect this fundamental distinction, perhaps by introducing separate sections for serious literary criticism and for general reporting about the contents of new books. Unsurprisingly, the newspaper reviewers at the seminar did not take kindly to this

---

14. Peckham, 26-27.
15. “Reviewers Debate Function of Reviewing,” *Publishers Weekly,* March 18, 1968: 30.
16. “Reviewers,” 31.
17. “Reviewers,” 32.

disparagement. One Midwestern reviewer countered by suggesting that the New York reviewers were an excessively insular group, writing only for each other and discussing only books and topics of interest to their in-group.[18]

These complaints about the state of book reviewing were tame compared to a well-known critique of the field written by novelist and critic Elizabeth Hardwick and published in *Harper's* in 1959. Titled "The Decline of Book Reviewing," Hardwick's piece scathingly described American book review media as characterized by toothless praise-mongering. [19] For Hardwick, American reviewers were incurably undiscerning and fundamentally unwilling to publish negative judgments. As she put it, "Sweet, bland commendations fall everywhere upon the scene; a universal, if somewhat lobotomized, accommodation reigns. A book is born into a puddle of treacle; the brine of hostile criticism is only a memory." Hardwick reserved her most vociferous criticism for the *New York Times Book Review*: its "flat praise and faint dissension" have made it into little more than a "provincial literary journal," ultimately no different from "all those small-town Sunday 'Book Pages.'" The result was a literary culture in which "the drama of the book world" was being "slowly, painlessly killed." All books are treated as if they were "somehow alike," a situation she describes as a "democratic euphoria" that will "hardly meet the needs of a serious work."[20] She does not limit herself to mere rhetoric, either; she cites a study from Wayne University that 51% of the reviews in *Book Review Digest* in 1956 were favorable, while an additional 44.8% were noncommittal.[21]

---

18. "Reviewers," 30.

19. Elizabeth Hardwick, "The Decline of Book Reviewing," *Harper's Magazine*, October 1959. https://harpers.org/archive/1959/10/the-decline-of-book-reviewing/

20. Hardwick.

21. She gives no indication as to whether the study suggests this number is different from the past.

The negative connotation with which Hardwick uses the word “democratic” is telling. One needn’t be Pierre Bourdieu to see that the derision that Hardwick and the other professional critics quoted above have towards newspaper reviewers is in part a defense of the traditional cultural authority of literary criticism. The problem with the universally-positive style of popular book reviewing, for these critics, is that it represents the abdication of this cultural authority. “The editors of the reviewing publications,” Hardwick writes, “no longer seem to be engaged in literature.”[22]

Hardwick’s impression of reviewing norms was shared by many of her contemporaries. Novelist Geoffrey Wagner, for instance, in an article with the exact same title as Hardwick’s, bemoans the “functionaries of the *Saturday Review* syndicate” who seem to imagine America as “so culturally backward that the production of a book, any book, is a triumph of pioneering ingenuity to be greeted with glee.[23] Hardwick’s polemic had a ready-made audience of like-minded thinkers. It is perhaps no surprise, then, that when she went on to co-found the *New York Review of Books* with Robert B. Silvers, Barbara Epstein, and A. Whitney Ellsworth in 1963, its first issue sold more than 100,000 copies. In this inaugural issue, the editors framed the project in opposition to the news-style of reviewing: “This issue of *The New York Review of Books* does not pretend to cover all the books of the season or even all the important ones. Neither time nor space... have been spent on books which are trivial in their intentions or venal in their effects…”[24]

---

22. Hardwick.

23. Geoffrey Wagner, “The Decline of Book Reviewing,” *The American Scholar*, Vol 26, No. 1 (Winter, 1956-67): 25.

24. Editors, “The Opening Editorial,” *The New York Review of Books*, February, 1963. https://www.nybooks.com/articles/2013/11/07/opening-editorial/

Although I have been unable to identify the study of the *Book Review Digest* cited by Hardwick, it is worth noting that, for some publications, a preference for positive reviews was (and remains) a conscious editorial policy. This policy is clearest in the case of publications targeted at library or publishing professionals responsible for making book orders on behalf of an institution, such as *Library Journal*, *Choice*, *Booklist*, and *Publishers Weekly*. These publications are designed with a fundamentally different vision of book reviewing in mind. For publications of this type, the purpose of a review is less to describe a culture or debate ideas than to process information: to sift through the enormous mass of new titles published each year, an endeavor which far exceeds the capacities of any individual reader or book purchaser. These publications design their reviews in large part to assist librarians or bookstore purchasers in making their selections. In this context, a negative review is simply less useful than a positive one and a positive evaluation may simply be implicit in the honor of having been reviewed at all. *Booklist*, for instance, pursues what it calls a "recommendation only" policy, exclusively publishing reviews for which the reviewer believes the book in question is a recommended purchase. The journal's aims are in large part functional and descriptive: simply to make librarians aware of a book in question. Its editorial policy reflects these aims: described as "the haiku of book reviewing," it limits its reviews to 150-175 words, including basic information such as context, synopsis, ideal audience, and "read-a-likes," or potential suggestions for similar books that a librarian might recommend to interested readers.[25] Although other publications allow negative reviews, they nevertheless similarly have a preference for positivity. A 1981 study of *Library*

25. "Write for Us," *Booklist Online*. https://www.booklistonline.com/writing-for-booklist.

Journal and *Choice* found that only about 18% of the reviews published by the two journals were negative.[26]

Importantly, publications that conceive of reviewing in this way also tend to publish an order of magnitude more reviews annually than literary magazines like the *New York Review of Books*. According to the data used in this chapter, the *New York Review of Books* published approximately ~13,000 book reviews between 1965 and 2000. The journal *Choice*, by contrast, published more than 220,000 in that same period. Other journals targeted at librarians and book professionals are similar: *Booklist* published more than 200,000 and *Library Journal* more than 210,000. Clearly, these journals are simply operating with a different vision of the scope and purpose of book reviewing. But it is a vision of the review market that would be a nightmare for somebody like Hardwick: hundreds of thousands of reviews, none of them scathing. Indeed, even *Saturday Review*, a target of scorn for both Hardwick and Wagner, published almost twice as many reviews as the *New York Review of Books*. In this data, *Saturday Review* published 13,993 book reviews between 1965 and 1984, as opposed to 7,634 from *New York Review of Books* in that same period.

There is an important way in which the reviewing culture imagined by Hardwick and other members of the New York literati requires a much smaller scale of literary production and conversation than the scale characteristic of industry publications, or even mainstream publications like *Saturday Review*. Hardwick's vision of book reviewing is personal, even intimate; she imagines a small number of reviews, passionately argued, in publications that favor reviewers with character and voice. She expresses dissatisfaction with the blandness of a

26. Beth Macleod, "*Library Journal* and *Choice*: A Review of Reviews," *The Journal of Academic Librarianship*, Vol 7, No 1: 23-28.

reviewing world without “involvement, passion, character, eccentricity.” But reviews of this sort take time to write—to say nothing of the time that must be spent reading (or at least skimming) the book being reviewed. The reviewing world she critiques is, in many ways, simply one that has adapted itself to the new size of the American literary world in the second half of the twentieth century. Her sense of the lack of passion of this world evokes a sense of industrialized, repetitious dullness: “Books pile up, out they go, and in comes the review… The wares offered by the better writers are apt, frequently, to be something less than their best.” Again, the *New York Review of Books’* opening editorial is telling: the editors do “not pretend to cover all the books of the season or even all the important ones.” Hardwick’s project necessarily requires a lot of selectivity; the style of review typical of the *New York Review of Books* could not be produced at the scale of an industry publication.

Journals less dedicated to this vision of the heroic reviewer will be better equipped to review a larger number of new books. But contra Hardwick, I think it would be wrong to assume that the neutral style of mass-review outlets like *Publishers Weekly* results in an absolutely “democratic” book review field. Although these publications do not perform very much evaluation in the actual text of their reviews, they nevertheless participate in an implicit system of literary evaluation insofar as they influence the allocation of literary attention. Even the most comprehensive review outlets only publish reviews for a fraction of all the new books released each year. Though 220,000 is already a huge number of reviews, the dataset features reviews for over 1.2 million unique books over the whole period. In other words, even heavy-hitters like *Choice* and *Booklist* are only publishing reviews for less than 20% of the books featured in the *Book Review Index*, to say nothing of the books that do not appear in the data at all because they were not reviewed by any of the indexed journals (self-published books, for example, or simply extremely obscure books). Despite the comparatively wide net cast by these journals, they nevertheless filter down the full set of all new books to some smaller portion that they feel are worthy of even receiving their attention in the first place. This kind of taste making is considerably more subtle than a book review that pans or praises its subject matter.

This quiet, background taste making function of book reviewing is the first line of defense of the existing structure of literary market. New books and authors have to get over this filtration barrier in order to even receive the attention of the readers of these magazines in the first place. Notably, publishers themselves sometimes describe the attention value of book reviews as more important than the actual recommendation value of the review. From the perspective of the marketing team, all press is good press. The book review market is like a great filter, controlling the distribution of publicity, though not necessarily of sales. Although a publisher cannot guarantee that a new novel will sell well, they can guarantee that information about that novel will appear in front of the eyeballs of people who might be interested. So, it is perhaps unsurprising that publishers continue to value book reviews and book reviewers; the process of attention filtration that they perform is something that publishers can actually directly influence.

3.2. Data from the *Book Review Index*

In this section, I provide a summary of some key aspects of the data and its distribution. I ask what the overall shape of the data's distribution tells us about the book review economy in this period. As I show, the book review market as seen through this dataset appears extremely polarized: most authors received only a handful of reviews, whereas a few influential authors were reviewed dozens of times, by a variety of different journals. These widely-reviewed authors include novelists who are today considered canonical literary authors of the post-1945 period, such as Thomas Pynchon. However, as I elaborate in the following section, it is misleading to consider authors only by their raw count of total reviews. Where they were reviewed—and where they were not reviewed—ends up mattering quite a lot for understanding the pattern of their reception.

But first, the overall distribution. What does the data look like as a whole? The full dataset features reviews of approximately 1,330,400 books. For a description of the process by which the data was generated, consult the appendix. Excluding titles without a listed author and obvious duplicate resulted in a working set of 1,218,257 books. The data extraction process recovered metadata on 3,355,879 reviews of these books. This is quite a large number, so it is worth beginning with some summary statistics. The first thing that one notices is that review counts are highly skewed: the mean number of reviews is ~2.75 and the median is one. In fact, out of a total of ~1.21 million books, 625,300—just over half—were reviewed by only one journal. **Figure 3.1** below shows a histogram that gives a sense of this skew (note that the y-axis is on a log scale). Most of the mass of the distribution is located on its left-hand side. Those few books scattered on the far-right end of the distribution are literary stars, at least in terms of prestige: they are those books that received an abnormally large number of reviews. **Table 3.1** lists the top three most-reviewed books for each period of the dataset: 1965-1984, 1985-1992, 1993-1997, and 1998-2000. This list includes a few names that will be unsurprising to scholars of post-45 literature—Philip Roth, for instance—as well as a few that are less obvious.

Figure 3.1. Histogram of total reviews of each book in the *Book Review Index* data.

title	author	reviews	period
Watership Down	Richard Adams	75	1965-1984
The Fate of the Earth	Jonathan Schell	69	1965-1984
Robert Kennedy and His Times	Arthur M Schlesinger, Jr.	63	1965-1984
A Bright Shining Lie	Neil Sheehan	124	1985-1992
The Counterlife	Philip Roth	112	1985-1992
The Golden Gate	Vikram Seth	104	1985-1992
The Bell Curve	Richard J. Herrnstein	59	1993-1997
The Robber Bride	Margaret Atwood	55	1993-1997
Preparing for the Twenty-First Century	Patricia Kennedy	53	1993-1997
Underworld	Don DeLillo	46	1998-2000
How the Mind Works	Steven Pinker	43	1998-2000
Mason and Dixon	Thomas Pynchon	42	1998-2000

Table 3.1. Most frequently reviewed books.

This list has not been filtered by genre; these are the most frequently reviewed books out of all books in the dataset. It is interesting, then, that literary fiction takes up so many of the top spots. Though we are used to thinking of literary fiction as a part of Bourdieu's "restricted field" of cultural production, this restrictedness is relative. It is true that authors like Pynchon and DeLillo have a much smaller audience than their contemporaries James Patterson and Danielle Steel. But within the field of books that are reviewed by book magazines and literary journals, the consecrated novelists are also the novelists with the largest "mass" appeal. They are able to command the attention of a large number of distinct magazines. Philip Roth's *The Counterlife*, for instance, was reviewed by elite journals like *New Yorker* and *Partisan Review*, industry magazines like *Publishers Weekly*, mainstream journals like *Time*, and many others. These novels are the crowd pleasers, if that crowd is made up of book reviewers. Besides award-winning novelists, other top-reviewed books include mainstream nonfiction, which are similarly able to capture the attention of a relatively wide variety of magazines. Many of these address then-contemporary political debates: Jonathan Schell's *The Fate of the Earth*, about the possible effects of nuclear war; or Richard J. Herrnstein's infamous *The Bell Curve*, the racist thesis of which provoked a storm of controversy and prompted a number of high-profile rebuttals (which would themselves contribute to the high review count for this book).

Compared to the full data, these books are exceptional; they were reviewed a far greater number of times than average. Many (though not all) have had lasting influence on elite literary culture. They have considerable overlap with some of the most famous high-cultural literary canons of the period. It is noteworthy, for instance, that these books include three out of four of the novelists whom Harold Bloom includes in his "canon of the American sublime": Pynchon, DeLillo, and Roth; only Cormac McCarthy is missing.[27] The fact that this overlap emerged simply by counting the raw number of book reviews, without transforming the data in any way, suggests that the amount of review attention a book received correlates with other ways of assessing a book's prestige or canonicity.

---

27. Harold Bloom, *The Daemon Knows: Literary Greatness and the American Sublime* (Spiegel and Grau: New York, 2015), 137.
See also Leonard Pierce, "Harold Bloom on *Blood Meridian*," *The A.V. Club*, 6/15/09. https://www.avclub.com/harold-bloom-on-blood-meridian-1798216782

But the majority of titles in the data are unlike the books listed above. Most of them are obscure, specialized books that received a very small number of reviews and that have already fallen out of cultural memory. Many are no longer in print. They include books like Mary Gentle's 1989 science fiction novel *Scholars and Soldiers*, which was reviewed only by *Locus: the Magazine of the Science Fiction and Fantasy Field*. The full data also has a greater diversity of genres and topics than the most-reviewed novels. It includes children's books like Peter Max's 1970 *The Land of Yellow*, specialized professional books like *The Handbook of Investor Relations*, academic monographs, books of poetry, and so on.

The distribution of reviews by journal is lopsided as well, but in a different way. As **Figure 3.2** shows, the total number of reviews published by each journal over the whole period roughly follows a lognormal distribution (the X-axis is on a log-scale). Most of the ~800 distinct journals in the data had between a few hundred and a few thousand book reviews in the index. A few outlying journals published far more, most of which I discussed earlier: library journals like *Choice*, *Library Journal*, *Booklist*, and *School Library Journal*; industry journals like *Kirkus Reviews* and *Publishers Weekly*; and mainstream publications with especially wide coverage, such as the *Times Literary Supplement* (~106,000) and the *New York Times Book Review* (~85,000). At the opposite end are specialized magazines that publish only infrequent book reviews, as well as some that only have a few reviews because they weren't in print over the whole period considered: *Byte Magazine* (14), *Model Railroader* (28), *Ballet News* (41). Journals in the middle ranges include academic journals, smaller mainstream magazines (*Irish Literary Supplement*), and genre-specific journals (*Locus*).

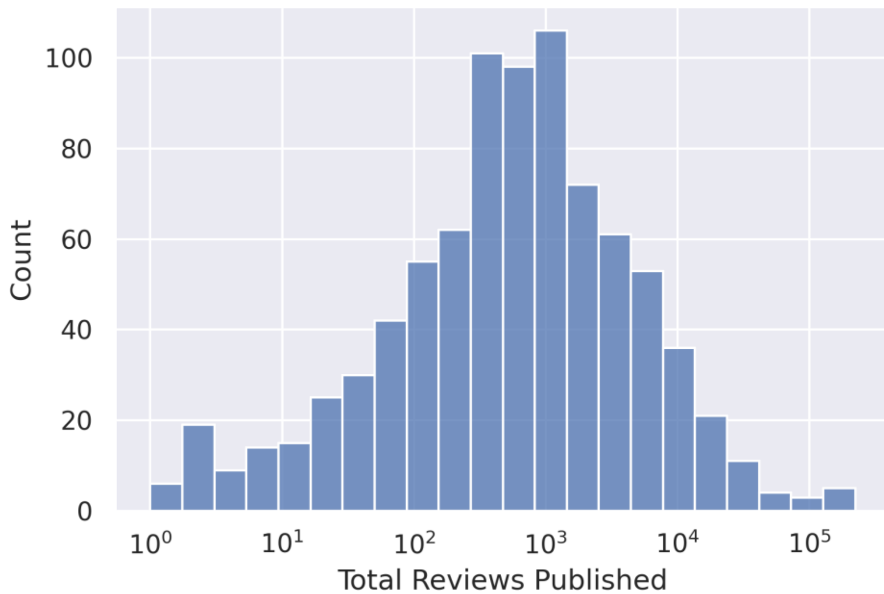

Figure 3.2. Histogram of total reviews published by journal. The distribution is a roughly lognormal distribution; the x-axis is on a log-scale.

Finally, there is the distribution of reviews by author. The three most frequently reviewed authors for each period are listed in **Table 3.2.** Notably, many of the names are different than in the earlier list. This is largely due to the fact that counting an author's reviews lends greater weight to more prolific authors: an author with many books has more opportunities for those books to be reviewed. Isaac Asimov, for instance, is one of the most prolific authors in the entire dataset, which helped him receive almost twice as many reviews as the author with the next highest total. He was also frequently reviewed because of the wide-ranging appeal of his work, among the many niches of the book review market. Few other authors could claim reviews by both the *New York Times Book Review* and *Nature.* Nevertheless, it is remarkable how much the implicit tastes of these names overlaps with those of the previous table. Canonical authors like Salman Rushdie and Mario Vargas Llosa have prominent positions. And even the conservative arch-aesthete himself, Harold Bloom, takes a spot among the most-reviewed authors between 1998 and 2000. But, this list does not simply include the literary elite. It also includes children's author Rosemary Wells (though one might argue that Wells has a comparable status within children's literature that somebody like Rushdie has within literary fiction).

author	reviews	period
Isaac Asimov	1380	1965-1984
Graham Greene	636	1965-1984
Robert Silverberg	630	1965-1984
Fay Weldon	439	1985-1992
Salman Rushdie	355	1985-1992
Mario Vargas Llosa	344	1985-1992
Kathryn Lasky	205	1993-1997
Joseph Bruchac	196	1993-1997
Eve Bunting	185	1993-1997
Harold Bloom	154	1998-2000
Rosemary Wells	141	1998-2000
Cynthia Rylant	140	1998-2000

Table 3.2. Most frequently reviewed authors. The table lists the three authors who have the greatest number of reviews for each of the four periods covered by the data.

3.3. Journals Find Their Niche

Of course, "total reviews" is a fairly simple metric by which to measure a book or author's position in the literary status market. It collapses the qualitative differences between authors and review outlets into a single number: reviews in *Library Journal*, *American History Review*, and the *Los Angeles Review of Books* are treated as commensurable. This masks the fact that not all authors listed above are "elite" in the same way; their frequent reviews did not all appear in the same journals. For example, in the 1998-2000 portion of the data, Rosemary Wells and Harold Bloom received book reviews by twenty five and forty one different journals, respectively. But of all those journals, only eight of them reviewed at least one book by both authors.[28] Clearly, these are authors who are respected for different reasons, or, at least, who are respected by different (if partially overlapping) segments of the book review market. Wells was disproportionately reviewed by journals specializing in children's literature, such as *Reading Teacher* and *Parents Magazine*. Bloom, meanwhile, was reviewed by what could be described as the elite-mainstream commentariat: the *New Yorker*, *Wall Street Journal*, *New Statesman*, *Christian Science Monitor*, and others (though very few academic journals, notably).

How to account for this limitation? In the simplest case, one could simply segment the data by journal and consider the most popular books or authors for each journal individually. Separating in this way gives one a view of prestige that is specific to particular literary micro-communities. For instance, the most-reviewed authors in the *New Yorker* between 1965 and 1984 are Georges Simenon, Isaac Bashevis Singer, and Vladimir Nabokov. Between 1985 and 1992,

---

28. Specifically, both authors were reviewed by *Booklist, Kirkus Reviews,* the *New York Times Book Review, Publishers Weekly,* the *Los Angeles Times Book Review, Time, School Library Journal,* and *Children's Bookwatch*. Most of Bloom's reviews in the latter two journals are actually edited collections of canonical works for which he wrote an introduction. The *Book Review Index* lists him as an author of these editions.

Poetry favored Charles Simic and Gerald Stem. If we do the same for *Time* magazine between 1993 and 1997, then the top authors are Peter Hoeg, James Ellroy, and John Grisham. One can imagine conducting this same process for every journal in the *Index*, resulting in a dizzying array of distinct "top ten" lists. These lists are unlikely to be exactly the same for any two journals and are instead reflective of the particular tastes and objectives of a particular journals' writers and editors.

But this method goes too far: it treats each journal as if it were completely distinct from each other journal. In actuality, journals are likely to overlap, at least partially. They are grouped into clusters of like-minded reviewers who tend to pick similar books to review. In statistical terms, each journal's reviews are not independent of each other journal's reviews. Reviews are correlated: if I know some of the journals that reviewed your book, I have a better-than-random chance of correctly guessing which other journals reviewed your book. If a book was reviewed in *Analog Science Fiction and Fact*, it's more likely that it was also reviewed by *Science Fiction Chronicle* than by *Poetry*. The same goes for *School Library Journal* and *Children's Book Review Service*, the *Journal of American History* and *American Historical Review*, and so on. The fact that journals are "grouped" into communities of reviewers with similar interests or even similar literary tastes can show us something about how reception is organized. It shows the cleavages and segmentations within the reception space, those "regions" of the literary market organized around particular shared standards of taste.

One way of modeling the distinct but partially-overlapping sets of journals that reviewed a specific author is with a method known as hierarchical clustering. A hierarchical clustering algorithm builds an inverted tree diagram—a "dendrogram"—such that each leaf of the tree shares branches with other leaves that are similar to it, according to some metric set by the researcher. The places at which branches split off in the resulting tree represent progressively more fine-grained groupings. Though the concept of hierarchical clustering is somewhat abstract, the resulting diagram is highly intuitive. For this chapter, I used the hierarchical clustering method implemented in the Python library sciPy. The similarity of two journals was assessed as the cosine distance between each journal's set of reviewed books. In simpler terms, two journals are considered similar if they reviewed many of the same books, and different if they rarely reviewed the same book. I used cosine distance as my similarity metric because it is less sensitive to differences in magnitude, which was important since some journals reviewed many more books than other journals. For the purposes of this clustering exercise, I considered only journals that were indexed by the *Book Review Index* in all four volumes, leaving a total of 285 journals. Since not all journals were in print over the entire thirty-five-year span of the data, failing to do this would cause journals to be grouped in part based on the period in which they were in print.

The resulting dendrogram is too large to be included in its entirety, so I reproduce only portions of it here. **Table 3.3** presents the ten largest clusters, when the threshold for cluster similarity is set at the median distance at which clusters were merged.[29] The overall picture is one of an opposition between the two largest clusters—which might be summarized as "mainstream magazines" and "children's literature"—and dozens of smaller clusters representing journals devoted to a particular academic discipline or professional community. Besides this,

29. This median value is ~0.94. Unlike "flat" clustering methods like K-Means, hierarchical clustering does not produce a set number of clusters. Its hierarchy can be "flattened" into a discrete number of clusters by choosing a distance threshold, below which all branches are considered part of the same cluster. A lower threshold produces a more "fine-grained" picture of the tree, whereas a higher threshold will group branches into fewer clusters of comparatively larger size.

description	sample journals
mainstream magazines	*Time — Los Angeles Times Book Review — London Review of Books — Observer — Publishers Weekly — Nation — New York Times Book Review — New York Review of Books*
children's literature	*Horn Book Magazine — Center for Children's Books, Bulletin — Social Education — Language Arts — Children's Book Review Service*
academic social science	*American Political Science Review — Journal of Politics — Sociological Review — Political Science Quarterly — American Journal of Sociology*
history	*Historian — Journal of Social History — American Historical Review — Journal of Modern History — English Historical Review*
religious studies	*Journal of Biblical Literature — Theology Today — Journal of Religion — Religious Studies Review — Theological Studies*
librarianship	*American Libraries — Library Quarterly*
philosophy	*Review of Metaphysics — Ethics — International Philosophical Quarterly — Philosophical Review*
American studies	*Journal of Southern History — New England Quarterly — American Literature — Journal of American Studies*
poetry	*Poetry — Hudson Review — Parnassus — American Poetry Review*
Canadian magazines	*Canadian Forum — Quill and Quire — Books in Canada — Maclean's — Saturday Night*

Table 3.3. Ten largest clusters discovered by hierarchical clustering. For each cluster, a heuristic description and a few sample journals have been added.

there are also clusters devoted to literary genres, such as a cluster made up poetry journals: *Poetry*, *Parnassus*, *American Poetry Review*, and others. Smaller clusters not present in the table below include a science fiction cluster (*Analog Science Fiction and Fact*, *Magazine of Fantasy and Science Fiction*), a women's magazine cluster (*Ms. Magazine, Signs*, *Women's Review of Books*), and others.

Importantly, this should not be interpreted as evidence that the many titles inside the "mainstream magazines" category are all very similar to each other. It just means that they are less different from each other than academic journals are from each other. Or, in other words, academic journals are extremely specialized; they simply don't review books outside of their discipline. Though the editors of *National Review* and *The New Yorker* have distinct tastes, this difference in taste does not compare to the disciplinary divide that separates *American Political Science Quarterly* and *Film Criticism*. The branches of the hierarchy that further subdivide the mainstream magazine group into smaller clusters are themselves informative. **Figure 3.3** below shows in detail the branches of the mainstream magazine portion of the tree. In general, the splits of the branches reflect intuitive divides in the periodical market. Wide-coverage magazines like *Publishers Weekly*, *Kirkus*, and *Booklist* are all together. British magazines occupy their own distinct branch. Explicitly conservative news and culture magazines all have their own section: *American Spectator*, *Commentary*, *Wall Street Journal*, and *National Review*. There is a separate wing for book reviews published in major urban newspapers, such as *Chicago Tribune* and the *Los Angeles Times Book Review*. And so on.

To get a more concrete sense of how literary attention is divided amongst these clusters, consider **Figure 3.4**, which depicts the most reviewed authors for each cluster. Each of seven categories of periodical is represented by one central, color-coded node. Each author node is connected to every periodical node for which it is one of the fifteen most frequently reviewed authors. Author nodes are colored based on whichever journal node they connect to; for those authors who appear in the top-fifteen list for multiple clusters, they are colored based on whichever of the clusters reviewed them the most. Most authors appear in the top-fifteen list for only one category, but a few authors manage to make connections to multiple different categories. Fay Weldon is in the top fifteen for both major city newspapers and mainstream wide coverage publications. Philip Roth has a similar honor for each of conservative magazines, liberal-progressive magazines, and literary magazines.

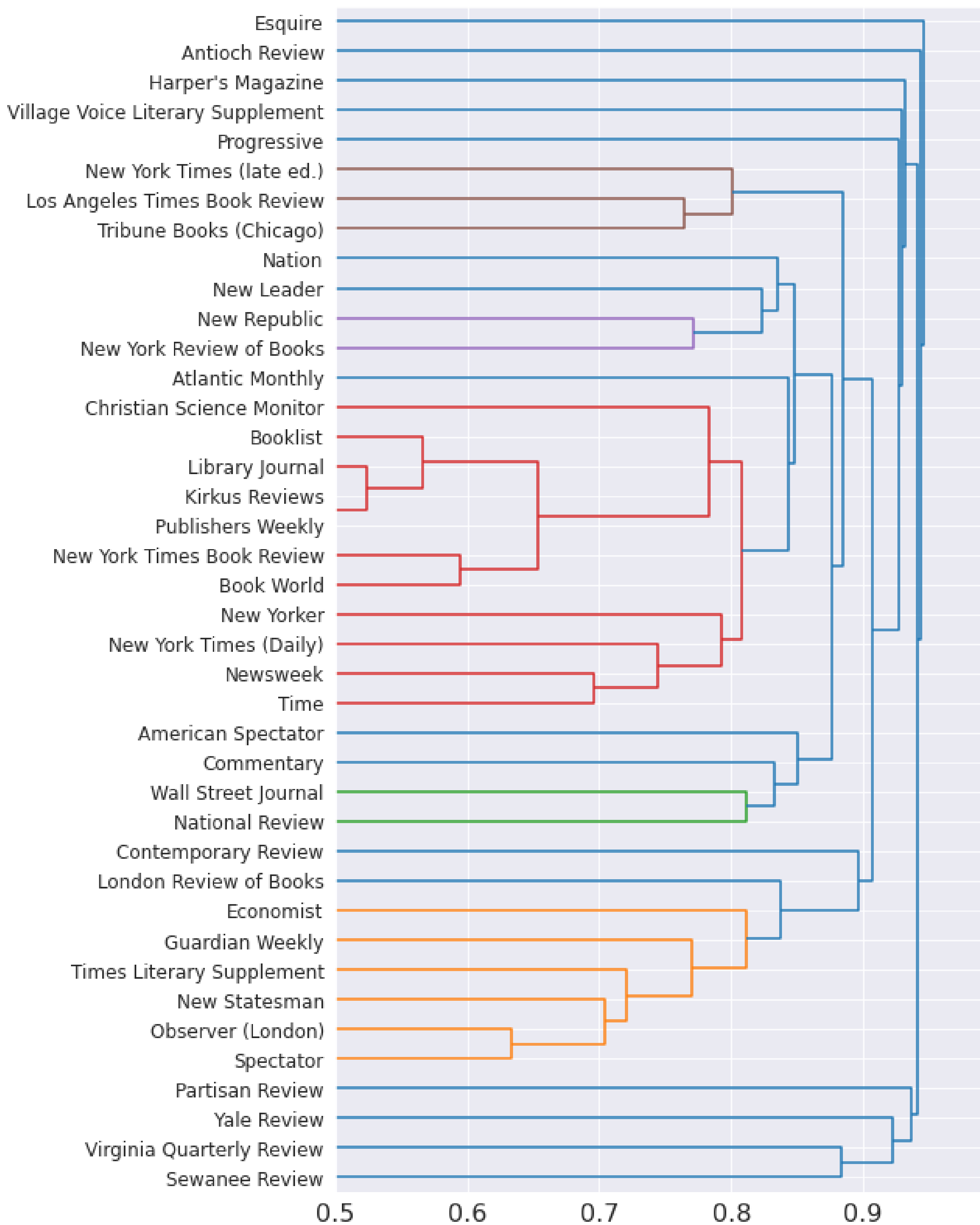

Figure 3.3. Mainstream branches of the journal dendrogram. The x-axis lists the cosine similarity breakpoint at which particular branches were merged into a single cluster.

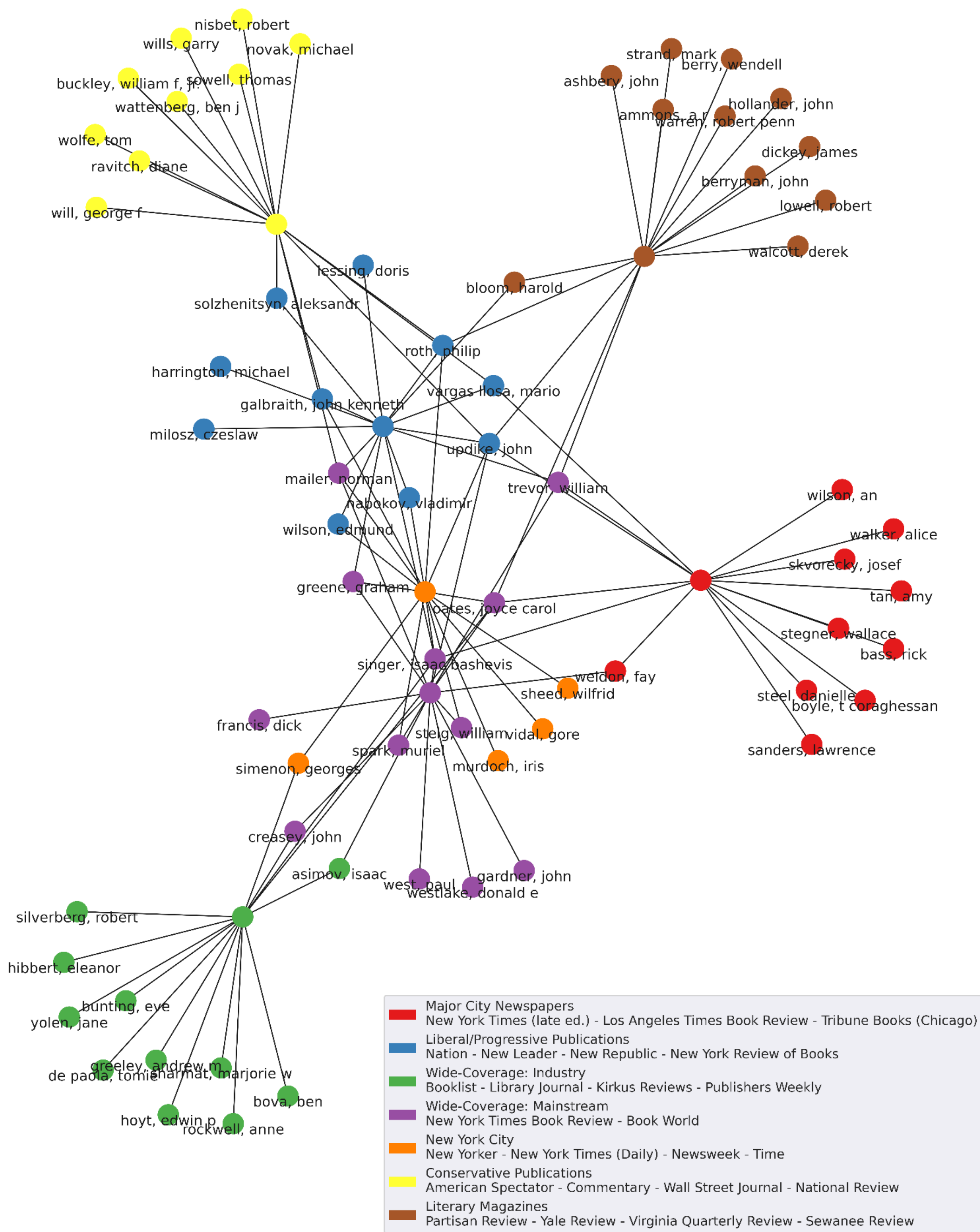

Figure 3.4. Network diagram of top 15 most-reviewed authors within selected clusters. Each cluster contains one central node, and each author node is connected to every central node for each cluster in which they are among the top 15.

This review network draws attention to the distinctiveness of each cluster's review profile. The overall network structure is one of isolated reviewing communities: the top-fifteen most frequently reviewed authors in four of the seven depicted clusters are almost entirely isolated. For instance, consider the authors reviewed by the cluster composed of literary magazines. Ten of fifteen of the authors most reviewed by literary magazines in this period are not in the top-fifteen list for any other cluster. All ten of these authors are poets, though several of them also wrote criticism, fiction, or plays. The absence of any substantial review presence for these authors outside of literary magazines is evidence of the social marginality of poetry in this period. More than any other genre, the field of poetry resembles Bourdieu's "restricted field" very closely; as a general rule, books of poetry simply aren't reviewed in newspapers, industry publications, or mainstream culture magazines. The insularity of poetry within the review world is surpassed only by academic publications. Indeed, even more than writers of fiction, poets in this period are very often academics themselves. One could argue that poetry is in this way very similar to a specialized academic discipline. At the very least, it occupies a homologous position within the reviewing field.

Another cluster that is highly "insulated" in its list of fifteen most-reviewed authors is the major city newspapers cluster, comprising the book reviews of the *New York Times*, the *Los Angeles Times Book Review*, and *Tribune Books* (of the *Chicago Tribune*). Unlike some of the other clusters, this cluster's top-fifteen list contains many authors who we would not recognize as "canonical" authors in the period. It includes at least two authors of genre fiction: Lawrence Sanders, a popular mystery author, and Danielle Steel, sometimes described as the most popular romance author of all time (though many of her books might be more aptly described as family tragedies). The cluster also features authors who might be called semi-canonical, such as Amy

Tan and T.C. Boyle. Though influential and well-regarded, neither author ever won a major national literary prize. Arguably the only truly canonical author who appears in only the top-fifteen list for major newspapers is Alice Walker. But even Walker's *oeuvre* is famous for finding an audience in both a popular and an establishment literary milieu.

Additionally, this cluster contains more women among its top fifteen than any other cluster in the mainstream category. I suspect that gender and demography go a long way towards explaining the coherence of the authors in this cluster. In terms of literary prestige, it is not clear what unites Danielle Steel with authors like Amy Tan and Alice Walker. The latter two authors are closer to the world of literary awards: Walker won both the Pulitzer and the National Book Award for *The Color Purple*, and, though Tan has never won either award, *The Joy Luck Club* was a finalist for both the National Book Award and the National Book Critics Circle Award. Steel, meanwhile, has never been nominated for a major national literary award. One possible explanation is that these publications are simply more willing to review books by women, perhaps because they have a readership with a greater proportion of women than some of the other publications featured in this dataset. This theory is purely speculative, though it becomes more plausible when we remember the claim by Stanton Peckham, editor of the *Denver Post*, that was quoted earlier in this chapter. Peckham suggested that many readers of newspaper book sections are women, remarking that one bad review he published had him "denounced" in every woman's club in town. If newspaper book sections are read by comparably more women than, say, literary magazines are, it would be unsurprising that women authors are more prominent in that region of the reviewing space.

In either case, the fact that only one cluster features a considerable number of women draws our attention to the importance of cultural politics for a consideration of reviewing practices. The decision to review a book is eminently political: it is an implicit judgement about whether a book is or is not worthy of the attention of a publication and its readers. This political dimension is most obvious in the schism between the conservative and liberal-progressive magazine clusters. Interestingly, the conservative magazines have comparatively little coverage of fiction; the majority of their top-fifteen most reviewed authors are primarily authors of nonfiction. This likely reflects the predominance of the intellectual leaders of the neoconservative movement, such as William Buckley, Jr., who primarily wrote nonfiction.[30] Theologians and religious historians are also important, such as Garry Wills and Michael Novak. The only novelist who appears exclusively in this cluster is Tom Wolfe. In the early days of the neoconservative movement, Wolfe was becoming famous for his satires of liberal culture, such his 1970 "These Radical Chic Evenings," in which he coined the phrase "radical chic." By the end of his career, Wolfe was pointedly positioning himself as someone outside the mainstream of the liberal New York culture in which he lived. In a 2004 interview, shortly before the release of his novel *I Am Charlotte Simmons*, Wolfe said, "[t]here is something in me that particularly wants it registered that I am not one of them."[31] He went so far as to say that he has "sympathy for what George Bush is trying to do," though he added that "obviously the excursion [into Iraq] [was] not going well." Wolfe's position within the reviewing network can be seen, then, as reflective of his explicit attempts to position himself outside and against the three clusters with the greatest amount of overlap in the network: the liberal-progressive cluster, the wide coverage - mainstream cluster, and the New York City cluster.

30. Buckley's presence here is notable, given that he helped found *National Review*, one of the four magazines that make up the cluster.

31. Ed Vulliamy, "'The liberal elite hasn't got a clue,'" *The Guardian*, Mon 1 Nov 2004.

This is a fairly stark example of what Bourdieu refers to as a “position taking,” a fundamental aspect of his model of the sociology of culture.[32] For Bourdieu, one’s position within a social field is only sensible relationally, in reference to other positions that one might take. If our review network is any indication, then the position against which Wolfe—and, perhaps, much of the conservative cluster writ large—is positioned can be summed up as “New York City,” or, perhaps more precisely, something like “New York liberal literary culture.”[33] These clusters are all closely related to each other—in fact, the orange cluster, which might be called “New York news and culture publications,” only has four authors exclusive to its own list. The remaining eleven of the authors on its top-fifteen list are shared with the liberal-progressive and wide-coverage clusters. Although the three clusters are united as a whole, it is notable that, internally, they are divided into tiers of popularity. The wide coverage cluster includes more authors of genre fiction, such as crime writers Dick Francis, John Creasey, and Donald E. Westlake. Meanwhile, more of the so-called “literary novelists” are reviewed by liberal-progressive publications. The New York news and culture publications split the difference.

What the discussions in this section and the previous section show is that reception looks very different when you consider where an author was reviewed, rather than simply how often they were reviewed. When an author’s reviews are decomposed into their constitutive parts, the fault-lines in the reception space fall into stark relief. One can easily see how journals are aligned into distinct reviewing communities, or even hostile reviewing camps, when you compare the authors who receive attention in conservative magazines to the authors who receive attention in

32. Bourdieu, 30.

33. Of course, the great irony of this position-taking is that all four of the journals in the conservative cluster are themselves based out of New York City. Thus, the schism between the conservative and liberal clusters is not reducible to the geography of their production (though the geography of their reception is another matter).

liberal-progressive magazines. It is not obvious how the tastes of these opposing clusters of journals can be summarized with a single metric of prestige; they have different standards of taste, and their social positions can only be fully understood in light of these differences. In the following section, I consider how an attentiveness to this conflictual aspect of prestige can inform future quantitative studies of literary reception.

3.4. An Alternative Model of Prestige

This clustering methodology provides an alternative way of modeling prestige in quantitative literary studies. Although quantitative projects on literary prestige have proliferated in recent years, prestige can be a slippery concept. And there is no single agreed-upon way to measure it. In literary studies, a popular definition of prestige or status comes from Pierre Bourdieu, who defines it as "the degree of recognition accorded by those who recognize no other criterion of legitimacy than recognition by those whom they recognize."[34] The definition seems to double inward and back on itself. But it is not a tautology; it just reflects the insularity and self-fulfilling quality of prestige. Any form of status or recognition only has meaning within the particular social group that confers it and continues to grant it meaning. One can, in theory, be highly prestigious within a particular social group but irrelevant or evenly low-prestige in a separate social group—especially if these social groups are situated in conflict with one another. Another way to put this is to say that prestige is intrinsically relational: one can only ever be prestigious with respect to the standards of some particular prestige-granting community. It

34. Bourdieu, 38.
Similarly, Bourdieu elsewhere in that same text writes that "the consecrated writer is the one who has the power to consecrate and to win assent when her or she consecrates an author or a work." Bourdieu,42.

doesn't make sense to speak of somebody as objectively or universally prestigious. For every category of prestige, there can be a person or group of people who just aren't very impressed by that type of prestige.

But this slippery, relational quality of prestige can be difficult to capture in quantitative projects. In the past decade, a number of projects in quantitative literary studies have explored different ways of operationalizing literary prestige, adopting proxy measures that allow researchers to compare the relative status of texts and authors in quantitative terms. Many of these projects were inspired by the terrain of literary prestige mapped by Bourdieu's "The Field of Cultural Production." In that piece, Bourdieu famously portrayed French literature as a field of conflicts for position carried out in two dimensions: an economic dimension, where success is measured in monetary profits or simply the size of one's audience; and a "consecration" dimension, where success is measured by the degree of recognition one achieves among those whose opinions are considered important in the field. For Bourdieu, these two dimensions run in tension with each other, as one's success in terms of consecration can often be increased by one's failure in economic terms. A popular novelist might be derided as a sellout or an unserious artist, whereas the stoic and even impoverished lifestyle of an unknown experimental writer might be celebrated as a commitment to art for art's sake. By plotting these dimensions on a simple grid, Bourdieu is able to create his famous two-dimensional visualization of the French literary field, which charts the positions of different literary genres and movements in terms of their relative financial success and degree of consecration, which is reproduced below.[35] In the bottom right is industrial art, reflecting big profits and low consecration; whereas the top left is "art for art's sake," reflecting high consecration but almost nonexistent profits.

35. Bourdieu, 49.

Bourdieu’s two dimensions of reception provide the framework for most recent quantitative projects on literary prestige. An early experiment in this vein was Pamphlet 11 of the Stanford Literary Lab, by Mark Algee-Hewitt *et al*, which set out explicitly to find a way of operationalizing Bourdieu’s framework.[36] In that project, the authors repeated Bourdieu’s analysis for late-eighteenth and early-ninteenth-century British literature, likewise mapping authors according to a prestige and a popularity dimension. They took the analysis a step further and assigned exact scores for each category. For popularity, they ranked authors by the number of times that author’s novels were reprinted or translated in the period, following the assumption—commonplace in scholarly bibliography—that popular novels were more likely to be reprinted. For prestige, they adopted an aggregate measure derived from the number of times an author appeared as the “primary subject author” in the MLA Bibliography and the length of that author’s entry in the Dictionary of National Biography. The result is a map of the ninteenth-century British literary field akin to Bourdieu’s, except that the distance between entities on the map has a specific quantitative meaning. The addition of a meaningful scale of reference yields interpretive dividends, such as the observation that Walter Scott’s novels put him in a world apart: one of only a few outliers in terms of prestige and without equal in terms of popularity.

Other projects have experimented with different measures of the relative prestigiousness or influence of canonical works. Ted Underwood and Jordan Sellers use book reviews in elite periodicals to compare rates of word-use in reviewed and non-reviewed English-language poetry between 1820 and 1919.[37] In this project, the authors trained a predictive model to differentiate

---

36. Mark Algee-Hewitt *et al*, “Canon/Archive. Large Scale Dynamics in the Literary Field,” *Pamphlets of the Stanford Literary Lab*, January 2016. https://litlab.stanford.edu/LiteraryLabPamphlet11.pdf

37. Ted Underwood and Jordan Sellers, “The *Longue Durée* of Literary Prestige,” *Modern Language Quarterly* 77:3 (September 2016).

between poetry collections that were and were not reviewed by elite journals, as a way of testing whether prestigious poetry had its own specific diction in different segments of their century-long period. This method led to the counterintuitive result that, not only was the diction of elite-reviewed poetry coherent at a statistical level, but it was also coherent over the entire one-hundred-year period. According to their experiment, elite reviewers tended to prefer poetry with "concrete language" that did not have a "sentimentally uplifting tone," such as poetry that deals with explicitly religious or political themes. [38] This preference was a relatively stable one for an entire century.

In both of these projects, as in Bourdieu, prestige is one-dimensional. An individual author's prestige can vary in magnitude, but not in kind. You can only have more or less of it and any one unit of prestige is the same as any other unit. Though there is nothing intrinsically wrong with modeling prestige in this way, it does limit the specificity of the claims about the literary field that one can make. More recent projects have sometimes acknowledged this limitation. In a project that maps popularity on Goodreads against citations in the MLA bibliography, J.D. Porter describes his portrayal of the literary field as only one among many possible ways of operationalizing prestige. As he notes, both the MLA Bibliography and Goodreads data are heavily skewed towards the present; each provides "a sense of the canon as it is today, rather than a picture of how it got that way."[39] Alternative methods of measuring canonicity or prestige might consider metrics more contemporaneous to the authors themselves, such as reprints and translations, historical library catalogues, or sales data. Nevertheless, a comparison of the popularity of an author among literary critics to that same author among Goodreads users can tell

38. Underwood and Sellers, 332.

39. J.D. Porter, "Popularity/Prestige: A New Canon," *Pamphlets of the Stanford Literary Lab*, Pamphlet 17, September 2018. https://litlab.stanford.edu/LiteraryLabPamphlet17.pdf

us much about the overlaps and distinctions in the attention of those two groups. Whereas Bourdieu seems to imply it is not actually possible to be both a consecrated and a commercially-successful author, Porter's experiment captures the fact that there are some authors—such as Shakespeare, Dickens, and Austen—who are widely read by both an "elite" and a "mainstream" audience.

As Porter points out, alternative measures of canonicity face a problem of data availability; the MLA Bibliography is readily accessible, unlike digitized versions of historical library catalogues. But even if we had access to data of this type, it would still only present a partial picture of the "canon." In other words, it is not obvious that it is always meaningful to speak of "the" canon in the first place, or, analogously, to describe "cultural capital" or "prestige" as a one-dimensional quantity that varies only in magnitude. Any measure of the popularity or esteem of a given author is necessarily a measure of that author's reception within a specific reading community. Strictly speaking, it is not entirely accurate to say that Shakespeare is both popular and prestigious; it would be more precise to simply say that he is widely talked about by literary critics and Goodreads users alike, whose interests may or may not align with other similar communities of reception. This is an obvious point, but I think that it is worth dwelling on its implications. Prestige or "cultural capital" is an intrinsically relational concept that only has any meaning in the context of the specific social group that recognizes it as valuable. Any description of something as "prestigious" can be followed up by the question, "for whom?" As James English puts it in a history of cultural prizes, any cultural "asset" valuable in a specific social or institutional context may be "relatively worthless" in a different context: "a Harvard philosophy degree or a demonstrable mastery of Sanskrit may be counted as capital in the academic world," but both are unlikely to carry the same cachet in the military or the world of “commercial activity.”[40] In Bourdieu’s sociological system, every social field has its own criteria for legitimation and contestation, by virtue of the fact that it is recognized as a distinct field in the first place by those operating within it.

Additionally, the boundaries of social fields are fuzzy. Although sociologists of culture generally speak of fields as if they are stable objects (something has value in the literary field), in theory one can always subdivide any social grouping into more specific and fine-grained subfields. Shakespeare’s position in Porter’s analysis would likely shift if we restricted MLA Bibliography citations to scholars working specifically in postcolonial literature, or Romanticism. Likewise, the most popular authors on Goodreads at large are probably very different from the most popular authors among Goodreads users who are self-proclaimed fans of science fiction. Cultural capital is a moving target, only meaningful within a specific slice of the population of readers (though, as Underwood and Sellers show, in practice, influential social groups may have a powerful ability to reproduce themselves and their standards over time). Cultural capital only exists insofar as those who believe in its value continue to do so. But this intrinsically-relational and infinitely-subdividable quality of social status is not well-captured by one-dimensional metrics. In fact, I would argue that Bourdieu’s own diagram does not capture the spirit of his own sociological system. It portrays recognition as a singular quantity, something applicable to the literary field as a whole. It is not obvious, for example, that vaudeville performers and audience members will agree with poets about the consecratedness of Symbolist poetry.

---

40. James F. English, *The Economy of Prestige: Prizes, Awards, and the Circulation of Cultural Value* (Cambridge, MA: Harvard University Press, 2005): 9.

Is there a way to capture this quality quantitatively, the fact that actors in different segments of the literary field may not agree on the prestige of a work of culture? The clusters identified earlier offer one way of doing so. Instead of treating prestige as a singular quantity we can portray prestige as multiple distinct quantities by segmenting the data into clusters of like-minded schools of reviewers. There are, in other words, many "prestiges"— one can be highly prestigious in the "conservative magazine" cluster without necessarily being prestigious in the "major city newspapers" cluster, and so on. By doing so, one can recreate the prestige plots surveyed above, except with different prestige-granting communities mapped along each axis. Three such plots appear below, which show varying degrees of correlation between authors' total reviews in two different clusters. The first plots New York City against Wide-Coverage: Mainstream, and has a moderate level of correlation (R^2 = 0.546). The second plots Liberal/Progressive Publications against Conservative Publications, and has a lower level of correlation (R^2 = 0.318). The final one plots Major City Newspapers against Literary Magazines, and has a very low level of correlation (R^2 = 0.08). Annotations for a few exemplary authors have been added as illustration.

The first plot shows an example of two "axes" of prestige that are closely related: authors who were frequently reviewed by New York culture magazines like *New Yorker* and *Time* were also likely to be reviewed in Wide-Coverage: Mainstream magazines, which include the *New York Times Book Review* and *Book World*. These two different clusters of magazines seem to be working with a similar definition of prestige; the most successful authors along these axes, those in the top-right corner of the plot, are authors whose work might be described as "mainstream literary," in the sense that they are authors who have won literary awards but who also have relatively mainstream name-recognition. Norman Mailer and Joyce Carol Oates are stand-outs in this regard: they both won major national awards (Mailer a Pulitzer, Oates a National Book Award) but also have more of a mainstream profile than somebody like Cormac McCarthy would have had during this time. There are exceptions, such as science fiction writer Gene Wolf, who was frequently reviewed by the wide-coverage mainstream magazines but never once reviewed by the New York magazines. But in general, reviews in these two clusters closely go hand-in-hand.

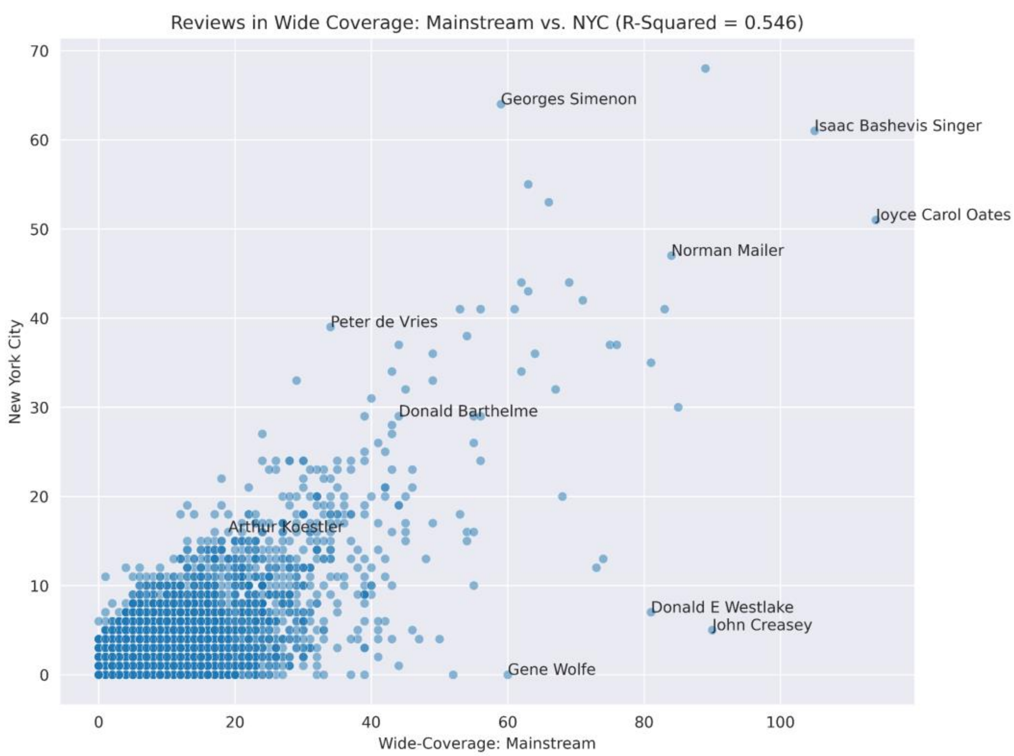

Figure 3.5. Authors' reviews in NYC magazines and mainstream wide-coverage periodicals.

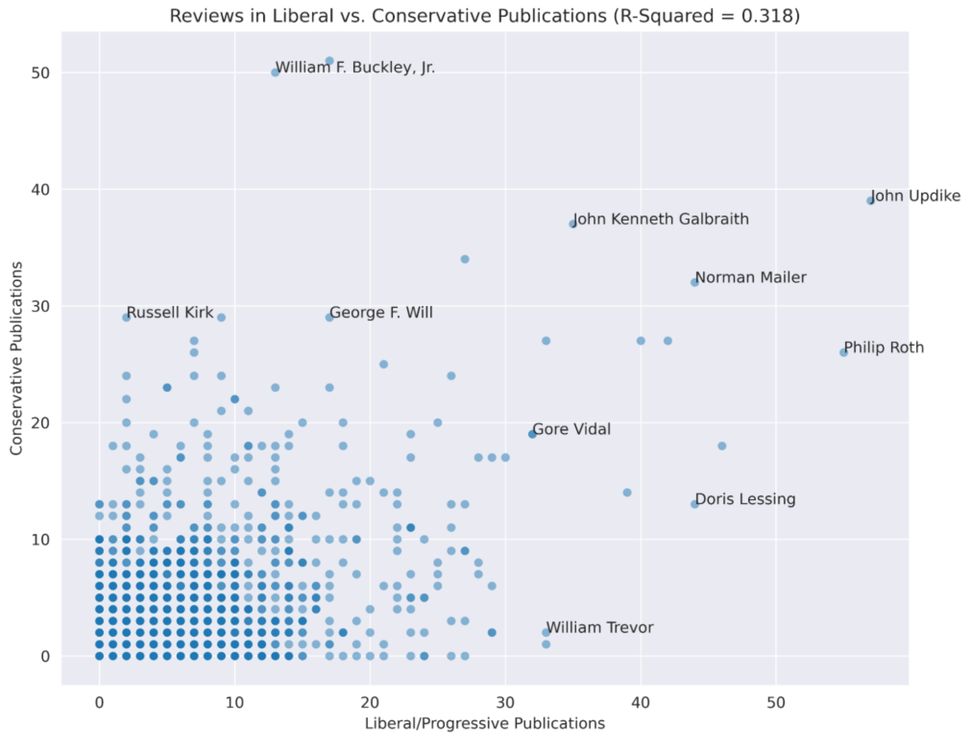

Figure 3.6. Authors' reviews in conservative publications and liberal/progressive publications.

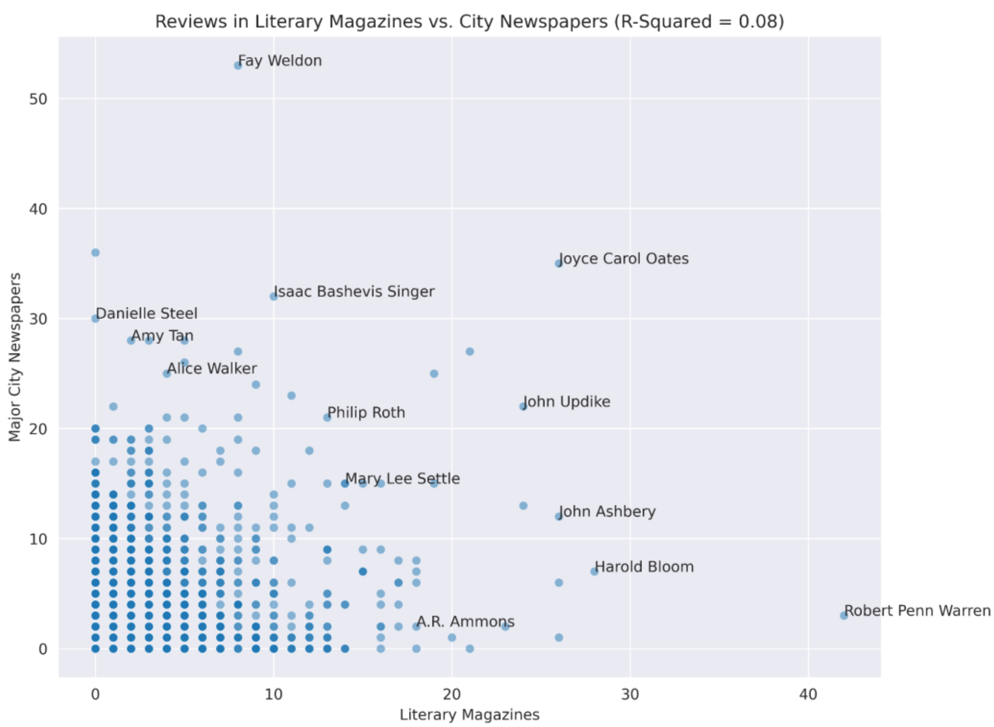

Figure 3.7. Authors' reviews in literary magazines and major city newspapers.

The correlation is a fair bit weaker in the next plot, which compares reviews in liberal and progressive magazines with those in conservative magazines. With an R^2 of 0.318, the correlation here is slightly more than half as strong as that in the previous comparison. There is some overlap between the two clusters: an author with many reviews in one cluster is likely to have at least some reviews in the other. Those who fall along the diagonal are authors who are political "compromises," people whose work received commentary from both conservatives and liberal-progressives. This includes novelists like John Updike and Norman Mailer, as well as public intellectuals like Gore Vidal and J.K. Galbraith.

In comparison to the previous graph, the disagreement between the two axes is more readily apparent. There are more datapoints scattered along the edges, authors who received a significant amount of attention from one cluster without getting a comparable amount of attention from the other. The identities of these "divergent" authors say much about the tastes of their respective clusters. Disproportionately conservative authors include conservative political theorists like Russell Kirk and William F. Buckley, Jr. Many of the authors in this category are, like these two, authors of essays or nonfiction. Their counterparts on the liberal side include non-American literary authors like Irish novelist and playwright William Trevor and British novelist Doris Lessing. There is a clear difference in focus between the two magazine clusters. The conservative magazines are focused on political theory, often nonfiction, which perhaps reflects the fact that several of these magazines were founded with the goal of fostering a political and intellectual movement; Buckley, for instance, founded *National Review* with the explicit intention of giving a unified voice and platform for the New Right.[41] The liberal-progressive magazines, by contrast, are interested in fostering a cosmopolitan or at least international Anglophone literary culture. It does not seem like prestige in one cluster automatically translates to prestige in the other; in fact, the "prestiges" of the two clusters are at least partially in conflict. Though the two axes are not perfectly orthogonal, they are clearly distinct.

This quality is most clear in the final chart, where the two axes have very little correlation at all. Here, there are many authors who were widely-reviewed in one cluster without being reviewed in the other cluster: reviews in literary magazines like *Yale Review* and *Partisan Review* do not automatically translate into reviews in major city newspapers like the *Los Angeles Times Book Review* or *Tribune Books*—and vice versa. Though some authors fall along the diagonal, there are almost as many of them who do not. Frequently-reviewed authors in literary magazines but not newspapers include many poets, such as A.R. Ammons and John Ashbery; across the aisle, in the opposite category, are popular novelists like Danielle Steel, as well as authors who might be called popular-literary, such as Alice Walker and Amy Tan. These are highly distinct reviewing spaces: the newspaper market has more women authors, as well as more authors whose work targets a "mainstream" audience. This plot probably corresponds most closely to the traditionally-Bourdieusian dichotomy between autonomy and heteronomy, "restricted field" and mass market.

---

41. In the magazine's original mission statement in 1955, Buckley describes *National Review* as "out of place because… literate America rejected conservatism in favor of radical social experimentation." Among the magazine's convictions, Buckley lists a number of ideas that would become closely associated with neoconservatism in the ensuing decades, including anticommunism, opposition to government expansion, skepticism towards "world government" as embodied in the United Nations, and a defense of society's "organic moral order." William F. Buckley, "Our Mission Statement," *National Review*, November 19, 1955. https://www.nationalreview.com/1955/11/our-mission-statement-william-f-buckley-jr/

What is clear in these three visualizations is that there is not a single quantity "prestige" that is fully captured by any of these metrics. The criteria by which *National Review* determines whether an author's work is worthy of comment is clearly different from the criteria used by *The Nation,* or by *Tribune Books*. No single one of these clusters of magazines has the "true" or objective standard of prestige; in fact, it might be most accurate to say that it is the nature of prestige that it tries to objectify itself, to persuade other would-be granters of respect that its own standards of respect-granting are best. But this process is often conflictual. Writers and critics may have conflicting visions of prestige and are frequently unwilling to adopt the standards of their intellectual rivals. Sometimes they create rival publications for the express purpose of challenging the intellectual influence of their opponents. One-dimensional measures of prestige struggle to capture this intrinsically conflictual quality of prestige. By modeling prestige as a multidimensional quantity, we can bring quantitative methods to bear on the sociology of literature without repressing this essential conflictual quality.

3.5. Conclusion: Clusters as Canons

In this chapter, I have tried to map out the dominant contours of the book review market in American letters between 1965-2000. The clustering methodology I adopt to do so provides a way of discovering journals with shared taste, understood as journals that tend to review the same authors and the same books as each other. As I have shown, journals' overlaps in book review choices reflect intuitive divides in the literary field, including divides around political affiliation, genre, market size, and geography. Clustering journals according to shared taste provides a useful macroscopic perspective from which to inspect where authors lie within the status cleavages of the field.

Another way to think about these discovered clusters is that they represent literary and intellectual canonicity that is limited in scope to specific reading communities. I think this quality of the discovered clusters is clearest in the case of the conservative cluster. Some of the most prominent authors in this cluster, in terms of total review count, are people who are now remembered as the intellectual forerunners of the American neoconservative movement, such as the authors William F. Buckley and Russell Kirk. Although no literary critic would ever think to call them canonical, they are nevertheless authors that a certain brand of conservative intellectual might consider "essential reading." Their work would likely receive reference in any intellectual history of American neoconservatism.

This clustering methodology, then, provides one method for thinking about the sociology of canon formation. Canons are, I think, one of the most fundamental and ordinary methods by which communities of readers manage the problem of abundance. Since nobody can read everything, readers appeal to the tastes of other like-minded readers—political allies, intellectual comrades, or simply other fans of the same authors—when choosing what to read. If a reading community manages to sufficiently universalize its canon, such as by institutionalizing it in school reading lists, then people may begin speaking of it as the canon of a culture. Then it has become a canon in sense in which this term was used during the so-called canon wars of the 1990s. This is very close to the definition of canonicity provided by Samuel Delany, who defines it as the whole process by which specific lists of canonical authors compete with each other for legitimacy:

> The canon is not a natural object… it is a social object… Though from time to time the canon presents itself as a paradigm, this is merely a flattened representation of a complex system, of a rhizome, of a syntagm, or simply of an abstract set of interrelations, too rich to be mappable with any sophistication in

> less than three—and more likely four—dimensions… [The canon is] the discursive machinery that produces the many contesting lists involved.[42]

Canons in the more general sense of the many competing collections of texts or authors considered interesting or valuable by various communities of readers are an ordinary part of everyday literary sociology. Though we can do away with a singular canon, with the canon, canons in the more general sense are probably an inevitable part of thinking in common. But one needn't come out in support of this or that canon in particular in order to study how they form, or how conflicts between competing canons emerge and are resolved.

---

42. Samuel Delany, *Shorter Views: Queer Thoughts and the Politics of the Paraliterary* (Middletown: Wesleyan University Press, 1999): 186-7.

Coda: On Literary Data

In this dissertation, I have given an account of three institutions which, between 1931 and 2009, were used by readers, publishers, and booksellers to manage the problem of abundance in the literary market, which refers to the fact that too many texts are in print at any one time for anybody to be able to consider all of them when making decisions about what to read or what to publish. These three institutions all manage abundance in different ways, and for social actors who occupy different occupations and social positions in the literary market. Bestseller lists manage abundance primarily on behalf of readers, by reporting information on a small sample of currently-popular books. As we saw, by doing this, bestseller lists do not just describe and report on the market but actively intervene in it. They provide valuable publicity to the authors who make it onto its pages. For this reason, booksellers and publishers adjust their behavior in response to the contents of the list, by providing discounts for bestselling titles or by prioritizing previously-bestselling authors when deciding what to publish. The brand-name author, meanwhile, manages abundance on behalf of both readers and publishers. By focusing marketing and publicity on books by authors whose names are already well-known, publishers entice readers to engage in a form of repeat consumption that makes publication more predictable. Finally, book reviews manage abundance for their own readers, who read their review sections to get news and commentary about the current state of some aspect of literary culture. Insofar as the market is too large and complex for all readers to read all reviews, reviewing publications generally target their reviews at some smaller, socially-meaningful segment of the whole market, which may correspond to political ideology, taste, geography, or some other category.

One way of thinking about these three institutions is that they are institutions for organizing literary reception. Although this would be an accurate and perfectly reasonable way of conceptualizing these institutions, I have intentionally avoided framing this dissertation as a project about literary reception. In my opinion, computational literary studies has often relied too heavily on making a conceptual distinction between literary production and literary reception or consumption. Such distinctions have been especially prominent in methodological debates about the appropriateness of using purely textual data for arguments about literary history. In response to Katherine Bode's critique of computational literary studies from the perspective of book history and bibliography, for example, Ted Underwood suggests that the difference between the textual approach and the book-historical approach is simply one of focus.[1] As he puts it, projects interested in "literary production" will necessarily use different methods than those focused on "reception," and there is "no reason to make a final choice" between these different methods. Laura B. McGrath interprets Bode's critique in similar terms. She describes Bode's own methods as capturing "not literary production, but reception and circulation," in contrast to "[m]any

1. "Scholars interested in literary production may want to approximate completeness, while scholars interested in reception prefer to focus on a subset of influential works. Some social questions hinge on the demographic identity of writers, others on readers… All of these sampling strategies have their uses, and there is no reason to make a final choice between them."

Ted Underwood, "A Genealogy of Distant Reading," *Digital Humanities Quarterly* 11:2 (2017).

For the piece to which this passage responds, see Katherine Bode,"The Equivalence of 'Close' and 'Distant' Reading; or, Toward a New Object for Data-Rich Literary History," *Modern Language Quarterly* 78, Vol 1 (2017): 77-106.

Similar rhetoric appears Underwood, Ted, Patrick Kimutis, and Jessica Witte. "NovelTM Datasets for English-Language Fiction, 1700-2009." *Journal of Cultural Analytics,* Vol. 5, Issue 2 (May 28, 2020).

For example: "…we do not think any single dataset will be universally valuable for all research questions. Some scholars are interested in literary production; others are interested in reception…" Underwood *et al*, 1.

distant readers," who "emphasize literary production."[2] Studies of production may be more justified in relying primarily or entirely on textual data by virtue of their research question.

Although I have no problem with computational analyses of textual evidence in itself, I think that this framing sidesteps the force of Bode's critique. The challenge facing computational text analysis is not just conceptual, a matter of expanding the conceptual frame to include data on circulation. Rather, the problem is that any data that one might plausibly use to ask questions about literary production that took place in the past has itself been mediated by years, decades, or even centuries of reception. Because all literary works are not equally likely to be widely printed, they are also not equally likely to have survived into the present. So, any dataset used for asking questions about literary production might exhibit trends or characteristics that are better accounted for by historical phenomena that we might want to call "reception." The first chapter of this dissertation provides a useful case study for this point. Although I modeled a trend that is textual in nature—the decline of topics relating to historical and domestic fiction, and the rise of topics relating to crime fiction—my explanation for the trend relied on reception. Or, to be even more precise, the trend cannot be accounted for except by considering production and reception as a single process. Publishers were moving paperback authors into hardcover: production. This, in turn, meant that mass market paperback genres like thriller were newly popular in hardcover outlets like bookstores and were achieving higher hardcover sales numbers: reception. Making a distinction between production and reception would obscure some part of the total social process.

I hope, then, that one thing readers can take away from this dissertation is a sense of the analytical power that can come from seeing the book-historical or so-called "reception" side evidence as an intrinsic part of computational literary studies, rather than as something distinct.

---

2. Laura B. McGrath, "More Specific, More Complex," *Post45*, May 8, 2019.

Literary data does not describe the literary past evenly or randomly. Literary history is itself a sort of sieve or filter that affects which data survives long enough to make it before the eyes of the literary historian. Every archive from which textual data might be derived has already passed through this great filter. It therefore reflects the biases of history itself, relating to historical reprinting practices, norms of archiving, and even the material qualities of different kinds of books (hardcover books, for instance, are more durable than paperbacks, and are therefore more likely to survive). Of course, this challenge is hardly unique to computational literary studies, but affects all disciplines that make use of historical evidence. Evidence about human prehistory is biased towards social groups that constructed stone buildings and groups that made a habit of preserving evidence of their lives in tombs or burial grounds. Likewise, the fossil record is biased towards matter like bones that is more likely to fossilize. It would be the height of absurdity to suggest that this means there is no point in studying archaeological sites or fossils in a systematic way. But it would be similarly absurd to make claims about historical evidence without accounting for these processes of selection and preservation. In reading this dissertation, I hope that the reader has come away with a renewed sense of the analytical utility of considering these historical processes as themselves an important object of study. Since evidence about literary production is itself structured by the history of literary reception, making a distinction between these two concepts is not always meaningful or useful.

Appendix

This section provides a description of the data used in Chapter Three and the process by which it was generated. Initial data was generously provided to me by Richard So and Daniel Sinykin, who extracted the text from the *Book Review Index* with OCR methods for two unrelated projects.[1] These projects only worked with a restricted portion of the data. Richard So, for instance, selects only the top 1% most frequently-reviewed English-language books written by American authors — a total of 1,776 books. These books represent one kind of literary elite; as So puts it, they are "the novels that commanded the greatest amount of 'literary attention.'"[2] Working with a restricted set allows So to make more fine-grained claims about the authors and books in his study. For example, he observes that a full 90%of those frequently-reviewed books were written by white authors.[3]

The data as provided to me came in the form of partially-parsed text files derived from optical character recognition (OCR) processed raw image scans of six cumulation volumes of the *Book Review Index* published by Gale. The volumes cover the following periods: 1965-1984; 1985-1992; 1993-1997; and annual cumulations for the years 1998, 1999, and 2000. The OCR process by which text data was extracted from page images did not recover very much of the logical structure of the *Book Review Index* itself, besides page breaks and line breaks that inconsistently mark the end of individual entries. In the best case, a segment of the raw text data might look something like the following excerpt. Note that this excerpt has been clipped to the

1. Richard So, *Redlining Culture: A Data History of Racial Inequality and Postwar Fiction* (New York: Columbia University Press, 2020).

Daniel Sinykin, *The Conglomerate Era* (New York: Columbia University Press, Forthcoming).

2. So, 70.

3. So, 70.

beginning and end of a single entry; in the real data, the boundaries of an entry are not explicitly marked.

CHOMSKY, Noam - The Political Economy Of
Human Rights
Mac - v93 - Ag 18 '80 - p37
NS - v99 - Ap 25 '80 - p623
Nat - v230 - F 16 '80 - pl81
New R - vl83 - S 6 '80 - p37
W - v25 - Je 23 '80 - p35

Although individual fields in an entry are not marked, they follow a consistent order: author name, book title, book reviews. Individual book reviews contain a number of dash-separated fields: journal abbreviation, volume number, issue number or season and year, page number, and, in later editions, a number summarizing the length of the review itself. The majority of entries, however, are not as clean as the excerpt above. More than half of scanned pages did not recover consistent line-breaks separating individual entries. Unwanted boilerplate text such as the title of the volume sometimes appears unmarked in the body of the text itself. Additionally, as with any OCR-derived text, rare but consistent spelling errors are randomly scattered throughout the text, including character substitutions, additions, and deletions. These errors include missing or transposed delimiters (the dashes seen above), which renders the start and endpoints of a field ambiguous.

The unstructured nature of this data makes it difficult to develop a reliable rule-based method for parsing individual fields, and impossible to develop a perfect one. For volumes published after 1985, the addition of a bracketed page-length field in each review entry considerably simplified the parsing process; these volumes came to me having already been parsed by previous researchers using a regular-expression based method.[4] For the longest volume, however, covering the years 1965-1984, I instead used a mixed procedure that combined statistical methods with rule-based string processing methods. First, I randomly sampled over 16,000 ordered tokens from the data and labeled by hand whether they belong to an author field, title field, or review field. Then, using the Python-based natural language processing library Flair, I trained a sequence labeler to automatically tag text according to that schema.[5] I chose the Flair library over more popular libraries such as spaCy for the fact that Flair's language models use character-level rather than word-level embeddings, which are better suited to the abbreviation-heavy style of the text in the *Book Review Index*. The F-1 score of the resulting model on a held-out validation set made up of 15% of the hand-labeled data was .994 for review fields, .872 for title fields, and .935 for author fields. In layman's terms, this means the model has a near-perfect ability to identify which sections of text describe a book review, but sometimes labels authors and titles incorrectly. As such, I exclusively relied on the review tags for disaggregating entries, exploiting the fact that reviews for a single book are always contiguous. I only used the author and title tags for distinguishing between author and title fields within an individual entry. It follows that the majority of parsing errors are to be found within individual entries, such as a token from a title field mistakenly included in an author field. Importantly, errors of this type have no effect on the journal clustering portion of my analysis, which is based only on relationships between entries. However, these errors could affect the results of my analysis in the portion on authors.

---

4. See So, 67-105 for a description of this data generation process.

5. Akbik, Alan and Bergmann, Tanja and Blythe, Duncan and Rasul, Kashif and Schweter, Stefan and Vollgraf, Roland, "FLAIR: An Easy-to-Use Framework for State-of-the-Art NLP," *NAACL 2019: 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations)*, 54-59. 2019.

As noted in the body of the chapter, the extraction process yielded entries for 1,330,480 books, which, when excluding titles without a listed author, leaves a working total of 1,218,257 books. The next step was to disaggregate individual review entries into entries for individual book reviews, as well as the additional information associated with that review, such as volume, journal, and issue of publication. I used Flair for this process as well, training a separate sequence labeler on ~2,700 tokens labeled for the purpose. Likely owing to the significantly greater level of regularity in the review fields than in the author and title fields, this labeler had greater than 99% accuracy for all fields except the month and day of publication, for which it was 98% and 94% accurate, respectively. Next, journal abbreviations reported by the *Index* were mapped to the full journal name that they represent. Due to OCR errors, this was not a perfect process and required using mostly heuristic, *ad hoc* approaches. For example, *Booklist*'s tag "BL" sometimes mistakenly appeared as "8L." Nevertheless, since these errors are relatively systematic, for each volume, between 97% and 99% of parsed reviews were successfully matched to a corresponding journal name in the *Book Review Index*. In total, with these methods, I was able to recover 3,355,879 book reviews from the data.

Bibliography

Adler, Bill and David R. Slavitt. *The Agent.* New York: Doubleday, 1986.

Adler, Bill and Thomas Chastain. *Who Killed the Robins Family?*. New York: William Morrow, 1983.

Advertisement. *Publisher's Weekly*, February 8, 1980: 32.

Algee-Hewitt, Mark, Sarah Allison, Marissa Gemma, Ryan Heuser, Franco Moretti, and Hannah Walser. "Canon/Archive. Large Scale Dynamics in the Literary Field." *Pamphlets of the Stanford Literary Lab*. January 2016.

American Book Producers Association. "About Us." https://abpaonline.org/about-us/.

Anthony, Carolyn. "New Players in the Mass Market." *Publisher's Weekly*, June 27, 1986.

Anthony, Carolyn T. "Are Reviewers Fair To Trade Paperbacks?" *Publishers Weekly*, October 15, 1982.

Appel, Allen. "Remembering Bill Adler Sr." *Publishers Weekly*, April 4, 2014.

Akbik, Alan, Tanja Bergmann, Duncan Blythe, Kashif Rasul, Stefan Schweter, and Roland Vollgraf. "FLAIR: An Easy-to-Use Framework for State-of-the-Art NLP." *NAACL 2019: 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations)*, 54-59. 2019.

Baer, Beverly, ed. Beverly Baer, ed. *Book Review Index: A Master Cumulation 1993-1997.* Detroit: Gale Research, 1999.

—. *Book Review Index: 1998 Cumulation.* Detroit: Gale Research, 1999.

Bakker, Gerben. *Entertainment Industrialized: The Emergence of the International Film Industry, 1890-1940.* Cambridge: Cambridge University Press, 2008.

Barbato, Joseph. "A Cheerful 40th Birthday for Bantam." *Publisher's Weekly*, October 18, 1985.

Barthes, Roland. "The Death of the Author," in *Image, Music, Text*, translated by Stephen Heath. London: Fontana Press, 1977.

Beech, Dave. *Art and Value: Art's Economic Exceptionalism in Classical, Neoclassical and Marxist Economics.* Chicago: Haymarket Books, 2016.

Benet, Lorenzo. *The Lives of Danielle Steel: The Unauthorized Biography of America's #1 Best-Selling Author*. New York: St. Martin's Press, 1995.

Bently, Lionel and Laura Biron. "The Author Strikes Back: Mutating Authorship in the Expanded Universe." In *Law and Creativity in the Age of the Entertainment Franchise*, edited by Kathy Bowrey and Michael Handler. Cambridge: Cambridge University Press, 2014.

Bloom, Harold. *The Daemon Knows: Literary Greatness and the American Sublime.* Spiegel and Grau: New York, 2015.

Bode, Katherine. "The Equivalence of 'Close' and 'Distant' Reading; or, Toward a New Object for Data-Rich Literary History." *Modern Language Quarterly* 78, Vol 1 (2017).

Bourdieu, Pierre. *The Field of Cultural Production: Essays on Art and Literature*. New York: Columbia University Press, 1993.

Britt, Ryan. "How George Lucas Made a Young, Anonymous Author Rich." *Inverse – Entertainment*. May 4, 2017.

Buckley, William F. "Our Mission Statement." *National Review*, November 19, 1955.

Burkett, Paul. "Nature's 'Free Gifts' and the Ecological Significance of Value." *Capital and Class*, Vol 23 (1999): 89-110.

Chrichton, Jennifer. "Warner Books: a Love Affair with the Bestseller." *Publisher's Weekly,* July 13, 1984.

Coser, Lewis A, Charles Kadushin, and Walter W. Powell. *Books: the Culture and Commerce of Publishing.* New York: Basic Books, 1982.

Crim, Kathryn. "Marx, *Silk Poems*, and the Pretext of Qualities." *Representations* 151 (2020): 96-126.

Dahlin, Robert. "Is There a Robot in Your Future? Yes, There Is. At least in a Harvest/HBJ Spring Paperback." *Publisher's Weekly*. December 5, 1977.

Davis, Kenneth C. *Two-Bit Culture: The Paperbacking of America.* Boston: Houghton Mifflin, 1984.

Deighton, John. "Marketing James Patterson." *Harvard Business Review*, Case Study, 2004.

Delany, Samuel. *Shorter Views: Queer Thoughts and the Politics of the Paraliterary*. Middletown: Wesleyan University Press, 1999.

Denning, Michael and Holly Allen. "'Who's Afraid of Big Bad Walt?' Disney's Radical Cartoonists." In Michael Denning, *The Cultural Front: The Laboring of American Culture in the Twentieth Century*. New York: Verso, 2010: 403-422.

Doyle, Jack. “The Love Story Saga, 1970-1977,” *PopHistoryDig.com*, June 29, 2011.

Driscoll, Beth. *The New Literary Middlebrow: Tastemakers and Reading in the Twenty-First Century.* Palgrave Macmillan, 2014.

Edge, Simon. “James Patterson - the bestseller who doesn’t write his own books.” *Express*, February 26, 2013.

Editors, “Big Gains for New Players.” *Publishers Weekly*, April 19, 2010.

Editors, “Fewer Titles = Good News: Some Call for More Cuts in Output.” *Publishers Weekly*, May 15, 2006.

Editors, “The Opening Editorial.” *The New York Review of Books*, February, 1963.

English, James F. “Now, Not Now: Counting Time in Contemporary Fiction Studies.” *Modern Language Quarterly* 77: 3 (September 2016): 395-418.

—. *The Economy of Prestige: Prizes, Awards, and the Circulation of Cultural Value*. Cambridge, MA: Harvard University Press, 2005.

Escarpit, Robert. *The Book Revolution.* London: George G. Harrap and Co; Paris: United Nations Educational, Scientific, and Cultural Organization, 1966.

Estate of Andrews v. United States. 850 F. Supp. 1279 (E.D. Va. 1994).

Finlayson, Ann. “Why We Call It Slush.” *Publishers Weekly*, May 16, 1986.

Fisk, Catherine. “Authors at Work: The Origins of the Work-for-Hire Doctrine.” *Yale Journal of Law & the Humanities* 15: 1 (2003): 1-70.

—. *Working Knowledge: Employee Innovation and the Rise of Corporate Intellectual Property, 1800-1930*. Chapel Hill: University of North Carolina Press, 2009.

—. *Writing for Hire: Unions, Hollywood, and Madison Avenue.* Cambridge: Harvard University Press, 2016.

Foucault, Michel. “What is an Author?” in *Textual Strategies: Perspectives in Post-Structuralist Criticism*, edited by Josue V. Harari. Ithaca: Cornell University Press, 1980.

Goldstein, Bill. “King of Horror.” *Publisher’s Weekly*, January 24, 1991.

Gramsci, Antonio. *Selections from Cultural Writings*, edited by David Forgacs and Geoffrey Nowell-Smith, translated by William Boelhower. London: Lawrence and Wishart, 1985.

Hackett, Alice P. “The ‘Best Seller’ Lists--and How they Grew.” *New York Times*, October 6, 1946: BR 22.

Hackett, Alice Payne and James Henry Burke. *80 Years of Best Sellers, 1895-1975.* New York: R.R. Bowker, 1977.

Hardin, Nancy. “For Paperback Houses, There’s No Business Like Show Tie-In Business.” *Publisher’s Weekly*, February 17, 1975.

Hardwick, Elizabeth. “The Decline of Book Reviewing.” *Harper’s Magazine*, October 1959.

Haynes, Christine. “Reassessing ‘Genius’ in Studies of Authorship: The State of the Discipline.” *Book History*, Vol. 8 (2005).

Heymann, Laura A. “The Birth of the Authornym: Authorship, Pseudonymity, and Trademark Law.” *Notre Dame Law Review*, Vol 80, Iss. 4 (2005): 1377-1450.

Jaszi, Peter. “On the Author Effect: Contemporary Copyright and Collective Creativity.” *Cardazo Arts & Entertainment Law Journal* 10 (1991-1992).

Jockers, Matthew and David Mimno. “Significant Themes in 19th-Century Literature.” *Poetics* 41: 6 (December 2013): 750-769.

Jockers, Matthew L. *Macroanalysis: Digital Methods and Literary History*. Topics in the Digital Humanities. Urbana: University of Illinois Press, 2013.

Kennedy, Maev. “Lawyer who uncovered JK Rowling’s Robert Galbraith Alter Ego fined £1,000.” *The Guardian*, December 31, 2013.

King, Stephen. “Frequently Asked Questions.” *stephenking.com*. Retrieved June 6, 2022.

Korda, Michael. *Making the List: a Cultural History of the American Bestseller, 1900-1999.* New York: Barnes and Noble, 2001.

Lottman, Eileen. “Novelizations: Are the Plums Drying Up?” *Publisher’s Weekly*, October 10, 1977.

Marx, Karl. *Capital: A Critique of Political Economy*, Volume 1, translated by Ben Fowkes. New York: Penguin Books, 1990.

—. *Capital, Volume 3*, edited by Friedrich Engels, translated by David Fernbach. London: Penguin Books, 1991.

—. *Capital, Volume 3*, edited by Friedrich Engels, translated by the Institute of Marxism-Leninism. International Publishers: New York, NY [n.d].

—. *Theories of Surplus Value*, Part 2. Progress Publishers: Moscow, 1968.

McGrath, Laura B. “More Specific, More Complex.” *Post45,* May 8, 2019.

Mehren, Elizabeth. “The Decade of the Mass-Market Hardcover.” *Los Angeles Times*, December 31, 1989.

Miller, Laura. *Reluctant Capitalists: Bookselling and the Culture of Consumption*. Chicago: University of Chicago Press, 2007.

Miller, Laura J. “The Best-Seller List as Marketing Tool and Historical Fiction.” *Book History* 3, Volume 1 (January 2000).

Mills, C. Wright. “The Cultural Apparatus.” In *Power, Politics, and People: The Collected Essays of C. Wright Mills,* edited by Irving Louis Horowitz. Oxford: Oxford University Press, 1963.

Montney, Charles B., ed. *Book Review Index: 1999 Cumulation.* Detroit: Gale Group, 2001.

—. *Book Review Index: 2000 Cumulation.* Detroit: Gale Group, 1999.

Moretti, Franco. “The Slaughterhouse of Literature.” *MLQ: Modern Language Quarterly* 61, no. 1 (2000): 207-209.

Patterson, Richard North. *No Safe Place.* New York: Knopf, 1998.

“Pearl Buck Scores Best-Seller Lists.” *The New York Times*, April 28, 1938.

Peckham, Stanton. “The Reviewer, the City and the Bookstore.” *Publishers Weekly*, March 20, 1961.

Perrin, Tom. “The Second-Greatest Stories Ever Told.” In *The Aesthetics of Middlebrow Fiction: Popular US Novels, Modernism, and Form, 1945-75.* New York: Palgrave MacMillan, 2015: 91-108.

Pierce, Leonard. “Harold Bloom on *Blood Meridian*.” *The A.V. Club*, June 15, 2009.

Piper, Andrew and Eva Portelance. “How Cultural Capital Works: Prizewinning Novels, Bestsellers, and the Time of Reading.” *Post45,* May 10, 2016.

Porter, J.D. “Popularity/Prestige: A New Canon.” *Pamphlets of the Stanford Literary Lab*, Pamphlet 17, September 2018.

Pruett, Jordan. ““NYT Hardcover Fiction Bestsellers.” *Post45 Data Collective*, 2022.

"*Publisher's Weekly* Paperback Bestsellers." *Publisher's Weekly*, January 8, 1988: 80.

Queen, Rob. "Interview with Alan Dead Foster." *SFF World*, November 22, 2007.

Radway, Janice. *A Feeling for Books: The Book-of-the-Month Club, Literary Taste, and Middle-Class Desire.* Chapel Hill: University of North Carolina Press, 2000.

—.*Reading the Romance: Women, Patriarchy, and Popular Literature*. Chapel Hill: University of North Carolina Press, 1984.

"Reviewers Debate Function of Reviewing." *Publishers Weekly,* March 18, 1968.

Rosen, Jeremy. Combining Close and Distant, or, the Utility of Genre Analysis: A Response to Matthew Wilkens's "Contemporary Fiction by the Numbers." *Post45*, December 3, 2011.

Rubin, Joan Shelley. *The Making of Middlebrow Culture.* Chapel Hill: University of North Carolina Press, 1992.

Sacks, Danielle. "Paper Lantern Lit Matches Undiscovered Writers With Publishers to Create Mass Market Hits." *Fast Company*, October 15, 2012.

Salzberg, Charles. "Book Packaging and the Big Time." *Publishers Weekly,* August 9, 1985.

"Schuster Assails 'Best Seller' Lists." *New York Times*, July 19, 1932: 15.

Shatzkin, Leonard. *In Cold Type: Overcoming the Book Crisis*. New York: Houghton Mifflin, 1982.

Sinykin, Daniel. *The Conglomerate Era.* New York: Columbia University Press, Forthcoming.

So, Richard. *Redlining Culture: A Data History of Racial Inequality and Postwar Fiction*. New York: Columbia University Press, 2020.

Sorensen, Alan T. "Bestseller Lists and Product Variety." *The Journal of Industrial Economics*, Volume LV, No. 4 (December 2007): 715-738.

Staiger, Janet. "*Kiss Me Deadly*: Cold War Threats from Spillane to Aldrich, New York to Los Angeles, and the Mafia to the H-Bomb." In *New Directions in American Reception Study*, Edited by Philip Goldstein and James L. Machor. Oxford: Oxford University Press, 2008.

St. Clair, William. *The Reading Nation in the Romantic Period.* Cambridge University Press, 2004.

Tanjeem, Namera. "10 Of The Biggest Book Deals In History." *Bookriot,* February 24, 2020.

Tarbert, Gary C. and Barbara Beach, eds. *Book Review Index: A Master Cumulation 1965-1984.*

Detroit: Gale Research Company, 1985.

Thompson, John. *Merchants of Culture: The Publishing Business in the Twenty-First Century*, 2nd edition. New York: Penguin Group, 2012.

Underwood, Ted. "The Life Cycles of Genres." *Journal of Cultural Analytics*, Vol. 2, Issue 2 (May 24, 2016).

—. "A Genealogy of Distant Reading." *Digital Humanities Quarterly* 11:2 (2017).

Underwood, Ted and Jordan Sellers. "The *Longue Durée* of Literary Prestige." *Modern Language Quarterly* 77:3 (September 2016).

Underwood, Ted, Patrick Kimutis, and Jessica Witte. "NovelTM Datasets for English-Language Fiction, 1700-2009." *Journal of Cultural Analytics*, Vol. 5, Issue 2 (May 28, 2020).

Vulliamy, Ed. "'The liberal elite hasn't got a clue.'" *The Guardian*, November 1, 2004.

Wagner, Geoffrey. "The Decline of Book Reviewing." *The American Scholar*, Vol 26, No. 1 (Winter, 1956-67).

Walker, Neil E. and Beverly Baer, eds. *Book Review Index: A Master Cumulation 1985-1992.* Detroit: Gale Research Inc., 1994.

Weedon, Alexis. "The Uses of Quantification," in *A Companion to the History of the Book, 2nd Edition,* edited by Simon Eliot and Jonathan Rose. Wiley-Blackwell, 2019.

"What Is This Thing Called Love?" *Publishers Weekly*, November 9, 1998: 43.

Whiteside, Thomas. "Onward and Upward with the Arts—The Blockbuster Complex, II." *New Yorker*, October 6, 1980.

Whiteside, Thomas. *The Blockbuster Complex: Conglomerates, Show Business, and Book Publishing.* Middletown: Wesleyan University Press, 1981.

Wilkens, Matt. "Contemporary Fiction by the Numbers," *Post45*, March 11, 2011.

Woodmansee, Martha. "The Genius and the Copyright: Economic and Legal Conditions of the Emergence of the 'Author.'" *Eighteenth-Century Studies* 17, no. 4, Special Issue: "The Printed Word in the Eighteenth Century" (Summer, 1984): 425-448.

"Write for Us." *Booklist Online*. https://www.booklistonline.com/writing-for-booklist

Writers Guild of America. *Understanding Separated Rights.* 2000.

Writers Guild of America West. "Separation of Rights." In *Television Credits Manual* [n.d.].