Nation, Ethnicity, and the Geography of British Fiction, 1880-1940

Elizabeth F. Evans and Matthew Wilkens

07.13.18

Among the most pressing problems in modernist literary studies are those related to Britain's engagement with the wider world under empire and to its own rapidly evolving urban spaces in the years before the Second World War.[1] In both cases, the literary-geographic imagination—or unconscious—of the period between 1880 and 1940 can help to shed light on how texts by British and British-aligned writers of the era understood these issues and how they evolved over

---

time. At the highest level, how can we characterize the international and domestic geographies of British writing? What roles, if any, did cultural identity play in contemporary writers' spatial imagination? What locations were over- or under-represented in their work and how, if at all, does the answer change when we group writers by national origin or by perceived ethnicity? What shifts in geographic attention marked the transition from the late Victorian period to the interwar era of high modernism? These questions, and others like them, have received much recent attention, both popular and academic.[2] In this essay, we explore what we learn when we ask them at scale with computational assistance.

Our goals in posing these questions are several. We seek first to assess the applicability of two specific, widely (though not universally) shared presumptions about the shape of British and British-aligned literature's engagement with the physical world during the period. These are its *internationalism*, by which we mean its interest in and use of locations outside the United Kingdom, and its *geographic intensity*, that is, the frequency of its reference to specific locations. Internationalism is attached to modernist literature in particular with such frequency that it can seem almost a truism. The early critical and polemical work of T.S. Eliot and Ezra Pound, classic studies by I.A. Richards and Hugh Kenner, and more recent scholarly turns to global modernisms and world literature are all premised on the decreasing significance in the early twentieth century of strictly national systems of literary production and on the central incorporation of a more cosmopolitan perspective into the era's literature.[3] How such a shift plays out in any given text or in the work of a single author is, of course, complex and unpredictable; certainly there were writers who remained steadfastly committed to their national frames. But we would be surprised to find that, taken as a whole, the literature of the early twentieth century was less international than that of the preceding decades.

The matter of geographic intensity is less widely debated, but no less interesting. Jon Hegglund has an explanatory mechanism in mind when he writes, in his excellent *World Views*, that

---

[2] An extended discussion of existing work on modernist-era literary geography follows below. Key studies include books such as Peter Brooker, *Bohemia in London: The Social Scene of Early Modernism* (Houndmills, Basingstoke: Palgrave Macmillan, 2004); Jon Hegglund, *World Views: Metageographies of Modernist Fiction;* Peter J. Kalliney, *Cities of Affluence and Anger: A Literary Geography of Modern Englishness* (Charlottesville: University of Virginia Press, 2006); Anna Snaith, *Modernist Voyages: Colonial Women Writers in London*, 1890-1945 (Cambridge: Cambridge University Press, 2014); Andrew Thacker, *Moving through Modernity;* Andrea Zemgulys, *Modernism and the Locations of Literary Heritage* (Cambridge: Cambridge UP, 2008).

[3] A contrary strain of *völkisch* nationalism was never absent from the period, most famously in Martin Heidegger's work, but clear in F.R. Leavis and, in the United States, in John Crowe Ransom and the Southern Agrarians.

> After the turn of the nineteenth century … many authors cease to in-
> clude maps in the front matter while at the same time maintaining,
> and often increasing, the amount of topographic detail in their nar-
> ratives. This overload of geographical particularity had the effect,
> ironically, of denaturalizing the 'background' spaces of fiction.[4]

But more attention to geographic space might also reasonably be expected to accompany the modernist era's increased internationalism and globalization, regardless of the specific fate of printed maps in its literature. On the other hand, a shift toward the representation of psychological interiority at the expense of the social world has often been associated with leading modernists from Virginia Woolf to James Joyce to William Faulkner. The resulting tension between outward-looking geographic intensity and inward-facing psychology opens space for a new quantitative intervention.

By characterizing the geographic attention of a large swath of British fiction published between 1880 and 1940, we hope not only to address these questions of internationalization and intensity, but also to detect other widespread spatial phenomena in the period and to better understand the dynamics of selected subgroups of authors and of texts in relation to one another. We ask, for instance, to what extent the London that took shape in more or less canonical writing of the period was representative of the imagined London of British fiction as a whole and, hence, to what extent canonical fiction is a reasonable proxy for period writing generally. How did foreign writers, especially those who identified as Black or Asian, resemble and diverge from native British authors in their treatment of the metropole, the nation, and the globe? Was there a distinctive form of regional fiction centered on London and, if so, how did it differ from other writing at the time? How, moreover, did any of these forms and groups change over the course of the sixty years leading up to the Second World War, or during what we might call the long modernist era?

It should be clear, then, that while we have a range of specific questions to answer, our work is also in part exploratory and recuperative. Our research concerning the imagined geography of foreign writers in Britain addresses a dearth of literary scholarship regarding writing by people of color within the nation prior to the more familiar influx of migrants from British colonies after the Second World War and, we hope, contributes to the ongoing recovery of this largely forgotten body of creative work. Finally, we aim to provide both quantitative and qualitative context for future research on the literature and culture of the period via computational means that are novel in the area.

---

[4]Jon Hegglund, *World Views: Metageographies of Modernist Fiction* (Oxford: Oxford UP, 2012, 9-10.

The sections below proceed by way of much new data, almost all of which is tied directly to the questions posed here. Our results lead us to three broad interventions in modernist literary studies. First, we argue that a modernist studies that values internationalism must devote significantly more attention to non-canonical literature. The mass run of fiction published between 1880 and 1940 was consistently and meaningfully more international than its better-known analogues. Writing by non-native British writers was radically more so. If critics are drawn to the outward turn in modernist texts, they can and should find a larger, earlier, and perhaps more important version of the phenomenon by looking beyond the usual suspects.

Second, we need to rethink London as it was encountered and described by outsiders. This isn't just a matter of turning away from the famous and the posh in favor of the neglected and the downtrodden (though there are worse places to start). It's about explaining, for instance, why foreign writers of color depict a more public, verdant London than their colony-born white counterparts, while devoting less of their attention to the East End and to notably international districts of the city. These patterns are either anecdotal or essentially invisible to conventional study. Computational methods make them available for nuanced literary-historical reinterpretation.

Finally, we argue against treating the years between 1880 and 1940 in terms that emphasize temporal discontinuity. Aspects of British fiction did change across this span of sixty years, and many of the differences we observe in the era's literary-geographic attention are genuinely important. But when we work at scale, it's very difficult to locate "on or about …" moments of sudden change across whole ranges of texts. We see instead situations of influence and drift or—and this is the rub—we find true ruptures only between corpora built around differing principles. The latter case, comparing corpora assembled to emphasize difference, is the one that resembles most closely the way in which modernist studies built its canons. Those canons and the practices they embed aren't simply errors, but they are deliberately and systematically nonrepresentative of large-scale literary history. Modernist literary critics would do well to grapple with that fact more directly than we often have.

There is long-standing interest in the literary geography of London. The lives and works of individual writers from Arthur Conan Doyle to Virginia Woolf, and the sites associated with particular literary and social networks, have been plumbed for named locations and marked on maps.[5] Woolf herself wrote of the

---

[5] See, for example, treatments of the London of Charles Dickens and of Arthur Conan Doyle in

fashion for literary guides in her day.[6] Though it's often said that London can't be synthesized (Julian Wolfreys, for example, asserts that London "resists ontology, and thus affirms its alterity, its multiplicities, its excesses, its heterogeneities"),[7] the number of popular and academic studies of London and its literature continue to grow.[8]

While there's little consensus concerning *which* locations matter most in London writing, it's now generally accepted that the overall geography of fiction matters very much indeed.[9] With the "spatial turn" of recent decades, narrative and cultural critics have found close relationships between the novel and its geography. Susan Stanford Friedman and Franco Moretti, drawing in part and in different ways on the work of Mikhail Bakhtin, have each influentially argued that setting gives rise to distinct varieties of narrative and that some narratives are inconceivable outside particular settings.[10] Richard Dennis voices a pervasive belief in current scholarship that, even if readers may not always grasp the significance of a novel's named locations, local geography matters because "locating characters and events geographically helped to locate them socially and symbolically."[11] Individuals shape and are shaped by geography, an insight that is often extended to explore the close relationship between conceptions of the metropolis and con-

---

two otherwise disparate sources, Franco Moretti's *Atlas of the European Novel 1800-1900* (London: Verso, 1998) and the online magazine the *Londonist*, "The London of Charles Dickens: Mapped" and "The London of Sherlock Holmes…Mapped." See Lynne Walker, "Vistas of Pleasure: Women Consumers of Urban Space in the West End of London 1850-1900," *Women in the Victorian Art World*, ed. Clarissa Campbell Orr (Manchester: Manchester UP, 1995) for an analysis of the interdependence between the West End of London and the late nineteenth-century women's movement.

[6]In her first review for the *Times Literary Supplement*, in 1905, Woolf critiqued two examples of the genre, *The Thackeray Country* and *The Dickens Country*. She faulted both for, among other things, having misleading titles that underplay the centrality of urban rather than rural environments in Thackeray and Dickens's writing. See Virginia Woolf, "Literary Geography," [1905] *The Essays of Virginia Woolf.* Vol 1. Ed. Andrew McNeillie (London: Hogarth Press, 1986).

[7]Julian Wolfreys, *Writing London.* Volume 2: Materiality, Memory, Spectrality (Houndmill, England: Palgrave Macmillan, 2004).

[8]Generalist cultural histories of London include books such as Peter Ackroyd, *London: The Biography* (New York: Doubleday, 2000), David Kynaston, *City of London.* 4 Vols. (London: Chatto & Windus, 1994-2001), and Roy Porter, *London: A Social History* (Cambridge, MA: Harvard UP, 1995).

[9]For example, Peter Brooker and Andrew Thacker's edited collection *Geographies of Modernism* begins with the observation that "we have recently witnessed a strong interest in the spatial and geographical dimensions of literary and cultural texts." Peter Brooker and Andrew Thacker, eds. *Geographies of Modernism: Literatures, Cultures, Spaces* (London: Routledge, 2005), 1.

[10]See especially Susan Stanford Friedman, *Mappings: Feminism and the Cultural Geographies of Encounter* (Princeton and Oxford: Princeton, 1998), Franco Moretti's *Atlas of the European Novel* and *Graphs, Maps, Trees* (London: Verso, 2005), and Mikhail Bakhtin, *The Dialogic Imagination: Four Essays*, ed. Michael Holquist, trans. Caryl Emerson and Michael Holquist (Austin: University of Texas Press, 1982).

[11]Richard Dennis, *Cities in Modernity: Representations and Productions of Metropolitan Space, 1840-1930* (Cambridge: Cambridge UP, 2008), 105.

ceptions of identity.[12] Named locations can also tell national stories; places have been "apt metaphors" for defining Englishness.[13]

The ways in which literary-geographic imaginations changed over time matter, too. "[T]he spatial history of modernism," argues Andrew Thacker, necessitates "an account of the precise historical fashion in which particular spaces and places were conceptualized and represented."[14] The sixty years spanning the last decades of the nineteenth century and the first four of the twentieth are generally accepted to entail significant changes in literature's subject matter and methods. The "critical consensus," as Eric Bulson writes, has it that "ways of representing the city and the world changed radically between Dickens' London and Joyce's Dublin."[15]   Hegglund's claim that, during this period, literature increased its topographic detail and particularity is attractive both because it affirms the importance of literary geography and because it is amenable to computational analysis.[16]   Below, we assess shifts in the intensity of geographical references over time at scale and compare historical changes in fiction's degree of attention to sites beyond national borders to determine if, as one might expect, the rise of modernist cosmopolitanism and the experience of a world war conducted mostly overseas produced more internationally-focused literature.

London's status as the political and conceptual center of a vast empire has likewise received increasing interest in recent years. Jonathan Schneer has argued that London in 1900 was an "imperial city" not only by virtue of its place in global trade and politics but also in its built environment. From Nelson's column in Trafalgar Square and Cleopatra's Needle on the Thames Embankment to the "revived classicism" of buildings constructed around the turn of the century, the "public art and architecture of London together reflected and reinforced an

---

[12] On cities and identity, see, for example, edited collections such as Dana Arnold, ed., *The Metropolis and its Image: Constructing Identities for London*, c. 1750-1950 (Oxford: Blackwell, 1999) and Felix Driver and David Gilbert, eds., *Imperial Cities: Landscape, Display and Identity* (Manchester: Manchester UP, 1999).

[13] Ian Baucom, *Out of Place: Englishness, Empire and the Locations of Identity* (Princeton: Princeton University Press, 1999), 4.

[14] Andrew Thacker, *Moving through Modernity: Space and Geography in Modernism* (Manchester: Manchester UP, 2003), 5.

[15] Eric Bulson, *Novels, Maps, Modernity: The Spatial Imagination, 1850-2000* (New York: Routledge, 2007), 2.

[16] Other critics have argued for changes in spatial representation but not always with such concrete and measurable claims. Nico Israel, for instance, finds that "locations beyond England and America appear frequently as the subjects of [modernist] texts, and in previously unimaginable ways" and observes an "immense conceptual distance between Stoker's or Kipling's sense of place and space and Joyce's or Beckett's—just as there is between Stoker's and Kipling's and Dickens's and Eliot's. Nico Israel, "Geography," in *A Companion to Modernist Literature and Culture,* eds. Bradshaw and Dettmar, 2006, 127, 131.

impression, an atmosphere, celebrating […] British imperialism."[17] Imperialism also shaped the city's population, as British civil servants traveled to occupied lands, industry and trade brought workers from around the world to London's Docks, and colonial subjects came to study at its universities. Imperial London was also the center of anti-imperial networks and activities. Elleke Boehmer, among others, has called attention to how the city was "an important meeting ground for Indian, Irish, African, and Caribbean freedom movements."[18] Pan-Africanism, "one of the major political traditions of the twentieth century, was largely created by black people living in Britain"[19] and London was "*the* focus for anti-imperialist agitation."[20] Historians have made important inroads in recovering the experiences of African, West Indian, African-American, East Asian, and South Asian visitors and residents in London prior to 1948, when the first of a "wave" of migrants from the West Indies arrived in Britain on the SS *Empire Windrush*, an event that still looms large in cultural narratives about the nation's ever increasing "foreign" population.[21]

---

[17] Jonathan Schneer, *London 1900: The Imperial Metropolis* (New Haven & London: Yale UP, 1999), 19. Felix Driver and David Gilbert have stressed how imperialism was also imbued in such apparently innocent spaces as gardens with "oriental" landscaping and exotic species and in the naming of streets and houses after imperial campaigns ("Heart"), while Ian Black has argued that the imperial classical design of City of London banks built in the 1920s and 30s aimed to bolster fading confidence in the City's status as the "heart of empire." Felix Driver and David Gilbert, "Heart of Empire? Landscape, Space and Performance in Imperial London," in *Environment and Planning D: Society and Space* 16 (1998): 11-28; Ian S. Black, "Rebuilding 'The Heart of the Empire': Bank Headquarters in the City of London, 1919-1939," in *The Metropolis and its Image: Constructing Identities for London, c. 1750-1950*, ed. Dana Arnold (Oxford: Blackwell, 1999), 127-52.

[18] Elleke Boehmer, *Empire, the National, and the Postcolonial, 1890-1920: Resistance in Interaction* (Oxford: Oxford UP, 2002), 20.

[19] Peter Fryer, *Staying Power: The History of Black People in Britain* (London: Pluto Press, 1984), 272.

[20] Roderick J. Macdonald, " 'The Wisers Who Are Far Away': The Role of London's Black Press in the 1930s and 1940s," in *Essays on the History of Blacks in Britain: From Roman Times to the Mid-Twentieth Century*, eds. Jagdish S. Gundara and Ian Duffield (Avebury: Aldershot, England, 1992), 166.

[21] This despite efforts to educate citizens that diversity was part of Britain from its beginnings, such as the Museum of London's project " 'The Peopling of London,' the aim of which was to call attention to 15,000 years of settlement" and websites like *Making Britain*. (John McLeod, *Postcolonial London: Rewriting the Metropolis* (London: Routledge, 2004), 6.) Significant historical studies include monographs by Elleke Boehmer, *Empire, the National, and the Postcolonial*; Barbara Bush, I*mperialism, Race and Resistance: Africa and Britain 1919-1945* (London: Routledge, 1999); Peter Fryer, *Staying Power: The History of Black People in Britain* (London: Pluto Press, 1984); Jeffrey Green, *Black Edwardians: Black People in Britain 1901-1914* (London: Frank Cass, 1998); C.L. Innes, *A History of Black and Asian Writing in Britain, 1700-2000* (Cambridge: Cambridge UP, 2002); Marc Matera, *Black London: The Imperial Metropolis and Decolonization in the Twentieth Century* (Oakland: The University of California Press, 2015); Sukhdev Sandhu, *London Calling: How Black and Asian Writers Imagined a City* (London: Harper Perennial, 2003); Jonathan Schneer, *London 1900: The Imperial Metropolis* (New Haven & London: Yale UP, 1999); Rozina Visram, *Ayahs, Lascars and Princes: The Story of*

But, while literary scholars have increasingly turned to the roles of race, ethnicity, national origin, and imperialism in London literature of the late nineteenth and early twentieth centuries, there remains a dearth of direct attention to period writing by people of color. The imperial city of "colonial writers" is often explored through white writers like Jean Rhys and Katherine Mansfield, who occupied an ambivalent place in imperial Britain's racial hierarchy but whose experiences and perspectives certainly were not representative of many colonial subjects in London.[22] "Black British" literature is often taken to begin with the postwar writing of Sam Selvon and George Lamming.[23] Perhaps still limited by a longstanding preference for stylistically experimental writing associated with high modernism, modernist studies' much discussed "expansion" hasn't yet resulted in a comprehensively reshaped canon, though one hopes that a handful of excellent and well-received recent books will continue to have an impact.[24] In the most recent edition of the *Oxford Companion to English Literature*, Bénédicte Ledent urges "it should not be forgotten that there had been a sizeable body of texts predating the work of pioneer figures like Samuel Selvon or George Lamming,"[25] but he finds space to name only two such writers, C.L.R. James and Una Marson, active in

---

Indians in Britain 1700-1947 (London: Pluto Press, 1986).

[22] In books that focus on the intersection of imperialism and London writing, writers of color prior to the 1950s are often silently elided, or given only passing mention. Jed Esty's important book on the postwar period, for instance, acknowledges that "writers from the colonies and ex-colonies had been a formative part of the London literary scene for decades" but names only white writers by way of examples. Jed Esty, *A Shrinking Island: Modernism and National Culture in England* (Princeton and Oxford: Princeton UP, 2004), 200.

[23] See Fred D'Aguiar, "Against Black British Literature," in *Tibisiri: Caribbean Writers and Critics*, ed. Maggie Butcher (Coventry: Dangaroo, 1989), 106-14; Bénédicte Ledent, "Black British Literature," *Oxford Companion to English Literature*. 7th ed. Ed. Dinah Birch (Oxford: Oxford UP, 2009), 17-22; and John McLeod, "Some Problems with 'British' in a 'Black British' Canon," *Wasafiri* 17.36 (2002): 56-59, for discussion of the history of the category "Black British" and its political benefits and limitations.

[24] "Expansion" is identified as "a single word to sum up transformations in modernist literary scholarship over the past decade or two" in Doug Mao and Rebecca Walkowitz, "The New Modernist Studies," *PMLA* 123.3 (2008): 737-48, and the trend has continued unabated in the decade since. Groundbreaking scholarship such as Jessica Berman, *Modernist Commitments: Ethics, Politics, and Transnational Modernism* (New York: Columbia UP, 2011), Sara Blair, "Local Modernity, Global Modernism: Bloomsbury and the Places of the Literary." *ELH* (Fall 2004): 813-838, Mary Lou Emery, *Modernism, the Visual, and Caribbean Literature* (Cambridge: Cambridge UP, 2007), and Anna Snaith, *Modernist Voyages: Colonial Women Writers in London, 1890-1945* (Cambridge: Cambridge University Press, 2014), has begun to illuminate how colonial writers of color responded to and shaped cultural forms throughout the coinciding periods of high imperialism and high modernism in varied and vital ways. Outside the British context, Richard Jean So and Hoyt Long have made important contributions to comparative modernist studies using computational methods, including social network analysis to map publishing cultures in the U.S., China, and Japan and machine learning to trace the evolution of the English-language haiku as a modernist form. Richard Jean So and Hoyt Long, "Network Analysis and the Sociology of Modernism," *Boundary* 2. 40.2 (2013): 147-82; "Literary Pattern Recognition: Modernism between Close Reading and Machine Learning," *Critical Inquiry* 42.2 (2016): 235-67.

[25] Ledent, "Black British Literature," 17

the half century preceding them, each of whom was writing of London in the 1930s (only the former gets an entry in the volume).[26] Though "black British writing" is an area of literature that has "received generously enhanced coverage" in the latest *Oxford Companion*, most of the attention goes toward more recent writing, a pattern evident in literary scholarship as well.[27] We do not minimize the importance of research on recent and contemporary writing by writers of diverse backgrounds (both of us write about and teach this material), but we hope to contribute to efforts to increase awareness of the varied writings by people of color in the era of British modernism, broadly conceived.

Did sites where large numbers of immigrants lived, studied, and worked receive greater proportional attention in writing by authors with similar geographic and demographic origins? According to the *London Encyclopaedia*, in "1911, one in twenty-five of London's residents was foreign born," but they were not evenly dispersed across the city.[28] In the nineteenth century, Chinese, Indian, and Caribbean seamen and dockworkers settled in Canning Town, Stepney, and Poplar in the East End of London. Soho was an important center for other immigrant populations and, by the turn of the century, Bloomsbury housed students from around the world. Were there differences in the proportional mention of these and other sites across different types of literature? These questions raise larger concerns about the relationship between identity and place, and about the relationship of historical circumstance and literary representation, which we explore below.

Corpora and methods

To address these literary-geographic questions, we assembled four corpora, each comprising books published in Britain between 1880 and 1940. These groups range from the comprehensive to the specialized, varying in size from over 10,000 books to as few as 130. In all, our research collections include 10,765 distinct volumes.

The largest of the corpora, labeled "Hathi" in the figures and discussion below and serving as a type of baseline for the others, contains 10,010 volumes of fic-

---

[26] Ledent also names the much earlier Olaudah Equiano (c. 1745-1797) and Mary Seacole (1805-1881).

[27] Dinah Birch, ed., *Oxford Companion to English Literature*. 7th ed. Oxford: Oxford UP, 2009, ix.

[28] Ben Weinreb, Christopher Hibbert, Julia Keay, and John Keay, *The London Encyclopaedia*. 3rd Ed. (London: Macmillan, 2008), 427.

tion published in the UK between 1880 and 1940 and held by the HathiTrust digital library. These volumes are the ones previously identified by Underwood et al., deduplicated and restricted to items having bibliographic metadata indicating publication in Great Britain.[29] The full text of each volume was included, except for basic preprocessing to remove paratext such as running heads and page numbers. The texts in all corpora were identically prepared; they were supplied by HathiTrust and processed non-expressively. Histograms of the four corpora by date of publication are presented in figure 1.

[29] Specifically, those with MARC publication location "enk" or an imprint entry beginning with "London." We accepted Underwood et al.'s selection of volumes at least 50% likely to be 80% precise in their classification as fiction. For details, see Ted Underwood, "Understanding Genre in a Collection of a Million Volumes," Interim performance report, NEH Digital Humanities Start-Up Grant. 29 December, 2014, as well as the code supplement to the present article. Researchers working with HathiTrust sources may be interested in the number of volumes culled at each stage of corpus assembly. Underwood et al. identify 101,948 volumes of English-language fiction in the HathiTrust public domain holdings and 104,043 such volumes in the in-copyright collection. Of these 205,991 volumes, 51,393 (25%) were published in the UK, of which 20,170 (9.8% of the total, 39% of British volumes) have publication dates between 1880 and 1940. See Ted Underwood, Boris Capitanu, Peter Organisciak, Sayan Bhattacharyya, Loretta Auvil, Colleen Fallaw, J. Stephen Downie. "Word Frequencies in English-Language Literature, 1700-1922 (0.2)" [Dataset]. HathiTrust Research Center, 2015. HathiTrust contains a significant number of duplicate volumes; discarding second (and subsequent) items that share the same author and title, we are left with 11,414 distinct volumes. We lack geographic information for a small percentage of these and have identified a handful of others as non-fiction, bringing us to our reported working corpus size of 10,010 volumes. Our other bibliographies lack geographic data for no more than a trivial number of identified volumes, though they face other limitations related to Hathi's holdings as described below.
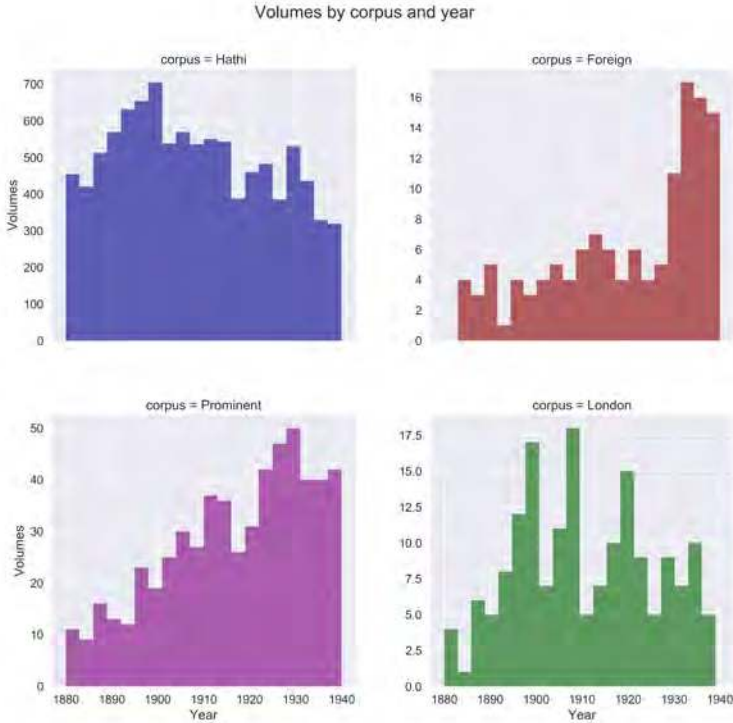
Figure 1. Volumes by corpus and year of publication in the research dataset. Note that the *y*-axes are not shared between corpora; the Hathi corpus is much larger than the others, which differ in turn from one another.

The other three corpora are more restricted, though they cover the same publication dates and broad context of publication. "Prominent British Fiction" (labeled "Prominent" below), comprises 576 fiction volumes drawn from three sources that describe and reinforce the canon of British novels. We used the chronological list of principal literary works in the widely circulating *Oxford Companion to English Literature*, now in its seventh edition (2009, ed. Dinah Birch). We excluded poetry, drama, and nonfiction works listed in the *Companion*; we included the small number of listed fictional works by non-British writers (mostly American and Irish). Second, we used Thomas Jackson Rice's *Bibliography of English Fiction*, 1900-1950 (limited by our dates).[30] As an additional measure to

---

[30] Rice employed three criteria to select authors for inclusion: "(1) all generally acknowledged 'ma-

insure that the Prominent corpus includes all fiction that rises to that name, and for added continuity across the sixty year period, we referred to the Norton *Anthology of English Literature* (both 8th edition [2006] and 9th edition [2012], the latter being the most recent available at the time of writing). Using the Norton as a proxy for authorial prominence, we verified that our corpus included every work of fiction by Norton authors that is listed in Oxford's *The Reinvention of the British and Irish Novel, 1880-1940*.[31] Through this process, for example, novels by Thomas Hardy and Jean Rhys are captured, though only poetry and short stories, respectively, appear in the Norton.

As the name suggests, the Prominent corpus comprises generally canonical or near-canonical fiction by writers published in Britain between 1880 and 1940. The prominence of the authors involved produces a corpus that is closer than any other in our dataset to the contours of canonical modernism. But we have been careful to avoid referring to it as such and, indeed, it contains important realist, popular, and topical fiction alongside aesthetically experimental texts by Woolf, E.M. Forster, D.H. Lawrence, and so on. When we examine these texts by author gender, race, and national origin, it reinforces what is generally known: the traditional canon of British fiction skews white and male. But the degree of gender and racial disparity might surprise. Of 576 novels, just 93 were by women, roughly 16%.[32] All 576 novels—100%—were written by authors classified as white. Nearly 90% of authors were born in Britain, the vast majority of whom

---

jor' novelists—Conrad, Joyce, Lawrence, and Woolf … and those one might call the 'second echelon' of major-minor novelists—Bennett, Ford, Forster, Galsworthy, Huxley, Maugham, Orwell, Waugh, and Wells; (2) all major men of letters who, though they may be better known for their achievements in other fields, have made a significant contribution to modern fiction—Aldington, Beerbohm, Chesterton, C.S. Lewis, and Wyndham Lewis; and (3) all so-called minor writers who have, nonetheless, attracted a significant amount of bibliographical, biographical, or critical commentary, and who have contributed significantly to the development of modern long and short fiction in Britain—Bowen, Compton-Burnett, Douglas, Firbank, Hartley, Lehmann, Macaulay, Mansfield, Munro, Myers, J.C. and T.F. Powys, Richardson, Rolfe, West, and Williams." Thomas Jackson Rice, "Introduction: The Scope and Format of this Guide," *English Fiction, 1900-1950: General Bibliography and Individual Authors: A Guide to Information Sources*. 2 Vol. (Detroit: Gale Research Co., c. 1979).

[31] We did not use *The Reinvention of the British and Irish Novel, 1880-1940* as a basis for our list because its (deliberate and admirable) inclusion of little-known works would have misrepresented what is commonly read, taught, and studied, which was the aim of this corpus. Patrick Parrinder and Andrzej Gasiorek, *The Reinvention of the British and Irish Novel*, 1880-1940. Vol. 4 of *The Oxford History of the Novel in English* (Oxford: Oxford UP, 2011), 609-33.

[32] For a survey of the long history of women writers' absence from anthologies (and, hence, from the domain of study), see Laura C. Mandell, "Gendering Digital Literary History: What Counts for Digital Humanities," in *A New Companion to Digital Humanities*, 2nd ed. Eds. Susan Schreibman, Ray Siemens, and John Unsworth (Hoboken, NJ: Wiley-Blackwell, 2016), 511-23. Regarding the recurring neglect of women writers in modernist studies, see Anne Fernald, "Women's Fiction, New Modernist Studies, and Feminism," *Modern Fiction Studies* 59.2 (2013): 229-40, and Jane Garrity, "Found and Lost: The Politics of Modernist Recovery," *Modernism/modernity* 15.4 (2008): 803-12.

(almost 84% of the total) were born in England. The next largest group may be described as hyphenated anglos (Anglo-Indian, Anglo-Irish, Anglo-Caribbean, Anglo-American), people of "Anglo-Saxon" ancestry who were born and perhaps raised in a colonized territory but who spent significant portions of their lives in Britain (Rudyard Kipling and George Orwell, for example).[33] So, "Prominent British fiction," as constituted by the existing critical literature, is overwhelmingly white, English, and male.[34]

The third corpus, London fiction, comprises the 171 available, relevant books named in three sources devoted to London as a literary site. The longest of these is the "London" portion of K.D.M. Snell's *Bibliography of Regional Fiction in Britain and Ireland, 1800-2000* (restricted to 1880-1940 for our purposes).[35] To insure that this corpus reflects recent scholarship on the literature of London, we included (place- and date-restricted) works named in the bibliography of fiction provided by Richard Dennis in *Cities in Modernity*, and relevant fiction discussed in peer-reviewed articles of the *Literary London Journal* since its inaugural issue in 2003. While the London corpus contains a number of highly canonical novels—*The Secret Agent, Mrs. Dalloway*—it is made up, on the whole, of much more obscure and "popular" texts than those on the Prominent list. This means that the corpus contains more genre fiction than do the others, especially detective stories, crime and sensation novels, and sociologically inflected accounts of extreme poverty and wealth. These 171 texts held by HathiTrust represent a disappointingly small percentage (48%) of the 359 entries in our consolidated London bibliography. This fact implies that our large Hathi corpus isn't as complete as researchers in the field might hope, particularly regarding less prominent works unlikely to be held by the mostly American universities that generate HathiTrust's archive. (By comparison, our Prominent source bibliography numbered 628, of which 576 [92%] are in the corpus.)

---

[33] Authors were identified as white or nonwhite on the basis of historical-biographical research. We required an unambiguous and uncontested attribution of ethnic identity in a published source. Where such an attribution was unavailable, ethnicity was not included in our dataset.

[34] For reasons of scope, gender is not a primary focus of the current article. We expect to examine the relationships between gender and literary geography in future work related to Evans's forthcoming *Threshold Modernism: New Public Women and the Literary Spaces of Imperial London* (Cambridge: Cambridge University Press, forthcoming 2018).

[35] Snell defines "regional fiction" as "fiction that is wholly or largely set in a particular geographical region, and which purports to describe or use recognisable and distinctive features of the life, customs, language, dialect, or other aspects of that area's culture and people." K. D. M. Snell, *The Bibliography of Regional Fiction in Britain and Ireland, 1800-2000* (Hants, England: Ashgate, 2002), 2. He also includes "fiction that conveys a strong sense of local geography, topography or landscape," as well as, beyond novels, "some items of a semi-autobiographical/fictional character" (2, 9). Snell notes that his definition of regional fiction "shares much with those adopted by other authors," including *The Concise Oxford Companion to English Literature* (2).

Finally, we assembled the fourth corpus, "Foreign authors in Britain" ("Foreign")—which is in some ways the primary object of our investigation—through reference to seven distinct critical studies. It includes 130 volumes in sum, drawn from: the bibliographic "Notes on Writers" in C.L. Innes's *A History of Black and Asian Writing in Britain, 1700-2000* (limited to the dates of our study); relevant books discussed by Antoinette Burton, Barbara Bush, Anna Snaith; fiction identified in David Dabydeen, John Gilmore, and Cecily Jones's *The Oxford Companion to Black British History*; and books identified in the archival work of David Killingray, emeritus professor of history at the University of London (Goldsmiths).[36] Again, the HathiTrust holdings of texts in this bibliography are less complete than those of the Prominent corpus, reflecting the deficits of contributing libraries, which were historically less likely both to acquire and to preserve these texts, as well as the texts' sometimes more obscure circumstances of publication; indeed, our Foreign bibliography lists 259 relevant volumes published between 1880-1940, almost exactly twice as many as the 130 available for inclusion in the corpus.

To be included in the Foreign corpus, a book must, in addition to having been named in the critical sources above, have been produced by a writer born and raised overseas and outside Europe who was resident in the UK for some period as an adult, generally as an outsider of one sort or another. The majority of these authors were from Britain's colonial possessions, especially in the Caribbean, South Asia, and Africa, and were of ethnicities other than white Anglo. Nevertheless, the Foreign corpus includes some books by writers who were generally identified as white. Most of these authors, like Eliot Bliss, William Plomer, and Jean Rhys, were born and raised in colonized territories (respectively, Jamaica, South Africa, and Dominica) as part of a minority population of colonial occupiers. Again, because we have aimed to focus on the works and cultural contexts of a subset of authors who have often been left out of literary study, we have not included white writers from white settler colonies (Australia, New Zealand, Canada, or the United States), nor have we included white authors from colonized territories who were educated in Britain and were more "insiders" than "outsiders" in the ruling society (such as Rudyard Kipling and George Orwell).[37]

---

[36] Antoinette Burton, *At the Heart of the Empire: Indians and the Colonial Encounter in Late-Victorian Britain* (Berkeley: University of California Press, 1998) and "Making a Spectacle of Empire: Indian Travellers in Fin-de-Siècle London," *History Workshop Journal* 42 (1996): 127-46; Barbara Bush, *Imperialism, Race and Resistance: Africa and Britain 1919-1945;* Anna Snaith*, Modernist Voyages: Colonial Women Writers in London, 1890-1945,* David Dabydeen, John Gilmore, and Cecily Jones,*The Oxford Companion to Black British History* (Oxford: Oxford UP, 2007.

[37] We might have reasonably chosen a different approach, as Caryl Phillips did in *Extravagant Strangers* when he compiled an "anthology of writing by British writers who are outsiders in the most clear-cut way—those not born in Britain" in order to show that "English literature has, for at

Including both white and nonwhite authors allows for comparisons between people who emerged from similar geographical areas but dissimilar social environments, broadly differentiated by perceived race; included in the results and analysis below are comparisons of the observed differences between nonwhite- and white-authored texts from this corpus.[38] Of course, such racial labels are troubled and complex, social rather than biological categories, based on constructions of difference that were enlisted to justify exploitation and abuse. But they had—and they continue to have—important, tangible political and social meanings.[39]

Foreign writers, especially those visibly identified as nonwhite, faced a context of publication different from their white British peers. Among the effects of this difference was an altered balance between novelistic fiction and other prose forms. Put simply, foreign writers often chose—or were forced—to produce boundary-crossing works of travelogue, memoir, narrative history, and expository essays. To exclude these forms would be to exclude a large portion of literary production by foreigners white and nonwhite alike during the period under consideration. It would also produce deeply misleading results, since the novel represented a uniquely minority form for foreign authors. We have therefore used, following our critical sources, an expansive understanding of narrative in the Foreign corpus, controlling as appropriate for the resulting generic diversity.

---

least 200 years, been shaped and influenced by outsiders." Caryl Phillips, ed. and intro. *Extravagant Strangers: A Literature of Belonging* (New York: Vintage, 1997), xiii. The anthology includes not only Kipling and Orwell but also Anglo-American T.S. Eliot, as well as black and Asian writers (but only one minority writer with a publication date between 1880 and 1940).

[38]Future research might encompass more white colonial authors, both from white settler colonies (Australia, New Zealand, Canada) and from British-colonized regions with a predominantly nonwhite population (India, etc.). Another potential area of inquiry would focus on Irish writers, who generally spent much of their formative years as one of the colonized majority, and "Anglo-Irish" writers, who often "returned" to the England of their ancestors and enjoyed more access to the resources of the capital.

[39]The Irish case highlights the difficulty of mediating between the racial discourses of the 1880s-1930s and those of the present day. In the late nineteenth century, the Irish were widely discussed in British writing as a race apart, one physiognomically distinct. But, while Irish anti-imperial activists met and traded information and strategies with anti-imperial activists from India and, to a lesser degree, from Africa and the Caribbean, their experiences of racial visibility differed significantly. Duse Mohamed Ali, a black Egyptian writer, publisher, and anti-imperial activist referenced this difference when, in 1920, he wrote in his London-based journal, *Africa and Orient Review*, "it behoves the coloured people of the world to show a solid front," since "[a]ll non-Europeans are labelled 'niggers' by Europeans." Duce Mohamed Ali, "The Final Word." *Africa and Orient Review* (May 1920): 45-46. The experiences of—and representations by—the Irish seem to call for separate treatment. We also have not identified Jewish writers to include among the "foreign"—though Jews (like the Irish) formed a large minority of London's population and encountered virulent discrimination—because their geographical and imperial history is significantly different from those of the colonial writers whose work is our focus here. These are obvious areas of future investigation.

Employing methods previously described[40] we extracted named locations from all corpus texts using the Stanford named entity recognizer (NER)[41] and associated each location string with detailed present-day geographic data via Google's Places and Geocoding APIs.[42] We then normalized location counts by volume length, reporting comparative measures on a per-100,000-words basis.[43] Potential sources of error between the conceptual formulation of the target literary formations and the finally extracted geographic data are several. We know that our bibliographies are subject to interpretation and to the vagaries of scholars' idiosyncratic selections (as well as our own). Many texts identified in the sources are unavailable via HathiTrust or, in a small number of cases, may be misidentified by our automated process of matching bibliographic records. HathiTrust texts were digitized via scanning and optical character recognition, and contain numerous mistranscriptions. The NER process is imperfect, and subsequent geocoding of even properly recognized locations can fail due to toponymic ambiguity. Still, our results are encouraging given the difficulty of the task. We find that we are able to identify correctly individual places with slightly less than 80% accuracy on average and that we identify the correct nation-level focus of individual volumes with better than 96% accuracy (and with region-level accuracy above 92%). Still, these limits suggest why we have chosen to emphasize comparative and large-scale analyses; our data are in most cases not sufficiently accurate to support high confidence in the geographic details of any one text, but they do much better when aggregated over many texts, especially when our object is to compare differing usage rates of prominent locations.

---

[40] Matthew Wilkens, "The Geographic Imagination of Civil War Era American Fiction," *American Literary History*. 25.4 (2013): 803-840.

[41] Jenny Rose Finkel, Trond Grenager, and Christopher Manning, "Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling." In *Proceedings of the 43nd Annual Meeting of the Association for Computational Linguistics* (*ACL 2005*). New Brunswick: ACL, 2005. 363-370.

[42] Acquiring, standardizing, and harmonizing global historical geographic data at levels from entire nations to individual buildings is a task beyond the scope of our project. The most notable consequence of using present-day geographic definitions is the separation of once-colonized areas from the nations that ruled them between 1880 and 1940. This includes the Republic of Ireland, which is not counted in our data as a British domestic location. Historical toponyms ("Constantinople") are, in general, correctly resolved to their closest modern equivalents ("Istanbul").

[43] The average volume length across our corpora is 101,598 words.

Results and analysis

We can begin to summarize our findings by moving from the largest to the smallest geographic scales. Figure 2 shows the distribution of global textual attention in the four corpora, summarized by nation. Marker sizes indicate the total number of mentions, within a given corpus, of locations at and below the national level (hence excluding supranational locations such as oceans and continents), scaled in proportion to that nation's share of the total number of place-name occurrences in each corpus. There is thus one marker per nation, each of which includes counts not just for invocations of the nation itself ("India," "United Kingdom"), but also for every place that falls entirely within the (modern) bounds of that nation ("New Delhi," "Trafalgar Square"). Markers are centered at the occurrence-weighted mean of the latitude and longitude of all locations within each nation. This choice for marker center means that marker locations bear more than the usual amount of information, since they indicate not just the identity of the nation in question and the fraction of attention devoted to it, but also summarize the spatial distribution of textual attention within its borders. This representation does, however, produce on rare occasions readable but anomalous results, as when Canada's marker edges into northern Minnesota (blame the prominence of Toronto and Montreal) and Portugal's falls somewhere between the island of Madeira (with its toponymic wine) and the mainland.
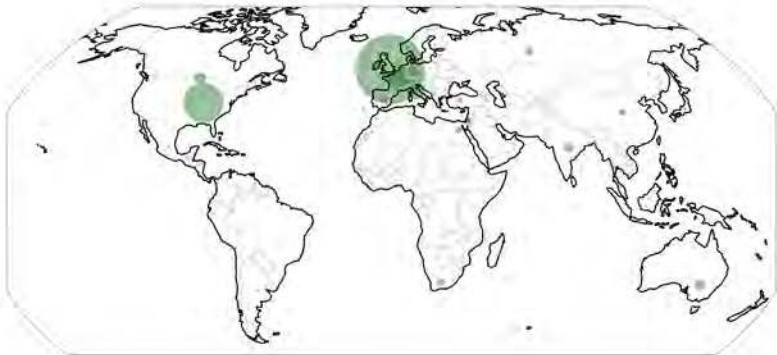
Hathi
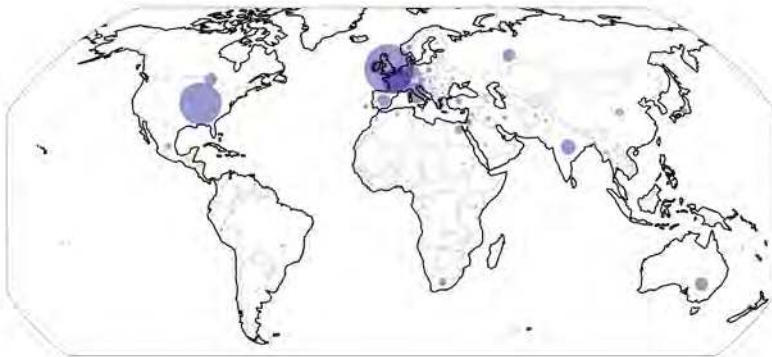
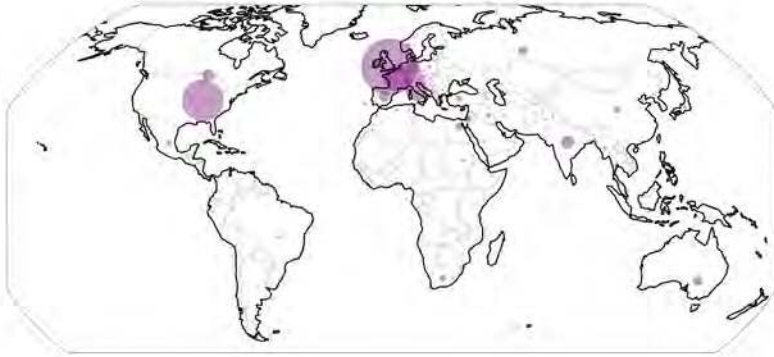**Foreign**
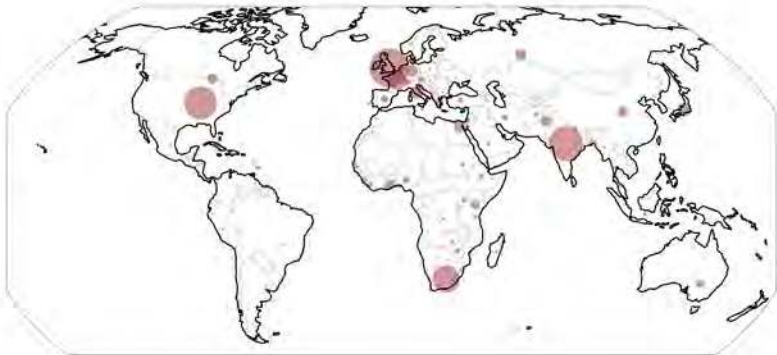
London

Prominent

Figure 2. Distribution of global attention in the four corpora, aggregated by nation and scaled by total location occurrences per corpus. See text for an explanation of marker sizes and locations.

While the patterns of attention at the global level do differ in important ways, we may note first that they are not *wholly* different. This phenomenon has been previously described in other national contexts[44] and the same conclusion holds here, albeit prospectively: the absolute magnitude of changes in the quantifiable characteristics of large literary corpora are often smaller than the conventional literary-critical emphasis on difference and discontinuity would lead us to expect. But small changes are not necessarily unimportant changes.

In the present case, the foreign-authored corpus reflects a substantially more diverse international outlook than does any of the other corpora. This is true not only with respect to the United Kingdom, which accounts for just 18% of all location occurrences in that corpus compared to as much as 57% of the London corpus, but also to other European nations and to the United States.[45] The geographic attention displaced from such Western sites is reallocated, among foreign writers, above all to India, China, Japan, South Africa, and the Caribbean, reflecting a widespread tendency of writers across corpora and periods to devote statistically disproportionate attention to their spaces of origin.[46] Foreign au-

---

[44]Wilkens, "The Geographic Imagination."

[45]We calculate 95% confidence intervals for all reported averages and p-values for all reported statistical comparisons in the code supplement to this article. Confidence intervals and p-values are included selectively inline. Unless otherwise indicated, we report no non-significant ($p > 0.05$) statistical comparisons. In the present case, the values are 14.9-20.6% UK domestic locations in the Foreign corpus vs. 54.1-59.3% domestic locations in the London corpus ($p = 1.4e\text{-}57$), taking the occurrence-weighted ratio of domestic to international locations in each volume as an observation.

[46]Concerning home-nation biases in literary fiction, see also Matthew Wilkens, "The Perpetual

thors were often writing for British and international audiences alike, and many of their books were specifically intended to decode British life from the comparative perspective of an "extimate" colonial subject.

In keeping with critical expectations concerning the internationalization of literature during the modernist era, we observe that the fraction of all locations that fall outside the UK in each of our corpora in the period 1880-1913, prior to the outbreak of the First World War, is less than that in the period 1914-1940 (with the exception of the London corpus, which is statistically flat). The changes in international attention over time are summarized in table 1 and visualized in figure 3. Beyond the observed rise in international attention, a comparison between the Prominent and Hathi corpora is telling. Although critics typically associate internationalization in the period most strongly with canonical fiction and especially with the thematic concerns of high modernism, the broad-spectrum record of the Hathi corpus shows meaningfully greater attention to locations outside the UK consistently across the period 1880-1940. We thus find our hypothesis concerning increasing international attention in the long modernist era to be well supported by our geographic evidence, but subject to a surprising inversion of expectations with respect to canonical and noncanonical (or, in Algee-Hewitt et al.'s formulation, "archival") sources.[47]

Corpus	Int'l fraction, 1880-1913	Int'l fraction, 1914-1940	p
Foreign	0.734	0.889	1.3e-8
Hathi	0.703	0.735	1.8e-12
London	0.425	0.443	0.51
Prominent	0.596	0.650	0.001

Table 1. Locations outside the UK as a fraction of all locations by corpus, grouped before and after 1914.

Hathi

Fifties of American Fiction," *Neoliberalism and Contemporary Literary Culture*, eds. Mitchum Huehls and Rachel Greenwald-Smith (Baltimore: Johns Hopkins UP, 2016, pp. 181-202).
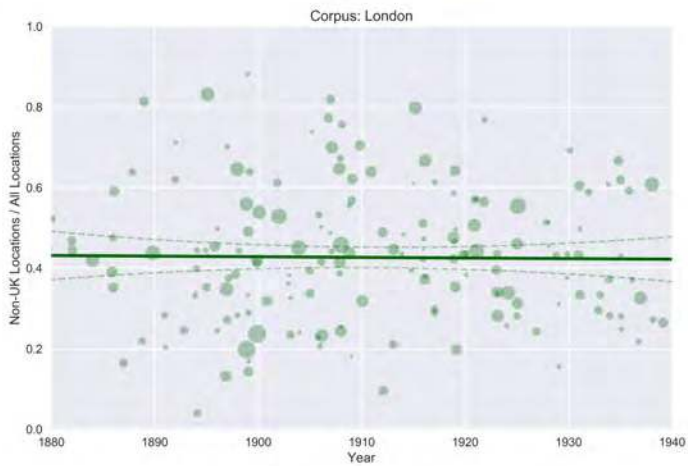
[47] Mark Algee-Hewitt, Sarah Allison, Marissa Gemma, Ryan Heuser, Franco Moretti, and Hannah Walser. "Canon/Archive. Large-scale Dynamics in the Literary Field," *Pamphlets of the Stanford Literary Lab*, vol. 11. Stanford, CA: Stanford University, 2016.
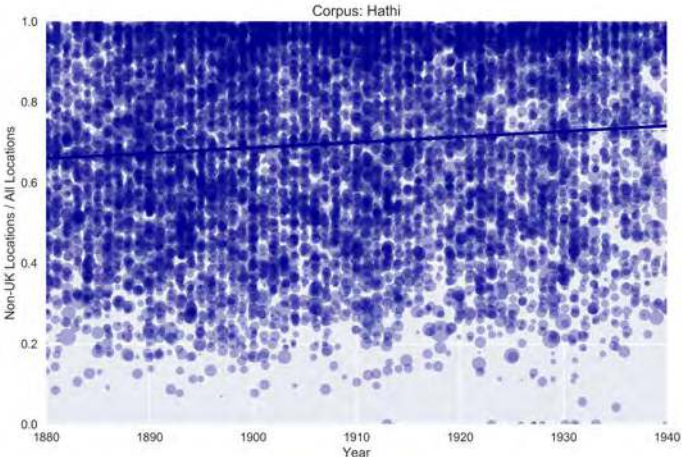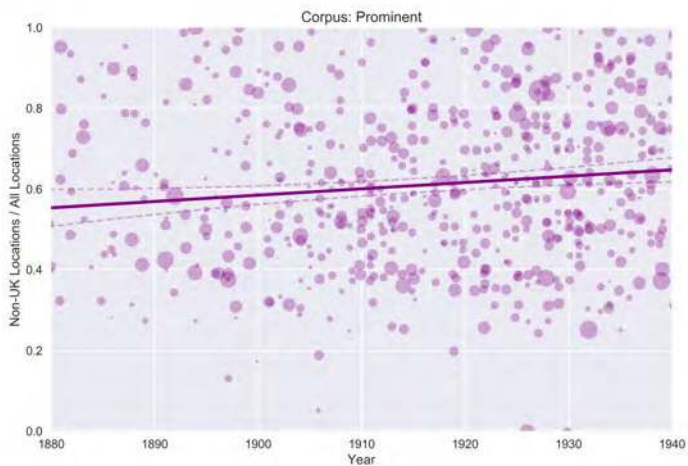
Foreign
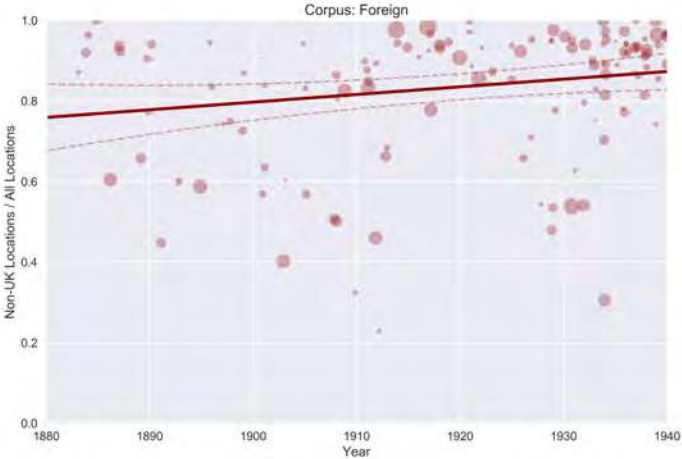
London

Prominent

Figure 3. Fraction of all location mentions devoted to places outside Great Britain, by date of publication. Each marker represents one volume, sized in proportion to the total location mentions per 100,000 words in that volume.

Books by foreign writers are also more geographically intensive overall, contain-

ing over twice as many location mentions per hundred thousand words as those belonging to any of the other corpora (which are statistically indistinguishable from one another in this regard). The magnitude of this result is due in large part to the generic diversity of the Foreign corpus; its nonfiction volumes use named locations at the rate of 843 ± 115 per 100,000 words, compared to 345 ± 66 in the fiction volumes (p = 7.3e-11). Still, Foreign fiction is significantly more geographically intensive than any of the other, fiction-only corpora (Hathi, 281 ± 4, p = 0.013; Prominent, 267 ± 13, p = 5.9e-4; London, 273 ± 22, p = 0.011). For reference, prominent writers whose work is near the average geographic intensity of volumes in the Prominent corpus include G.K. Chesterton, Elizabeth Bowen, Graham Greene, and James Joyce. Virginia Woolf's books use fewer location mentions (212 per 100,000 words, on average). (Readers may also wish to explore an interactive visualization of geotypicality in five dimensions among authors in all of the corpora.)[48]

The case is slightly more complex if we ask, in keeping with our hypothesis concerning intensity, whether these values rose over time. In the Foreign and Prominent corpora, the answer is no; in those groups, intensity fell across the 1914 boundary, dropping in the former to 533 location mentions per 100,000 words in the interwar period from 781 (p = 0.004) before the war and in the latter to 253 from 288 (p = 0.006). The Foreign result is complicated by a corpus composition shift toward fiction in the latter period (to 48% of volumes from 41%), but the Prominent result is unambiguous. In the Hathi and London corpora, geographic intensity is statistically unchanged across the periods before and after the war. To a first approximation then, the hypothesis of rising geographic intensity is not supported.

A reasonable alternative reading of Hegglund, however, might interpret "topographic detail" as not so much a matter of frequency as of specificity; perhaps interwar literature preferred the specific to the general, "Bloomsbury" to "Britain" and "Bombay" to "India." To assess this version of the claim, we calculated the fraction of location mentions that fall below the level of the country or the city, both in the UK and abroad. The results, however, are no different in sum. The Hathi and Prominent corpora became somewhat *less* specific in the interwar period, while the Foreign and London corpora were, for the most part, statistically unchanged.[49] We note that, unlike the intensity case, we did not observe perva-

---

[48] As a piece of trivia, we note that either Graham Greene or Ronald Firbank is the most geographically "typical" author in the Prominent corpus (depending on whether the average is calculated relative to the Prominent corpus alone or to all four corpora). Rebecca West and Aldous Huxley are also near the mean. For details of the method, see the linked visualization.

[49] It is not an accident that statistical significance is more difficult to achieve in the smaller corpora, especially when subdividing them by publication era.

sive generic effects with respect to specificity in the Foreign corpus; fiction and nonfiction texts were geographically specific at roughly equal rates.

We did, however, observe significant differences between corpora in their rates of specificity across the full period 1880-1940. Within the UK, the London corpus was the most likely to mention locations below the city level, containing nearly twice the fraction of such uses relative to the other corpora (40.2% vs. 17-27%, $p <$ 1.4e-17). Books in the Prominent corpus were more likely than those in the larger Hathi corpus to use sub-city-level locations in the UK. W. Somerset Maugham, George Orwell, and Thomas Hardy were near the Prominent average; Virginia Woolf was far above average, James Joyce somewhat below (though much higher in Ireland, which, to recall, is not counted as a UK location).

Globally, the Foreign corpus again stood out, although not in the way one might expect. Foreign volumes, despite their high geographic usage rates and preponderance of international locations, used the lowest fraction of locations below the country level outside the UK. That is, books in the Foreign corpus were more likely to refer to nations as such rather than to districts, cities, landmarks, and more specific locations within those nations than were the books in any of the other corpora (50% below the country level in the Foreign corpus vs. 67-69% in the others, $p <$ 4.5e-19). We tentatively attribute this finding, which accords well with research by Boehmer, Fryer, and Macdonald, to foreign writers' greater investment in geopolitical subject matter compared to their Britain-born peers, with the result that texts by foreign writers were more likely to name nations and their relations, while other writers, when they used foreign locations at all, used them disproportionately as settings rather than political entities.

Texts drawn from the London corpus featured the highest ratio of UK domestic locations; these are, after all, books selected specifically for their focus on London as a British regional site. We note in passing that it was the Continent that supplied most of the sacrificed international attention in the London case. The Prominent corpus hewed more closely to the average domestic-international split observed in the broad Hathi corpus, but still skewed domestic in comparison (37% domestic in Prominent vs. 28% domestic in Hathi, $p =$ 3.0e-20). This is a surprising finding, given canonical literature's reputation for internationalism in the period.[50]

Within the UK, these trends were in some ways reversed: where foreign authors made use of significantly more international diversity in their texts, their UK lo-

---

[50] Ford Madox Ford, Henry James, Aldous Huxley, E.M. Forster, and H.G. Wells were all near the Prominent average in their rate of use of locations outside the UK. Joyce was notably high—Ireland being outside the UK—while Woolf was very low.

cations were more heavily concentrated in London than were those of the large Hathi corpus. (A cartographic overview of the differences between the corpora is provided in figure 4.) It is the Hathi corpus, in fact, that was the most domestically diverse, making use of locations in Greater London (including references to "London" itself) for just 40% of its total UK place occurrences.

Hathi
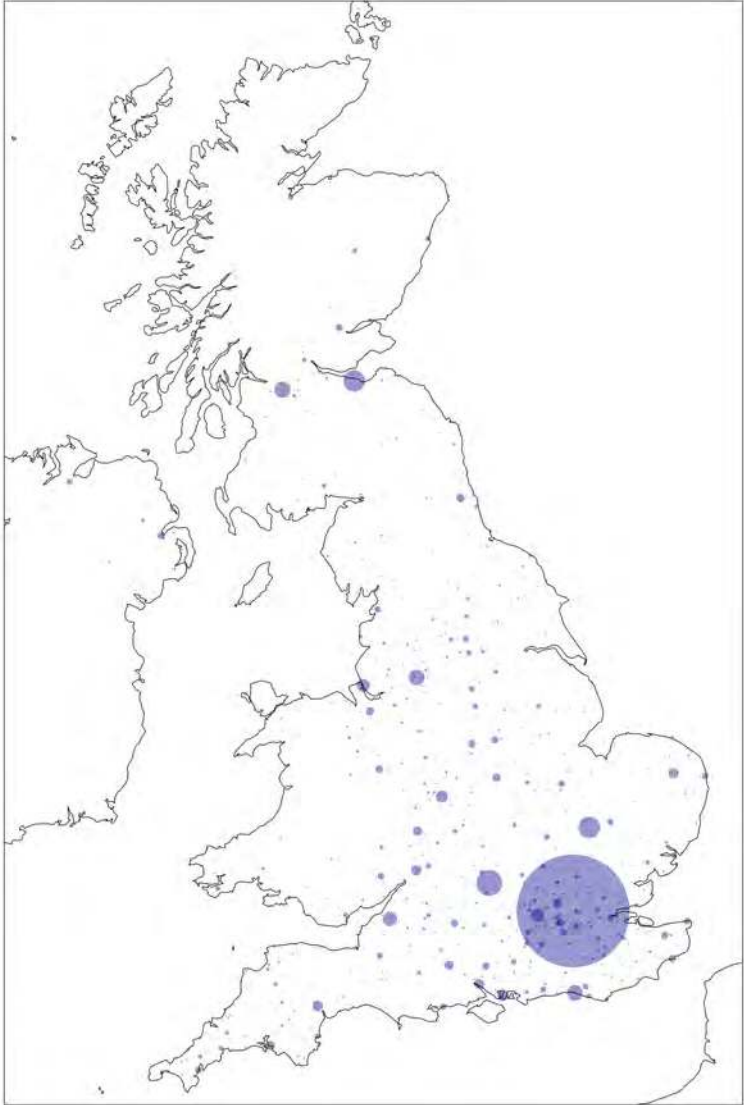
Foreign
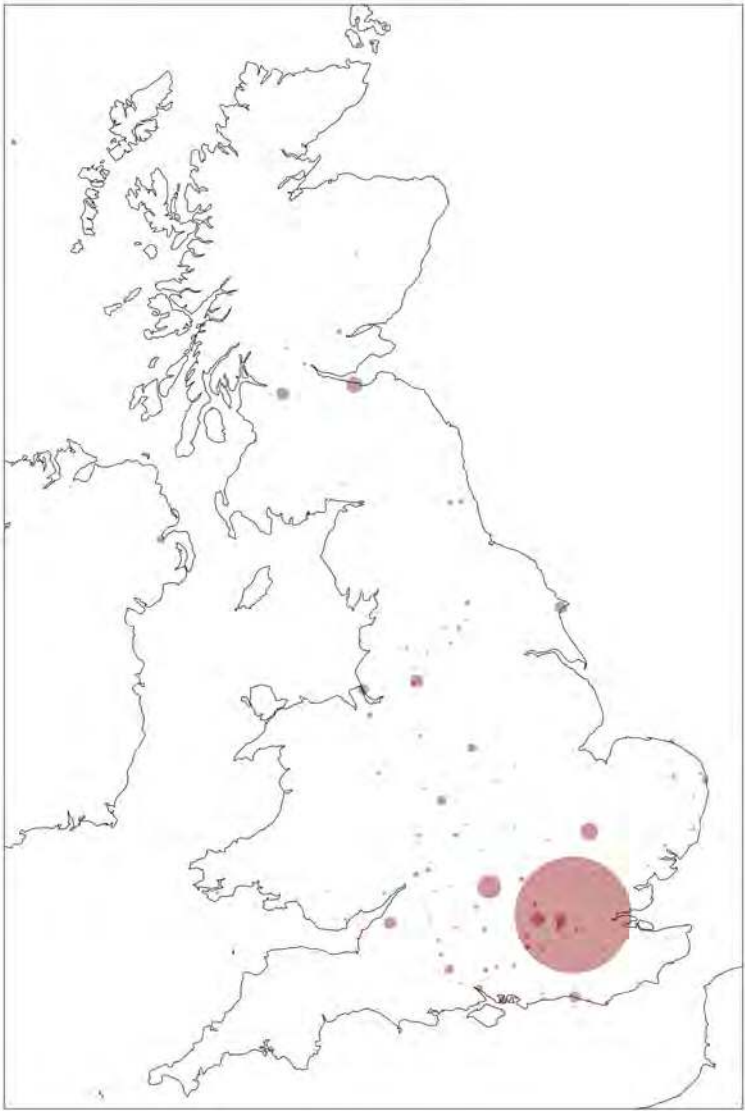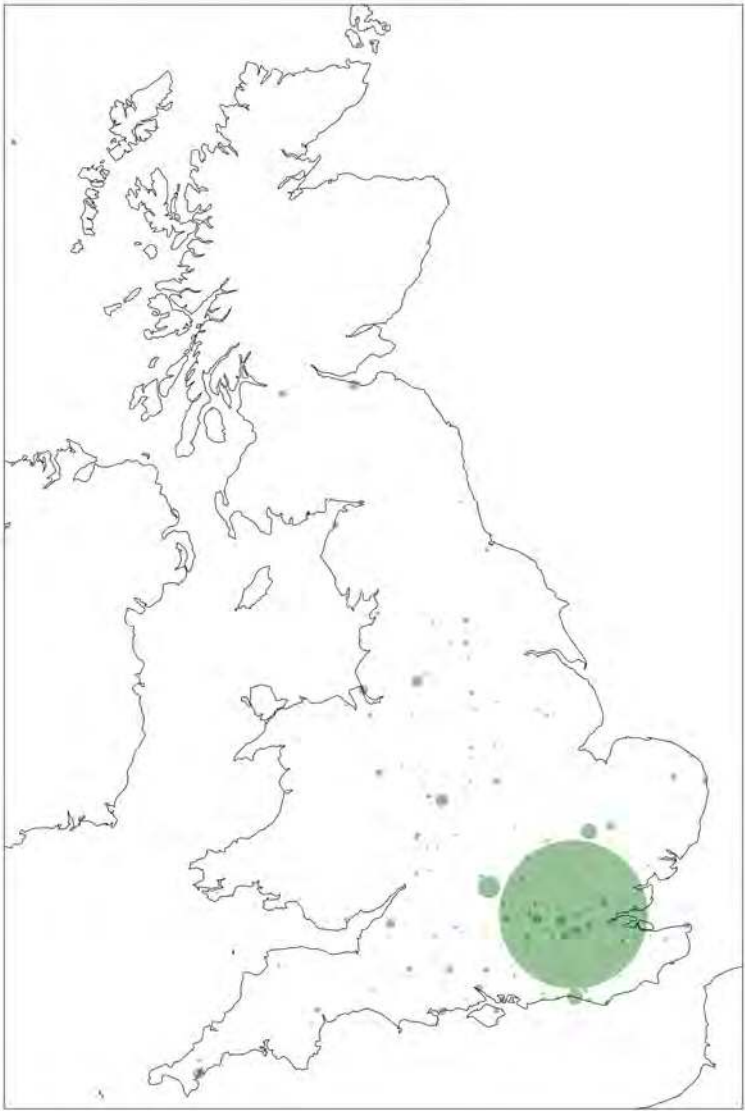
London

Prominent

Figure 4. Distribution of UK attention in four corpora, aggregated by locality (city) and sized by fraction of total location mentions in each corpus.

The London corpus was, little surprise, the most focused on London places, which account for 69% of all its UK domestic location occurrences. The Prominent corpus c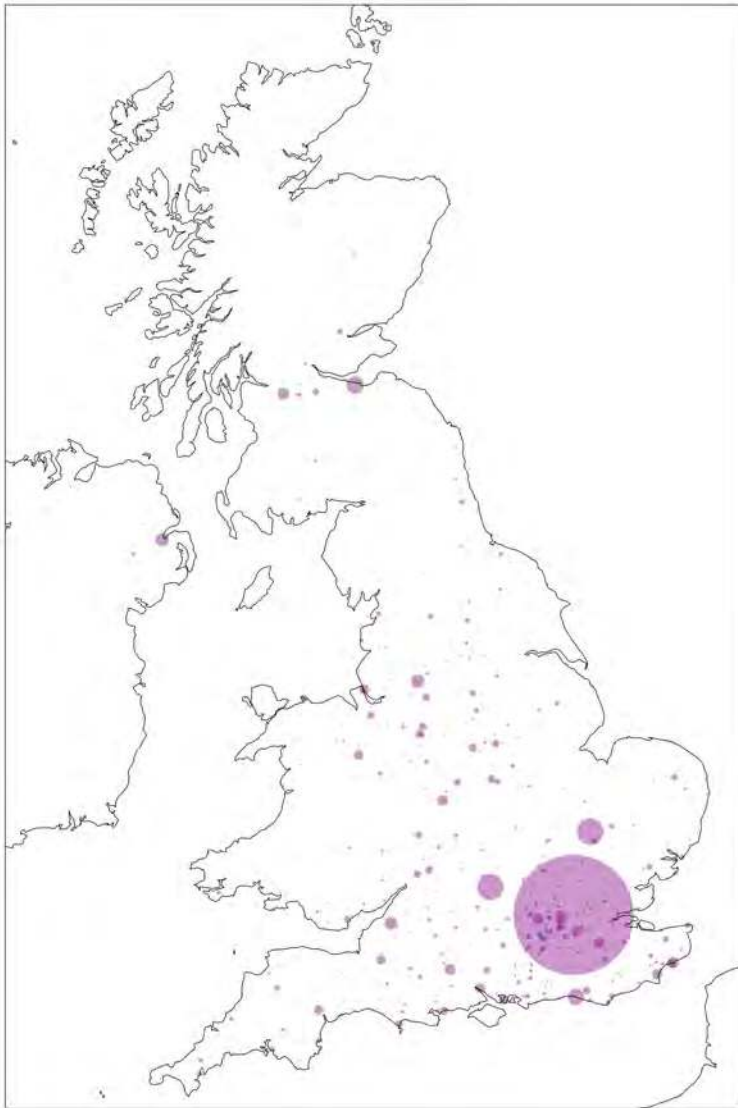losely resembled the foreign-authored set on this metric, the two using London locations for 46% and 43% of their UK mentions, respectively ($p = 0.08$, not significant).

If we take London usage rate as a plausible proxy for attention to urbanization and urban issues (subject to some caveats explored below), we find general confirmation of the long-standing critical claim that relatively canonical writing at the end of the nineteenth century and the first half of the twentieth century (and modernism in particular) was especially preoccupied with urban spaces. Indeed, texts in the corpus of prominent British fiction devoted about 6 percentage points more of their domestic attention to places in London than did those in the Hathi corpus. So it does appear that there's a detectable difference, a bend toward the urban, in texts we think of as canonical compared to British fiction more generally.[51] Foreign writers weren't much different from Prominent writers in this regard, both groups finding, perhaps, a comparatively congenial home in the metropole. But the existence of the London corpus and its singular geographic focus demonstrate that there existed a parallel trend in less elevated British writing of the same period. So, while it may be true that outsider and (relatively) canonical period texts differentiated themselves in part through their interest in the city, theirs was an investment shared by pockets of popular genre and regional fiction as well.

Interesting variations between the corpora also appeared in their relative attention to locations within London, which are mapped in figure 5 and summarized via centers of gravity (weighted mean latitude and longitude of their London locations) in figure 6. A list of the most frequently occurring London locations is presented in table 2. As is clear in the summary measure of figure 6, books in the Prominent corpus favored, on average, locations in the wealthier and more fashionable West End, while the London corpus leaned toward the working-class East End. The Hathi corpus stood in the middle; the Foreign corpus edged west, toward the Prominent average, but was statistically indistinguishable from either the Hathi or the Prominent corpus. We can observe notable change over time in each of the corpora with the exception of the Foreign set (where statistical uncertainty is large). The center of gravity of the London corpus moved further east after 1914 (the change is relatively large, but not statistically significant due to the

---

[51] Katherine Mansfield was near average in the fraction of London locations among all British locations in her work, as were Evelyn Waugh and Rebecca West. Woolf was well above average, Joyce well below. The case of Joyce emphasizes, again, both his imperfect fit as a British writer and the potential limitations, in any single case, of using London as a proxy for the urban. Further details of London usage across the corpora follow immediately below.

small size of the corpus), while the Hathi corpus drifted slightly west (borderline significant) and the Prominent corpus moved mostly north.

The dynamics driving these differences, both between corpora and across time, are complex. To understand the details (including those not well captured by averages), to explain the intermediate positions of the Hathi and Foreign corpora, and to see why the hypothesis of rising literary-geographic intensity strikes many critics as plausible despite not being borne out at scale, we need to examine specific locations.

Hathi

Foreign

London

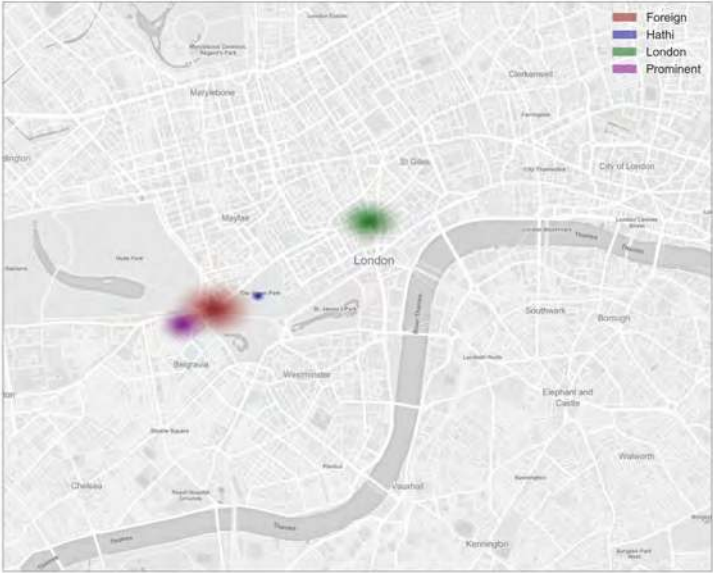**Prominent**

Figure 5. Distribution of geographic attention within London, aggregated by point location, for each of the four corpora. Note that generic references to "London" are excluded from these maps. Markers are scaled to represent each location's fraction of total location mentions (not restricted to London) in the relevant corpus.
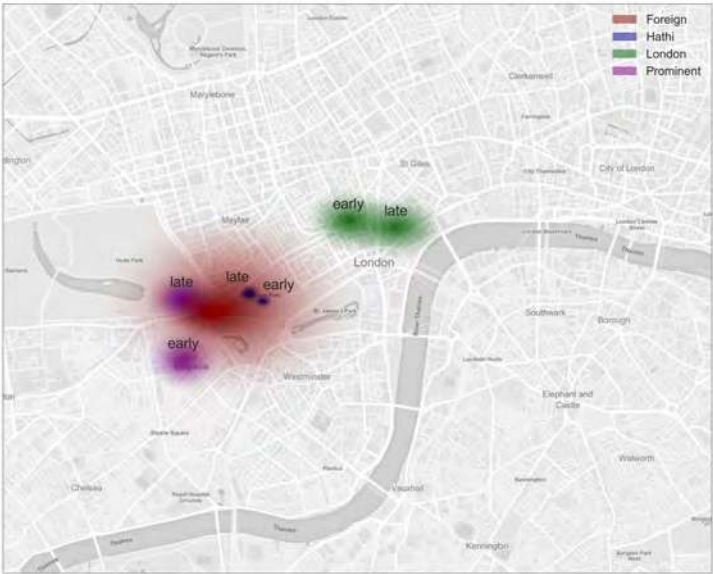
(a) By corpus

(b) By corpus and era

Figure 6. London centers of gravity with probability density clouds (a) by corpus and (b) by corpus and era (before and after 1914).

#	Hathi		Foreign		London		Prominent	
	Text	Frequency	Text	Frequency	Text	Frequency	Text	Frequency
1	London	20.89	London	35.70	London	46.49	London	26.40
2	Richmond	0.33	Richmond	0.66	Bloomsbury	3.05	Richmond	0.75
3	Thames	0.29	Thames	0.48	Paddington	2.20	Soho	0.44
4	Kensington	0.20	Hyde Park	0.46	Oxford Street	0.93	Oxford Street	0.39
5	Hyde Park	0.19	Oxford Street	0.43	Hampstead	0.82	Hampstead	0.39
6	Hampstead	0.18	Greenwich	0.33	Hyde Park	0.82	Putney	0.38
7	Westminster	0.16	Westminster	0.32	Greenwich	0.78	Kensington	0.38
8	Covent Garden	0.16	Kew	0.32	Limehouse	0.77	Thames	0.35
9	Pall Mall	0.15	Kensington Gardens	0.31	Soho	0.77	Hyde Park	0.34
10	Scotland Yard	0.15	London Bridge	0.29	Kensington	0.74	Chelsea	0.29
11	Chelsea	0.15	Kensington	0.29	Hoxton	0.68	Notting Hill	0.29
12	Oxford Street	0.15	Hampstead	0.27	Islington	0.66	Croydon	0.27
13	Greenwich	0.13	Brixton	0.26	West End	0.65	Bloomsbury	0.25
14	Kew	0.13	Buckingham Palace	0.25	Clerkenwell	0.63	Trafalgar Square	0.25
15	Kingston	0.12	Bloomsbury	0.23	London Bridge	0.62	Scotland Yard	0.24
16	Chatham	0.12	East London	0.23	Scotland Yard	0.58	South Kensington	0.24
17	Paddington	0.12	Downing Street	0.22	Park Lane	0.55	West End	0.23
18	Soho	0.12	Kingston	0.16	East End	0.52	Kensington Gardens	0.23
19	London Bridge	0.12	Chelsea	0.15	Thames	0.52	Greenwich	0.21
20	Park Lane	0.11	Camden Town	0.14	Highgate	0.51	Roehampton	0.21
21	Islington	0.11	Kensal Rise	0.14	Bayswater	0.49	Park Lane	0.20
22	Bloomsbury	0.10	Park Lane	0.13	Trafalgar Square	0.49	Covent Garden	0.19
23	West End	0.10	South Kensington Museum	0.13	Covent Garden	0.48	Paddington	0.18
24	Kensington Gardens	0.10	Chatham	0.13	Chelsea	0.44	Westminster	0.18
25	Putney	0.09	Trafalgar Square	0.12	Grosvenor Square	0.44	Tottenham Court Road	0.17
26	Haymarket	0.09	Wardour Street	0.12	Wimpole Street	0.41	Pall Mall	0.16
27	East End	0.09	Holborn	0.12	Tottenham Court Road	0.39	Bayswater	0.15
28	Grosvenor Square	0.09	Mansion House	0.12	Praed Street	0.39	Berkeley Square	0.15
29	Trafalgar Square	0.08	Hammersmith	0.12	Haymarket	0.38	Red Lion Street	0.14
30	Croydon	0.08	East End Of London	0.12	Richmond	0.37	Hyde Park Corner	0.14

Table 2. London locations occurring most frequently per 100,000 words in the four corpora. Full lists are available in the data deposit.

Class is a structuring force of London geography, with the West End known for its theaters, clubs, high-end shopping, and exclusive residential neighborhoods, and the East End long condemned, mourned, or studied as a region of poverty, crime, and degeneracy (it was "Darkest England" in one best-selling book of that title). The East End also included the greatest diversity among its residents' ethnic and religious affiliations, perceived race, and national origins. We selected specific sites that might serve as shorthand for the opposed regions and their particular connotations in order to illuminate their relative rate of representation in the corpora. To compare the degree of attention to wealthy areas across the corpora, we grouped mentions of Chelsea, Mayfair, Belgravia, Bond Street, and

Park Lane. To capture differences in the rate at which poor, East End locations were mentioned, we grouped together Whitechapel, Limehouse, Mile End Road, West India Dock Road, and East India Dock Road, as well as the broader labels East End and East London.

The cluster of rich places received the greatest proportional attention in the London and Prominent corpora, with Hathi following closely and Foreign trailing well behind (cross comparisons yield statistical significance only in the cases of the Foreign corpus vs. the others). It is not surprising that the Prominent corpus, with its interest in well-to-do "society" would devote ample attention to the environments of the rich. The near equal rate of occurrences in the London corpus is less expected but the sheer number of named London locations in that corpus makes it more likely that any particular place within the city would be named. The relatively small degree of attention paid by the Foreign corpus may, we hypothesize, be explained by these "outsider" writers' lesser degree of access to the homes and social spaces of the elite. The well-to-do Mayfair, for example, didn't make it into the top 100 named London locations of the Foreign corpus, though it appeared on those of the other three.

Our collection of poor, East End locations was, by a large margin, used most in the London corpus. As table 2 shows, the London corpus was the only one to include Limehouse or East End among its top twenty named London locations. Collectively, the relative frequency of poor areas in the London corpus was more than four times that of any of the other corpora, which each named those locations at roughly the same rate. The London corpus contained a greater number of books that capitalized on the popularity of quasi-sociological or sensational depictions of poverty (Israel Zangwill's *Children of the Ghetto* and Thomas Burke's *Limehouse Nights*, for example) and on the enduring draw of crime fiction (such as Arthur Morrison's *The Hole in the Wall*), leading to the easterly skew to its center of gravity, already remarked. It is notable that our group of foreign-born writers were statistically no more likely than the majority of their British counterparts to write about the East End even though residents who shared the writers' nations of origin tended to live in that part of the metropolis.[52] It may be that the writers represented in the Foreign corpus, who were overwhelmingly students and intellectuals from Britain's colonies, felt as unwelcome in areas of poverty as they did in areas of wealth or were not predominantly interested in the dock workers and manual laborers who most often lived in the East of the city.

The most familiar neighborhood of London for readers of literary modernism is Bloomsbury, which we find named relatively frequently in all of the corpora. It captured the highest number of relative occurrences in the London corpus,

---

[52]Of course, the trend may have been different for other groups of foreign-born writers.

in which it was the single most frequently named location within Greater London.[53] While the fact that many writers devoted attention to Bloomsbury isn't entirely surprising, neither is its striking prominence straightforwardly expected; the Bloomsbury group of writers and artists, after all, wasn't active until late in our period, and little of the London corpus is made up by their books. Texts from the more squarely canonical Prominent list did not mention Bloomsbury as frequently, nor did those in the Foreign corpus (Bloomsbury ranked 13th and 15th, respectively, among those corpora's London locations). But the fact that Bloomsbury occupied such a notable place in the geographic imagination of a wide range of writers provides further evidence for Sara Blair's assertion that Bloomsbury was important as a London hub before it gained fame through its association with its toponymous artistic group. What we know of Bloomsbury on historical grounds, moreover, makes its frequent use within the London corpus especially interesting; Bloomsbury was among the most "international" of domestic locations in the city and in the British nation, suggesting that, even within domestically oriented writing of the type associated with our London corpus, there was a legible desire for foreign heterogeneity.

Relative mentions of Bloomsbury were lowest in the Hathi corpus, confirming what we suspect, that the books we read—exemplified by the Prominent corpus—were not typical of all that was published, though it should be noted that they were also not remarkably far off in this case (Bloomsbury comes in 22nd in the larger corpus). This pattern of lowest mention in Hathi continues when we compare the relative frequency of mentions of the primary roads that defined the Bloomsbury area (Tottenham Court Road, Euston Road, and Gray's Inn Road) and of Russell Square at its center, which provides some reassurance that the Bloomsbury toponym was indeed primarily geographical.[54]

More startling, perhaps, is the discovery that Soho significantly outshone Bloomsbury in the Prominent corpus (it was second only to Richmond, which lay outside of London in the period concerned). Soho was of least interest in the Foreign corpus, though, in 41st position, it was hardly overlooked. This is an interesting discovery because Soho, like Bloomsbury, was an area known for cosmopolitanism. The foreign residents of Soho were primarily from Europe, but in the last decades of our period the always-lively neighborhood became known for its black night clubs.[55] Strikingly, then, the Foreign corpus demonstrated relatively little interest

---

[53] Bloomsbury is an exemplary case of the fundamental need for interpretive interventions concerning toponyms. Mentions of the neighborhood and of the group alike are counted in our data, on the theory that the latter also represent meaningful reference to a specific place.

[54] The smaller arteries Bloomsbury Way and Theobalds Road also bordered the area (as described by Weinrob et al., *The London Encyclopaedia*, 78) but they are not in the top 100 of any of the lists.

[55] See Judith Walkowitz's book on Soho, *Nights Out: Life in Cosmopolitan London*, (New Haven:

in two of London's most heterogeneous neighborhoods, just as it paid relatively little attention to areas of extreme wealth and poverty.

If Bloomsbury was not of particular interest to the writers of the Foreign corpus, the same cannot be said of the area's most famous landmark. The British Museum had the greatest rate of mentions among Foreign writers (whose corpus was the only one to include it in the top 100). The Museum was a popular tourist site throughout the period under investigation, as it is today, but its singular importance for writers born outside Great Britain may also have reflected the significance of the British Museum Reading Room for international students and intellectuals who came to London to read and to write. When the British Museum is considered together with several other cultural sites—the National Gallery, London Library, and South Kensington Museum—the relative cumulative mentions of these cultural institutions were highest in the Foreign and Prominent corpora, in which they were a statistical tie. These cultural institutions were least important in the London corpus, perhaps another reflection of the preponderance of genre fiction within it.

The comparison of top London locations across the corpora produced other discoveries about the singularity of the Foreign corpus that we find compelling. It revealed that the Foreign corpus contained a significantly higher rate of mentions of three types of public sites. First, Foreign writers paid greatest relative attention to the city's parks and gardens. When we combine central London's Hyde Park, Green Park, St. James's Park, Regent's Park, and Kensington Gardens with the further-flung Kew (south of the Thames, now in the London Borough of Richmond), the Foreign corpus had over 30% more relative mentions than any of the other corpora.[56]

Peter Kalliney has pointed to the importance of the public parks as "one of the few venues in which recent immigrants and white Londoners meet on relatively equal terms" in the "emerging postimperial metropolis" of Selvon's *The Lonely Londoners* (108 and 106). Our findings suggest that parks were also significant in the earlier experiences of colonial subjects in pre-1940 imperial London, particularly for writers of color, as we discuss below.[57] A second, perhaps related,

---

Yale UP, 2012) for more on both these aspects (especially pp. 22 and 232-52).

[56] Of London's many parks and gardens, Hyde Park was the most frequently mentioned green space across all the corpora. In the London corpus, the only green spaces to fall in the top 50 named London locations were Hyde Park and Greenwich. We exclude Greenwich from the collection of parks and gardens because its maritime history gave it a distinctive cultural function not shared by the others. Popular among tourists and London day-trippers alike, Greenwich was frequently mentioned across all corpora, with the least attention in the Prominent corpus.

[57] The prominence of parks in this corpus is an area for further exploration, for, as Kalliney discusses in the context of Woolf's *Mrs. Dalloway*, London's public parks, replete with imperial monu-

type of public space that received greatest relative interest in the Foreign corpus was river places. The Thames and four of its bridges - Westminster, Waterloo, London, and Blackfriars - received, in sum, 20-40% more attention in the Foreign corpus than in the others, though only the comparison with the Prominent corpus approached statistical significance.

Finally, we note that the Foreign corpus contained the greatest proportional mentions of the politically inflected sites Westminster, Downing Street, and Hyde Park Corner, taken as a group. One might explain the greater prominence of these locations in the Foreign corpus as deriving from the sites' touristic value, as could be the case with the parks and cultural institutions discussed above. However, the anti-imperial movements increasingly active between 1880 and 1940 suggest that at least part of their significance among foreign writers was political in nature. That Scotland Yard, by contrast, received the least relative attention in the Foreign corpus suggests both that writers born and raised outside of Great Britain were more interested in policy than policing and that they tended to work outside the generic concerns typical of popular fiction.[58] The reverse could be said of the London corpus, in which Westminster was ranked lower than in any of the other sets. The Foreign corpus, then, was distinct from the corpora dominated by native British writers in that it attached greater significance to London's parks and gardens, river places, and political sites. It was notable, too, for its comparatively high rate of interest in cultural institutions, equivalent to that in the Prominent corpus, and for its lesser attention to neighborhoods of extreme wealth and poverty and to cosmopolitan Soho.

To explore the effects of ethnic difference on the literary-geographic attention of non-native authors, we compared the distribution of attention within the Foreign corpus between writers who identified (or who were identified) as white and those who did (or were) not. This was an obvious, important, and often visible source of social differentiation that corresponds only very imperfectly to any of the other divisions between our corpora. In many ways, the observed differences are modest; nonwhite and white foreign authors' usage rates of international and domestic locations, of our selected political and cultural sites in London, as well as London's share of their British spatial attention, were statistically indistinguish-

---

ments, "were one of the primary spaces in which the English celebrated their colonial triumphs, articulated their imperial vision, and educated themselves about the gravity of their position as rulers." Peter J. Kalliney, *Cities of Affluence and Anger: A Literary Geography of Modern Englishness* (Charlottesville: University of Virginia Press, 2006), 86.

[58] For more on the dynamics of genre in the nineteenth and early twentieth centuries, considered from a similarly computational angle, see the cluster of articles on that theme published in CA in late 2016.

able, and the patterns of their global usage, while distinct, seem to reflect primarily the specific identities of their home countries, a pattern that we have already seen at both large and small scales and that seems linked more closely to national origin than to ethnicity.

Meaningful differences did exist in several areas. White writers in the Foreign corpus used global locations that were specific below the city level at rates higher than their nonwhite counterparts (21% vs 17%, $p = 0.01$). If our hypothesis that such specificity was inversely correlated with political content in the literature of the period is correct—that is, that specificity is often a matter of political abstraction vs. literary-narrative detail—this fact suggests that the nonwhite Foreign writers in our corpus were responsible for texts that were comparatively politically intensive. This finding accords well with both the existing critical literature on colonial-era writing and with our own sense of the texts that make up the Foreign corpus.

Also telling is the map of each group's geographic usage within London, where they were alike outsiders, but often of importantly different types.
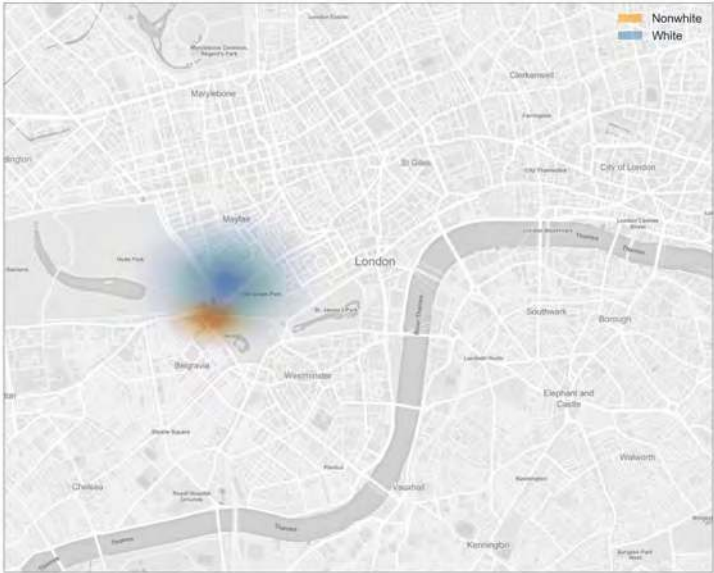
(a) Nonwhite
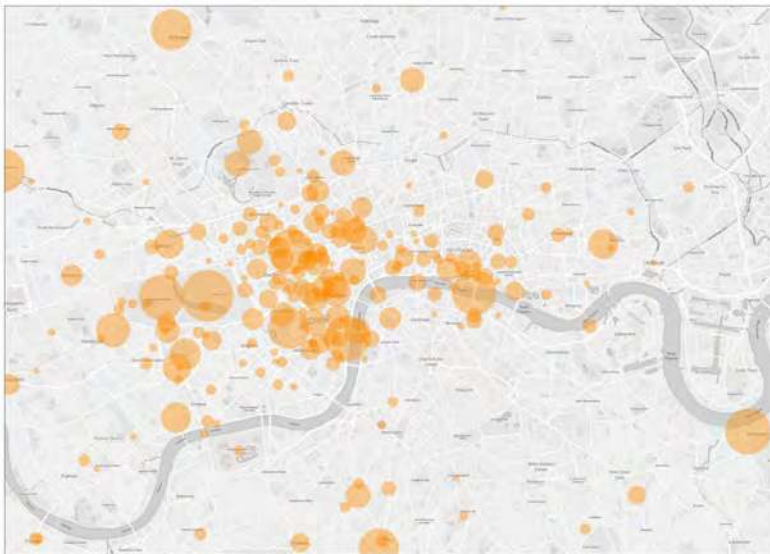
(b) White
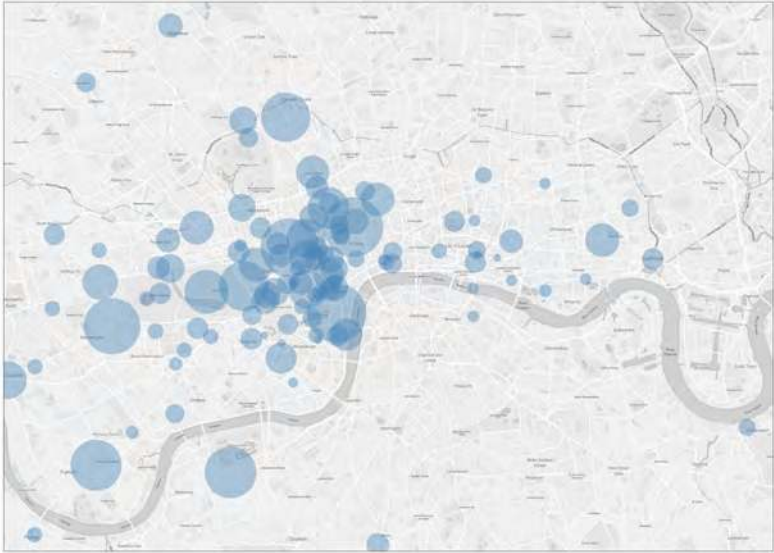
(c) By ethnicity

(d) By genre

Figure 7. Distribution of geographic attention in London within the Foreign corpus (a and b) aggregated by point location and distinguished by author ethnicity class; (c and d) centers of gravity by ethnicity class and by genre.

The greater variety of places mentioned by nonwhite writers is in part an artifact of their larger representation in the Foreign corpus; more books mean a greater chance that rare locations will be mentioned at least once. But the overall pattern is subtly different. The center of gravity of places used by nonwhite authors lies further south and a bit to the west of that for white authors. Nonwhite writers paid more attention to locations directly on the Thames, especially bridges and tourist-associated points. This pattern, in fact, more closely matches that of the Hathi and Prominent corpora than does that of the white foreign writers, potentially suggesting that, while the two groups were alike shaped by their status as outsiders, nonwhite writers presented a London slightly more closely aligned with that of their domestic British literary peers than did their less racialized foreign counterparts.

When we examine the set of prominent parks and green spaces that were used more frequently in the Foreign corpus than in any other, we find that it is texts by nonwhite writers that drive the whole of the difference; books by white au-

thors use parks at rates similar to those of the other corpora (which are made up almost exclusively of white-authored volumes). It is difficult to avoid the conclusion that what Kalliney argued with respect to Sam Selvon's post-1945 cultural milieu—that such public spaces were especially linked to the literary and political imagination of nonwhite writers in Britain—was broadly applicable in the period before World War II. It is our hope that they will be the object of increased critical attention in the future.

Conclusions

The amount of information we have presented is formidable and our results point in several directions. How can we summarize the patterns of shared and differential attention we have observed in our corpora? For one thing, we have produced new data that bear directly on existing critical arguments concerning the development of British fiction in the late nineteenth and early twentieth centuries. We have found that, in keeping with broadly shared expectations, the texts in all of our corpora devoted more of their geographic attention to locations outside Great Britain in the period after the Great War than they had in the decades leading up to it, suggesting that postwar literature was indeed marked by its increased internationalization. This change was steady, cumulative, and not obviously linked to any one moment or event. As we noted at the outset and detailed above (see figure 3 and related discussion), however, while a historical rise in international attention was shared across the corpora, the baseline levels from which that rise occurred were meaningfully different. Texts in the Prominent corpus were approaching by 1940 a level of non-British geographic use that Hathi texts had attained, on average, more than 60 years earlier. Both corpora lagged behind the Foreign set, especially by the end of the period.

We have, in this case, one instructively ambivalent answer to the question of whether or not the modernist-era books that critics are most likely to read serve as a reasonable proxy for the period's literature. The international trend that is recognizable in well-known texts reflected a development that was taking place more widely at the time; in that sense, insights gained from canonical fiction do not mislead. But well-known fiction was notably conservative when it came to representing the world beyond British borders and was therefore a trailing indicator of the wider literary culture's outward turn. In that sense, the books in the Prominent corpus are, collectively, a poor object on which to build an understanding of how British literature evolved between 1880 and 1940.

At some level, most critics already know this. They (and we) care about Woolf and Joyce and the like exactly because those authors' work did not represent the average or typical product of the early twentieth century. But we want to insist that there are cases—the present one very much included—in which such an appeal to exceptionalism fails. It fails because the phenomenon of interest within the canon—an increasing investment in the world beyond the borders of Britain—is one that emerged earlier and more dramatically elsewhere. It's hard to know, in general, when such a temporal misalignment will have been the historical case, because canons exist to contain the abundance of that "elsewhere." In this instance, however, there is good evidence that a move beyond the usual sources has given us a better, richer, more deeply contextualized understanding of canonical and archival sources alike.

The representational value of canonical texts is similarly mixed in the case of Hegglund's carefully articulated argument concerning "topographic intensity." The interwar period's writing does not appear to have become notably more specific or more intensive in its use of geographic space, whether in canonical sources or elsewhere. We have offered two possible explanations for this fact, both of which were probably at work across the diverse volumes of our corpora. In the case of specificity, less detailed use of geographic space can be found in some varieties of political writing, especially those devoted to international and colonial relations in which the actors and objects of attention are more likely to be nations than cities or individual sites. If such political content was increasingly incorporated into interwar literature, it would have resulted in lower average specificity across the corpora, all else being equal.

Concerning geographic intensity, which, like specificity, was flat or decreasing across the full period 1880-1940, we note only that this fact may be relevant to longstanding critical debates concerning the role of psychological interiority as a signal feature of modernist literature, on the assumption that narratives focused on states of mind may find geographic detail less suited to their purposes than would more externally directed accounts. If descriptions of interiority did consume a greater share of literary attention in the interwar years, this fact would help to explain why the era's geographic intensity failed to rise even if there existed a slight increase in baseline geographic or topographic interest. Critical investment in certain categories of spatial use may also explain why Hegglund's argument that geographic intensity or specificity rose in the modernist era seems plausible despite not being supported by our evidence. Select, notably famous locations such as Trafalgar Square, Bloomsbury, Soho, Buckingham Palace, and the like did see collectively increased use in the years after 1914. If critical attention were devoted to these places in preference to the full range of named locations

that appeared in the era's literature, then it is not difficult to see how an impression of rising geographic intensity, at least with respect to British fiction's most prominent city and in its most widely read books, could have formed.

We also recall that our task was in part descriptive: what was the distribution of literary-geographic attention in British fiction written and published during the six decades preceding the Second World War, previously unknown at large scale? To this question, we have some answers. British fiction of the period, considered as a pseudo-whole across more than 10,000 volumes and a billion words, was notably international, mentioning places outside the UK more than twice as often as those within it. When considering domestic locations, London was dominant even relative to its large population, accounting for more than twice the literary location share (circa 40% of the UK total) as the area did population share in 1900.[59] These figures represent a substantial difference from the American case, where literary attention in the nineteenth and twentieth centuries alike was more domestically focused, yet more diversely distributed within the home nation.[60] The sources of this national difference are several and difficult to trace in full, but are, we speculate, related to Britain's historically greater international role at the time (including its status as a more globally expansive imperial power), its relative prosperity and proximity to the Continent, its comparatively compact size, and, domestically, the singular importance of its largest city.[61]

In all of the corpora, geographic attention remains focused largely on a notional map of early- to mid-period Western industrial modernity, comprising principally locations in the UK, US, and Western Europe, though with historically conditioned attention to India that serves as one indication of the role of empire in shaping cultural production. The fit between this imagined map and, for example, global population is poor. We observe—as did Heuser et al. in their studies of earlier texts[62]—a similar distortion with respect to the population density of London itself, where literary locations represent only imperfectly the distribution of the city's habitation, reflecting a self-perpetuating set of traditionally

---

[59] For the historical populations of London and of Great Britain, see Julie Jefferies, "The UK Population: Past, Present, and Future," in *Focus on People and Migration* (London: Office of National Statistics, 2005) and "Population Growth in London 1939-2015," (London: Greater London Authority, 2015). In 1901, these numbers were 6.5 million and 38.8 million, respectively, meaning London accounted for about 17% of total UK population (excluding what is now the Republic of Ireland but including present-day Northern Ireland) at the turn of the last century.

[60] Wilkens, "Geographic," "Perpetual Fifties."

[61] France would thus provide an interesting comparative case, given its close alignment with the UK in many of these respects.

[62] Ryan Heuser, Mark Algee-Hewitt, Annalise Lockhart, Eric Steiner, and Van Tran, "Mapping the Emotions of London in Fiction, 700-1900," *Literary Mapping in the Digital Age*, edited by David Cooper, Christopher Donaldson, and Patricia Murrieta-Flores (London: Routledge, 2016), 25-46.

significant sites. This fact in turn suggests one of the sources of the stability we note across corpora and time: cultural forms have conservative inertial weight, making their sites, styles, forms, and subject matters more likely to be carried forward in preference to new forms that compete to displace them without the benefits of incumbency.

Given the overall conservatism of aspects of literary geography, even modest differences between corpora and subsets can be vitally important. It is our sense that one of the challenges of quantitatively oriented literary criticism in the years ahead will be to calibrate our expectations concerning the magnitude of important shifts in the measurable properties of our texts. While there is obviously no single standard to be applied across research projects, these are important conversations to foster by way of specific results. In the present case, we have seen statistical significance attached to differences as small as a few percent, especially in the large Hathi corpus, but we note that the technical sense of "significant" doesn't always imply the colloquial sense of importance. In the discussion above, we have taken care to present findings that we believe to be meaningful in the latter sense while also testing for the former sense. That said, some instances are easier; there are times when we observe strikingly large quantitative and qualitative differences between texts at scale, as, for example, in the cases of the international and London fractions of the four corpora. When we construct a corpus around a specific textual facet, it is possible to see very large deviations in that facet compared to a baseline corpus. This won't always be the case, of course, but the fact that it's possible argues for the standard use of reference corpora for computational and qualitative criticism alike.

What we do not observe are large, sudden changes in the features we examined within any one corpus. Surely such changes do sometimes occur, but we don't see them here, nor have they emerged in many other quantitatively oriented literary projects devoted to the study of large corpora. While this fact does not necessarily imply that critics' existing concepts of periodization are wrong, the ways in which many critics model periodizing change could probably use revision in their light. In the course of reading a few books, it is both easy and (often) desirable to emphasize the myriad differences between them. A model of periodization that scales up differences between books to differences between eras will, almost inevitably, translate the genuinely sharp distinctions between individual texts to the much larger classes of which those texts are presumptively representative members. But representation and typicality, on the evidence presented here—and, surely, in the light of critical experience—do not work in such direct terms, especially when the distinctions at issue are diachronic and continuous rather than synchronic and cross-sectional. Modernist studies perhaps finds

models of discontinuity even more attractive than do other areas of literary studies; the avant-garde movements that form one of the field's core objects certainly did. Modernists would be well served to balance that tendency with a robust appreciation of the roles played by continuity and incremental historical change within the literatures and cultures they seek to decode.

Concerning the foreign corpus that has been the object of our particular attention, we find support for the view that foreign writers were more likely to act as social outsiders, distributing their attention—within a larger framework of continuity—to locations that were more lightly engaged by native British authors. Foreign writers were in some cases less likely to adopt the spatial conventions of genre fiction. This may be a matter of privilege, or lack thereof; detective stories may not be especially high prestige, but the market for them is not equally open to all writers. At the same time, there were likely selection effects at play. The foreign writers who lived and worked in Britain before the Second World War were disproportionately students and intellectuals, often drawn from the colonial elite, and frequently returned to their home countries after a period of some years in the UK. This meant, in turn, that their imagined audiences were not necessarily exclusively native-born white British readers, a fact that likewise shaped the market for their work. Further, the foreign writers' greater interest in locations of government and iconic sites of British history, along with their greater geographic intensiveness overall, suggests that these writers, and especially colonial writers of color, may have employed topography to express their knowledge of and, by extension, their power within, the imperial metropolis and to place themselves on the map at "the heart of empire."[63]

Paul Gilroy emphasized almost twenty years ago the urgent, continuing need for attention to the geography of colonial-era London, arguing that "before we can be plausibly post-anything, we have to comprehend the colonial character of this city in a more profound manner than has happened before. Secondly, we have to produce histories of the city … which allow the presence of diverse colonial peoples and their stubbornly non-colonial descendants far greater significance than they have been allowed in the past."[64] The numbers within our Foreign corpus are relatively small, just 88 volumes by nonwhite authors and 42 by whites. This fact forecloses the more detailed cross-ethnic comparison that is obviously desirable with respect to the otherwise flattening category "nonwhite." But even "just" 88 volumes by nonwhite writers in Britain published before 1940 is an order of magnitude or more than have been treated in all but a few previous studies of

---

[63] See Elizabeth F. Evans, *Threshold Modernism*, (esp. ch. 5) for an extended exploration of this dynamic in a handful of works by colonial writers of color.

[64] Paul Gilroy, "A London Sumting Dis…" *Critical Quarterly* 41.3 (1999): 57-69, 60.

the field.

To the extent that our findings accord with the existing criticism—as when we observe, on the part of racialized foreign writers, increased geographic diversity at the international and intra-London levels, but a decreased range of other locations within Britain—we provide new detail to the longstanding project of recovery and expansion in British literary studies. Where our results differ—on the relatively close large-scale alignment of ethnically heterogeneous foreign writers in the period, for instance, or the declining geographic intensity and specificity of texts published after the First World War—we hope to have offered new pieces of evidence toward a more robust understanding of the forces that shaped ethnically and politically diverse literature in the years between 1880 and 1940. The work that follows from our results on this front is as likely to be small-scale and critical as it is large-scale and quantitative. That there are opportunities for such blended research within modernist studies, informed and shaped by quantitative literary geography, is the last of our interventions and the guiding principle of our own future research.