Familial Places in Jim Crow Spaces: Kinship, Demography, and the Color Line in William Faulkner's Yoknapatawpha County

Johannes Burgers

ARTICLE INFO

ABSTRACT

Hosted out of the University of Virginia, *Digital Yoknapatawpha* is an international collaboration between scholars of William Faulkner and technologists at the Institute for Advanced Technology in the Humanities. The project team has encoded all the locations, characters, and events in Faulkner's Yoknapatawpha fictions into a relational database that powers an open-access web-portal. Users can avail of an atlas of "deep-maps," data visualizations, archival material, and aural and visual resources to explore, teach, and research his works. Using techniques common in ecology and demography, this paper leverages the data to investigate the relationship between race, kinship, and space. It concludes, tentatively, that the social world of Yoknapatawpha is far more rigidly bounded along racial lines than current scholarship suggests. In particular, most interactions between characters from different races happen in a familial context, and are the result of racialized labor exploitation or outright enslavement of African-American families by Anglo-Americans. The lack of interactions outside of this context underscores just how little agency non-white characters have in Faulkner's fiction.

In the summer of 2019, the *Digital Yoknapatawpha* (*DY*) project quietly passed a resounding milestone: after eight years of laborious coding, all of the Yoknapatawpha fictions had been entered into the database. [1] The data represent a corpus of all of the locations, characters, and events from a significant share of William Faulkner's oeuvre.[2] While this data is continuously being updated and refined, it is robust enough to start probing commonly held critical assumptions about Faulkner's work.[3] A key area of investigation is the social world of Yoknapatawpha. The database makes it possible to study interactions between characters within individual texts, and across the corpus. In some respects, the data confirms much of what has already been written about the racial, class, and gender tensions between the different social groups in Faulkner's work. This is

unsurprising. Faulkner has benefitted from some of the most sustained and perspicacious literary scholarship of any American author. That being said, while *DY* is able to demonstrate the value of this scholarship in a new light, it also reveals novel insights into character relationships that would otherwise remain hidden. In particular, ecological and demographic models suggest that, on the whole, Faulkner's social world was more rigidly bounded by race, class, and gender differences than previous scholarship has documented. Notably, interactions between characters from different backgrounds occur almost exclusively within the context of economic relationships between wealthy white families and the African-American families they enslaved or employed.

At scale, the image that emerges of Yoknapatawpha's social world is therefore a deeply conservative one. True, some of Faulkner's fiction questions the boundaries of the white supremacist power structure, but the data suggest that a significant share of his work reproduces it too. To what extent Faulkner's reproduction of the color line is a vehicle for demonstrating its constructedness, and to what extent this is a projection of his own social world, is not easy to parse. Nor is it easy to square Faulkner's version of the South with other accounts of the lived experience of enslavement, and, subsequently, segregation in a small Mississippi town during this period. Not exclusively of interest to Faulkner scholars, these interactions across the color line also raise critical questions about race relations in the Jim Crow South more generally. The relative paucity of interactions between different races outside of the spatial intimacy created through racialized labor practices, speaks to the complex network of formal and informal social codes necessary to maintain segregation.

Within the Digital Humanities there is an extensive and growing body of scholarship on identifying character entities and their relationships in fiction. The bulk of this work has been done through supervised and unsupervised Natural Language Processing, character network analysis, or a combination thereof. These broader methodologies have utilized specific techniques including, but not limited to, inferring latent character types, sentiment analysis, and conversational and co-occurrence network visualization.[4] These forms of character analysis can be incredibly powerful, because they operate at a very large scale.

This paper contributes to this scholarship from quite a different perspective. Operating at a much smaller scale, it was possible to identify all of the characters manually, and give them attributes consistent with biographical and demographic data; something that NLP methods still struggle to do with literary fiction.[5] This character data can be situated in space in time by connecting it to location and event data. Using techniques from the biological and social sciences, it is possible to create corpus-wide, dynamic demographic views into how characters from different backgrounds interact with one another in Yoknapatawpha. While these techniques provide an incredibly rich view into character interactions, the critical issue is that fictional characters, of course, are neither like the animals biologists study nor the people described by the social sciences; they have an ontology that is uniquely their own. As Andrew Piper explains, characters are not necessarily "singularly deep mirrors of ourselves," but rather the practice of characterization hews more closely to its "etymological origins as a sense of being representative, general, or 'characteristic.'"[6] The extent to which these biological and social science models can articulate meaningful explanations of the interactions between characters in fictional texts is therefore by no means a settled matter.

Literature Review: Faulkner and the Color Line

In her important work on Faulkner's treatment of race in his later novels, Theresa Towner remarks, that "race" and "art" become "functions of one another."[7] Reading Faulkner and reading race are, quite often, the same thing, and work in this area has a long and rich genealogy with landmark studies such as Eric Sundquist's *Faulkner: The House Divided*, Philip Weinstein's *What Else But Love?*, and Thadious Davis's *Games of Property*. It is undoubtedly one of the central themes of his work, and scholars have been wide-ranging in their approaches and theoretical perspectives, to wit: John Duvall has shown how Faulkner uses performative "figurations of blackness," to probe "white masculine identity"; Michael Wainwright has unearthed the connections between Faulkner's writing and Darwinian thought; in a quite different fashion, Dorothy Stringer explores the uneasy associations between representations of blackness, trauma, racial violence, and the phallus with respect to psychoanalysis; taking a more historical perspective, Leigh Ann Duck highlights Faulkner's rejection and internalization of Jim Crow's ideological framework; Keith Cartwright highlights his indebtedness to African cultural forms; and, most recently,

Kristin Fujie points to the aural and visual connections between his fiction and African-Americans on film.[8] These works represent just a sliver of the available scholarship. There is no way to do justice to the nuances and differences between them, but by dint of their methodological frameworks, they all marshal close-readings to demonstrate how Faulkner engages the larger structural issues of race by unpacking his epistemology of racialization within select scenes from his extensive corpus.

By and large, these works recast Faulkner's representations of race in a more revisionary light. Whereas Faulkner's own life was marked by conservatism and racial views that were consistent with that of white moderates in the segregated South,[9] there is a sizeable body of scholarship that contends that the most significant parts of Faulkner's work deploys strategic subversions of white supremacist ideology. Thus, while Davis admits that she initially found Faulkner's representations of African-Americans "reactionary" and "retrograde," subsequent readings revealed to her that *Go Down, Moses* is "still problematic," but through an extended series of close readings of the mixed ancestry, enslaved character, Tomey's Turl, she demonstrates that he functions to "transgress and disrupt the power and authority of whites over both lives and stories."[10] John Duvall's analysis of *Mosquitos*, *The Sound and the Fury*, *Sanctuary*, and *Go Down*, *Moses*, likewise recognizes that while Faulkner, the public persona, could be racist, his "fiction is populated by white characters that are not exactly white," which is "a testament to an artistic imagination that recognized the contingency of racial identity."[11] A more radical version of this premise is one articulated by Christina Thyssen. Through a reading that relies largely on two select stories from *Go Down, Moses*, she argues that in "Pantaloon in Black," Faulkner arrives at a form of "literary ghosting where words and events arrive from a space of both pre- and postcategorical indefinition that blasts open the plantational text as the emblematic figure of historicism and the suicidal thought models of Western traditions."[12] Each of these accounts represents a trend in Faulkner studies to acknowledge his problematic pronouncements on race, while at the same time offering a more generous reading of race relations in a select number of his texts.

What is left fragmentary, implied, or otherwise unclear, is the picture that emerges when all of the scenes are tied together and viewed in the aggregate. This is not due

to any scholarly oversight, but without the robust dataset available in the *DY* database, such analysis is simply not possible. To be sure, some of the above scholars have been invaluable contributors to *DY*, and this paper is therefore both complementary to and complimentary of their foundational work.

Having said that, the aggregate view of Yoknapatawpha is far whiter than even Faulkner might have imagined. The demographic makeup of Lafayette, during this period consisted of an African-American population of around 35-40%, and, in any event, never exceeded the number of white residents.[13] Faulkner actually doubles this percentage on his map for *Absalom, Absalom!* in 1936, and claims African-Americans are 68% of the county's population.[14] Yet, within the corpus, characters with African-American ancestry only represent a little under 20% of the characters present at events. Curiously, African-Americans are over-represented in Faulkner's own demographics, but far less visible in his text. When they are visible, it is usually by virtue of a white character being present. It is very rare that characters with African-American ancestry are present in an event without white characters there as well (3.5%).[15] In this respect, the scenes that Thyssen highlights for her close reading constitute .04% of the entire corpus, and are far from representative of Faulkner's work in its entirety. Arguably, their singularity makes the scenes so notable and rich, but it is much more specious to make larger claims about his views on race based on these very exceptional instances alone.

The real power of the *DY* database, though, is that it allows for complex relational queries that show which characters are where and under what conditions. These queries, once again, cannot understand the nuances of race in specific situations, but they do reveal a generalized pattern of what types of character combinations are more or less common. Establishing a pattern of regular character combinations, in turn, signals the importance of those instances in the corpus where there is a deviation from the pattern, prompting textual exegesis. To name but a few notable exceptions to the regular order of things in Faulkner's world: Quentin Compson meeting an African-American outside of Yoknapatawpha in *The Sound and the Fury*, a white woman being alone in a house with an African-American man in "A Rose for Emily", or the fact that the only time the Chickasaw are ever present in the town of Jefferson is when they give up their land deed in *Requiem for a Nun* and attend Thomas Sutpen's wedding in *Absalom, Absalom!*.[16] Reading at scale therefore

does not obviate previous close readings, but instead underscores a similar point made by Piper that with regard to characterization, "small cues carry a large effect."[17]

Data Operationalization: Characters as People

The *DY* database consists of three related tables: locations, characters, and events. Locations are those spatial entities that either exist as a physical place in the text or are mentioned, and characters are those human entities that are either present or mentioned at locations. An event is a textual unit that demarcates one continuous action by one or more characters at one location for a discrete period of time. The three tables are keyed to each other in a relational database that allows the data to be viewed from the perspective of locations, characters, events, a combination of the three, or all three simultaneously, as with the main interface.

The process of data entry raises a whole host of practical and theoretical questions that, though interesting in their own right, are not germane here.[18] Suffice it to say that any data extracted from the database is measuring very specific things in a very specific way. A full overview of the data model and the encoding parameters is available through the instructions to the editors on the data creation portal.[19] Nevertheless, there are a few qualifications to the data that are pertinent to the subsequent analysis.

The data does not represent the "totality" of Faulkner's work. Primarily, this data is only for all the texts that either take place in or feature characters from Yoknapatawpha. Novels such as *Soldier's Pay*, *Mosquitos*, and *The Wild Palms* are not part of the data set. A further constraint is that the data has been encoded using specific editions of Faulkner's texts, although variance between editions has been noted, at the moment there is no way to represent the variances. More philosophically, totality implies that Faulkner had a grand vision, when, in fact, it is clear that he returned to Jefferson and Yoknapatawpha over and over again throughout his career in continuously different ways. From a computational perspective, totality is also a problematic term. There are different ways to aggregate and represent the data, but to do so in a meaningful way requires subsetting and highlighting specific attributes. With regard to specific data sets, this study

operationalizes a subset of attributes from the character, location, and event tables. Data creation for these three tables was highly interpretive and, going forward, it is useful to qualify what this data means.

Characters – Editors make several critical assumptions about the way Faulkner conceived of race and class. Written from 1929-1962, most of the Yoknapatawpha stories reflect how Faulkner conceptualized, articulated, and subverted notions of "race" before the Civil Rights Movement. In doing so, Faulkner, at times, adopted the language and ideological precepts of the Jim Crow South, while at other times called them into question. As much as possible, *DY* tries to reproduce his racial views with the understanding that these have since been rejected and are offensive to modern sensibilities. A central part of this methodology is the assumption that every interaction in Yoknapatawpha has a racial component, and when the race of a character is not stated in the text explicitly, the editors draw on previous scholarship and contextual clues to establish race. One guiding principle is that Faulkner tends to make the race of his non-white characters explicit, and leaves white characters racially unmarked. In Faulkner, whiteness is frequently the absence of race.[20]

Although class is a fluid concept, Faulkner's is a highly hierarchical society, and generally a character's source of economic sufficiency can be used to classify him or her as upper class, middle class, lower class, or "poor white." While this method is certainly not ironclad, usually white characters can be placed on the social ladder with a high degree of confidence. Meanwhile, Native Americans do not fit into this class system, and their status is based on their position in the tribe. Meanwhile, African-Americans are severely limited in their economic and social mobility due to the South's racial apartheid regime. They might have class status within their own community, but in *DY*'s interpretation of Faulkner's sociology African-Americans are either enslaved or free.

Perhaps the most fundamental hurdle to entering "all" the characters in Faulkner's world are uncountable collections of people. There are scale differences between "suitors,"[21] "a committee of Baptists,"[22] "gang,"[23] and Jefferson townspeople,[24] but there is no way to consistently distinguish what the orders of magnitude are. They have been given the qualifier "group," but this serves a descriptive function and cannot be quantified. Events that include groups constitute roughly 20% of the

corpus, and are therefore not insubstantial. Still, since groups tend to produce a large number of unknowns, they have been left out of this analysis. In the future, it would be interesting to test if there is a difference between character interactions when groups are present.

Locations – The most notable issue with locations is their positional variability. Stephen Railton and Christopher Rieger demonstrate that Faulkner was consistent with the distance ratios of his county, but locations can vary quite a bit, even within texts.[25] This, coupled with the inherent problem of locating places through textual clues,[26] has led to locations shifting from text to text. As such, there is no "total" map of Yoknapatawpha, only a series of individual maps for each story.

Events – Events measure when a character or characters occur at a location, and store the chronological order, order in the text, and narrative mode. Despite the richness of this information, as of yet no way to consistently establish how long an event is in terms of text time or word count, its duration in narrative time, its saliency, or what specific roles characters play. The practical result is that if two characters occur in one place *n* times, it may be that one character is there for a lengthier period of time and plays a more important role. Therefore, iterating characters across locations and events results in an unweighted frequency model, where each individual count is not necessarily equal.

Birds of a Feather: Co-Occurrence Analysis and Demography

The caveats above should not detract from the fact that there is an incredible amount of data available that allows users to explore Yoknapatawpha from a considerable variety of perspectives. One established method for investigating character interactions in fiction is through social network analysis. *Digital Yoknapatawpha* makes this option available to users as well. Team member Rafael Alvarado has created a series of visualization tools that allow users to investigate character interactions through character-character force directed graphs, and character-location co-occurrence through both force directed and bipartite graphs. Both types of graphs map connections between characters or between a character and a location weighted by the frequency of the occurrence.[27] There are several limitations to such graphs for the purposes of understanding social interactions. First, as Markus John

et al. point out, generally such social network diagrams tend to be static summations of an entire text, whereas the social connections make more sense when they are related to different moments in the text.[28] Second, such diagrams can show character centrality or group homophily, but it is harder to detect and quantify the patterns of character interactions based on an intersection of their demographic attributes. Relatedly, co-occurrence networks create edges between nodes whenever two characters occur in the same event. This correspondence makes sense for pairs, but this has the risk of overdetermining relationships when multiple characters are present in the same space. It may be that a character's presence in an event is incidental and not part of a pattern.

To understand the social world of a text requires knowing what type of characters interact with one another, how often, and under what conditions. One powerful technique for understanding population patterns is co-occurrence analysis. In ecology, co-occurrence analysis is used to understand the "rules of assembly" in an ecosystem by noting the presence and absence of species at particular sites at particular times.[29] Inferential statistics are then used to determine if a co-occurrence is:

1) **Positive:** If species is present x then species y is present.
2) **Negative:** If species x is present then species y is absent.
3) **Random:** Species x and species y occur together in a pattern that is neither positive nor negative relative to their share of the total population.

There are several ways of doing this, and a particularly accurate version is the combinatorics method developed by Joseph Veech. This calculates the ratio between all the possible combinations of observed co-occurrences, and the total number of possible co-occurrences.[30] While the computation is involved, Veech and a group of researchers have turned the algorithm into a library for the statistical language R for wide-spread use.[31] The advantage of this method is that it circumvents the various weighting issues with events. Instead, it establishes the frequency of co-occurrence between two types of characters and whether that co-occurrence is positive, negative, or random. These relationships – positive, negative, and random – should be understood as purely statistical. For instance, a pair of lovers and a slaver and an enslaved person can both co-occur positively, but for significantly different reasons. In a similar sense, as in nature, nothing in Faulkner is ever truly random. All species

within one ecosystem have a probability of co-occurring together by virtue of their area of habitation.[32] Just the same, Faulkner's characters can co-occur because they are in the same fictional world. The crucial distinction is that some co-occurrences are more or less stochastic than others.

Co-occurrence analysis establishes the contours of relationship patterns in Yoknapatawpha, and can also be used to determine what attributes affect social interactions. In Faulkner's version of the American South, one of the key features of social organization was kinship. Though the population of Northern Mississippi has always been transitory, communities like Oxford were established, in no small part, through familial chain migration.[33] In these early antebellum years, Wainwright explains that, "[s]ociety had meant accumulated kinship."[34] Power in these communities was usually arrogated by the prominent slave-holding families and the paternalistic plantation societies they created, where white families lived side-by-side with African-American families for generations.[35] When these societies collapsed after the war, industrialization and modernization significantly increased labor mobility and migration into the South, enhancing the potential for exogamy and rendering the erstwhile rigid social boundaries more brittle.[36] Situated within these social forces, Faulkner's characters use family and kinship ties as a bulwark against the erosion of social barriers, even if notions of familial purity were based on fantasies of unadulterated and un-adulterous patrilineal descent.[37] Thus, to understand possible transgressions of these boundaries requires looking at family interactions.

The families in *DY* include:

1) Benbow
2) Compson
3) Gibson
4) Issetibbeha | Ikkemotubbe
5) McCallum | MacCallum
6) McCaslin
7) Sartoris
8) Snopes
9) Stevens
10) Strother
11) Sutpen

12) Bundren
13) De Spain[38]

Running the co-occur analysis reveals that the thirteen families can be paired in 78 different ways and that there is a positive co-occurrence between 12 of them (see Table 1).[39]

Family 1	Ancestry	Family 2	Ancestry	Co-Occurrence
Sartoris	Anglo-American	Strother	African-American/Anglo-American	501[*]
Compson	Anglo-American	Gibson	African-American	165[*]
Issetibbeha	Native American	McCaslin	Anglo-American/African-American	91[*]
Compson	Anglo-American	Sutpen	Anglo-American/Anglo-American	81[*]
Benbow	Anglo-American	Sartoris	Anglo-American/Anglo-American	77[*]
De Spain	Anglo-American	McCaslin	Anglo-American/African-American	71[*]

Table 1. Co-Occurrence of Families

Note. There is no way to do co-occurrence analysis in Faulkner without running into the issue of double counting. Faulkner worked short stories into novels throughout his career. For example, in the data above, the set of *Saturday Evening Post* stories about the Civil War that will eventually be used for the novel *The Unvanquished* are all counted, meaning that the Strother and Sartoris families are counted for both the short stories and their ultimate novelization. In some cases, these stories vary very little, in others there are significant revisions. Distinguishing whether a text is a reinvention or a duplication raises serious issues of textual genetics. Suffice to say, that all of the texts included were different enough from each other to merit encoding into the database.

[*]*p* < .05 and top 50th percentile

What stands out are the top three families. The first two are relationships between families in which one upper class white family either enslaved an African-American family (Sartoris-Strother) or is exploiting them for domestic labor (Compson-Gibson). The third family relationship of significance is that of Issetibbeha and McCaslin, which, once again, points to familial connection between two families from different social backgrounds. It should be noted that the Issetibbeha-McCaslin co-occurrences are mostly from *Go Down, Moses* (83%), and, largely the result of the relationship between Sam Fathers – who is part Native American, African-American, and Anglo-American – and Ike McCaslin – who is white upper class. These top three co-occurrences constitute 60% of all positive family co-occurrences,

and suggest that family connections make interactions between different social groups possible in Yoknapatawpha.

A more fine-grained way of investigating character interactions is by looking at the intersection of attributes. *DY* has a comprehensive set of controlled vocabularies for race, class, and gender, and by concatenating them characters can be turned into "species." Because some character attributes were rarely used, they were consolidated into adjacent categories,[40] resulting in the controlled vocabularies in Table 2. The terms below can be combined in 50 different ways, 37 of the combinations occur in the fiction.[41]

Race	Class	Gender
Black	Upper Class	Female
Indeterminable	Middle Class	Male
Indian	Lower Class	
MixedBlackIndian	Poor White	
MixedBlackWhite	Indeterminable	
MixedIndianWhite	Indian Tribal Member	
MixedIndianWhiteBlack	Enslaved Black	
White	Free Black	

Table 2. Controlled vocabularies used for creating character types.

To test the effect of family on the diversity of co-occurrence, the concept of familial relationships requires some refining. There are multiple ways to conceive of how families might interact with one another:

1) **Intra-familial:** within families
2) **Inter-familial:** between families
3) **Extra-familial:** between a member of the family and a character not identified with a family
4) **General population:** co-occurrence in which there is no interaction with a major family.

In the figures below, Family Present has been defined as all intra- and inter-familial events, which are those where at least two characters are from different families or from the same family. Family Absent includes any extra-familial and general

population characters, and indicates those events where there is at most one character present from a major family. These charts can be visually overwhelming, but the general pattern is that there are more positive co-occurrences of different types of characters with family (48) than without family (22).
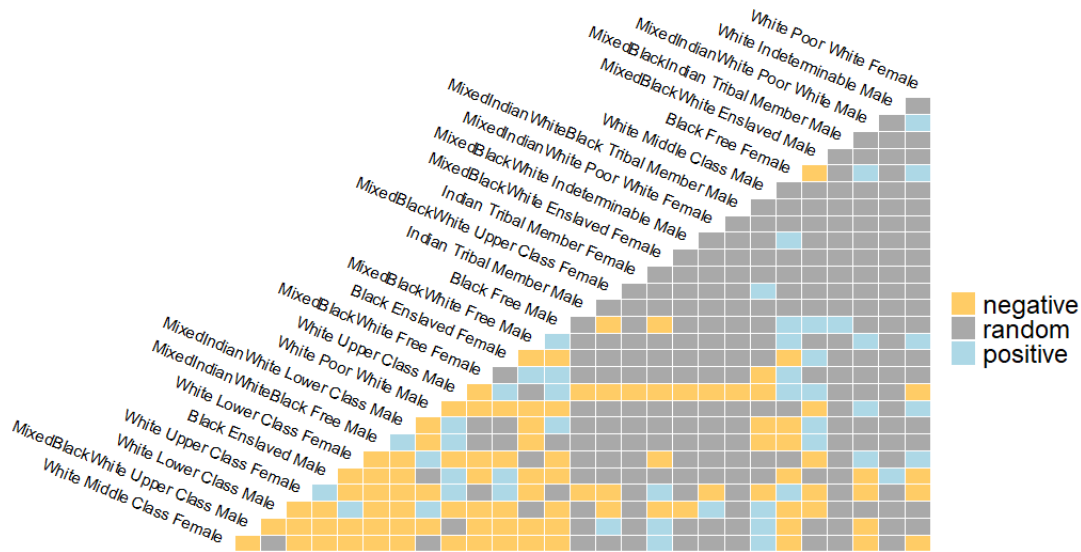

Figure 1. Family Present Co-Occurrences. Analysis when there were intra- and inter-familial co-occurrences resulted in 256 pairs. 48 were positive, 100 were negative, and 108 were random.
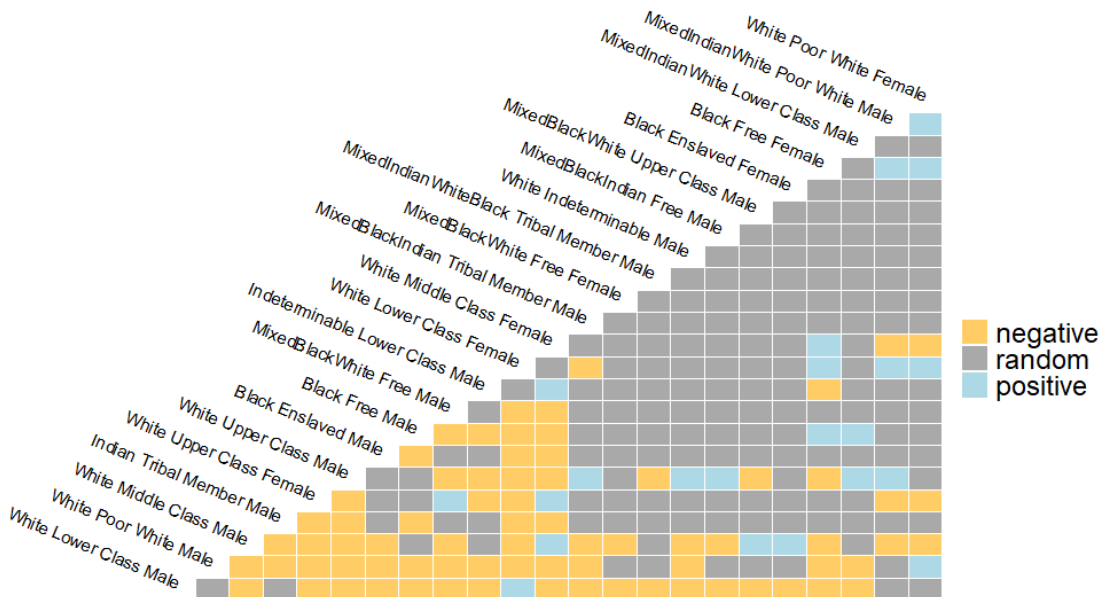

Figure 2. Family Absent Co-Occurrences. Analysis when there were extra-familial co-occurrences and co-occurrences among the general population resulted in 165 pairs. 22 were positive, 74 were negative, and 69 were random.

Breaking down the individual co-occurrence pairs reveals that the majority of relationships in the corpus are those between white upper-class families and African-Americans who are either enslaved or exploited for domestic labor. Table 3 merges the 90th percentile of both the Family Absent and the Family Present probabilities tables.

Pair Change	Character Type 1	Character Type 2	Family Absent	Family Present	%Change
Added	White Upper Class Female	White Upper Class Male	85	781*	818%
Added	Black Free Male	White Upper Class Male	103	359*	249%
Added	Black Enslaved Male	White Upper Class Male	10	304*	2,940%
Added	Black Free Female	White Upper Class Male	29	215*	641%
Added	Black Enslaved Male	White Upper Class Female	1	185*	18,400%
Unchanged	White Lower Class Female	White Lower Class Male	318*	143	-55%
Unchanged	White Middle Class Female	White Middle Class Male	129*	46	-64%
Dropped	MxdIndianWhite Poor White Male	White Upper Class Male	53*	63	18%

Table 3. Family Absent and Family Present Co-Occurrence Changes Positive Co-Occurrences

Note. The combinatorics algorithm was run on two different data sets: Family Present (2666 events) and Family Absent (3677 events). By separating out the two data sets, certain pairs occur either more or fewer times. This increase or decrease causes changes in the probability of the co-occurrence being positive, negative, or random. Hence, the added rows were not positive pairings in the Family Absent data set. Meanwhile, some pairings remained positive even if the total amount of co-occurrences varied between the Family Present and the Family Absent data sets. Only one was dropped – the positive co-occurrences between the MixedIndianWhite Poor White Male and the Upper Class White Male. This is the relationship between Boon Hoggenbeck and Lucius Priest in *The Reivers*. Boon co-occurs positively when only Lucius is present, but not when Lucius is with his family.

* $p < .05$ and 90th percentile

The contrast between the two data sets is marked. The added pairs are almost exclusively indicative of enslavement or exploitative domestic service. White upper-class characters from major families co-occur with enslaved or free African-Americans over seven times more often (1,063) than when there is no familial connection between African-American and Anglo-American (143). Generally,

interactions between Anglo-American and African-American characters outside of a labor relationship between families are exceedingly rare. For example, with regard to enslaved African-Americans, only *once* in the entire data subset of nearly four thousand events does it happen that a white upper-class woman not from a major family co-occurs with a black enslaved male. Significantly, this is in *Light in August* when Gail Hightower's mother meets her father-in-law, and, he, knowing his son's abolitionist politics, takes his two enslaved servants with him.[42] Even more striking, African-American men and women in any context, co-occur with each other ten times fewer (110) than they do with their Anglo-American enslavers or employers. More to the point, the narrative spends more time on African-Americans in relation to Anglo-Americans than it does on African-Americans in connection with each other. Whatever racial politics may play out on a local level within an event, globally the picture that emerges is one in which African-Americans rarely exist outside of their relationship with Anglo-American characters. These relationships are, by and large, the result of racialized labor exploitation, if not outright enslavement.

More generally, the data reveal several different social groups that are consonant with the broader sociology of the ante- and postbellum South, and reproduce the paternalistic system enforced by the plantocracy.[43] African-Americans were bonded to upper class Anglo-Americans by virtue of enslavement and, later, racialized labor practices. These relationships were intimately tied to family, and are consistent with nineteenth century practices of framing economic exploitation of African-Americans as a paternalistic relationship that normalizes conditions of white supremacy. With familial names like "Uncle" and "Aunt" for enslaved African-Americans and, later, servants, white families could maintain a façade of familial comity to mask deeply coercive and abusive labor practices.[44] Importantly, by the time Faulkner is writing, a whole family of African-American domestic workers living with their white employers, as is the case in *Flags in the Dust* and *The Sound and the Fury*, would have been anachronistic. Most domestic work was performed by African-American women whose family were seen as an unnecessary distraction to white employers, and, consequently, single African-American women were highly regarded.[45] In contradistinction to this, Faulkner is reproducing the plantation social order well into the twentieth century. To be sure, Faulkner complicates plantation relationships in works like *Absalom, Absalom!* and *Go Down, Moses* where kinship between families was not always imagined, but the result of illicit

relationships between Anglo-Americans and African-Americans. Still, even here there are echoes of paternalism.[46]

Meanwhile, the non-upper class white population stood outside of this dyad as a different social grouping that had more contact with one another than with wealthy Anglo-Americans and African-Americans. Indeed, Faulkner's separation of the social worlds of the Old South plantocracy and the masses of dispossessed whites of the New South, is something early critics like Malcolm Cowley already observed.[47] Yet, within this broad contrast there is quite some nuance. There is pronounced stratification between the lower class and middle class characters. The fourth group in Yoknapatawpha, Native Americans, do not co-occur with any character type positively in either a familial or a non-familial context, and are quite separate from Yoknapatawpha at large.[48]

Similar techniques also provide a more detailed look at how family affects racial interaction at the scale of the novel. Figure 3 shows the number of positive and negative cross-gender and cross-racial pairs divided into four different categories of familial relationships and bucketed by novel. For this chart, co-occurrence analysis was run on all of the novels individually and subsetted into intra-familial, inter-familial, extra-familial relationships, and the general population. To highlight the cross-racial relationships exclusively, this data model eliminates class as an attribute for character type creation, leading to 28 pair possibilities. Also eliminated were any mixed-gender co-occurrences between characters of the same race. All texts that did not produce a statistically significant co-occurrence pattern were left out.[49]
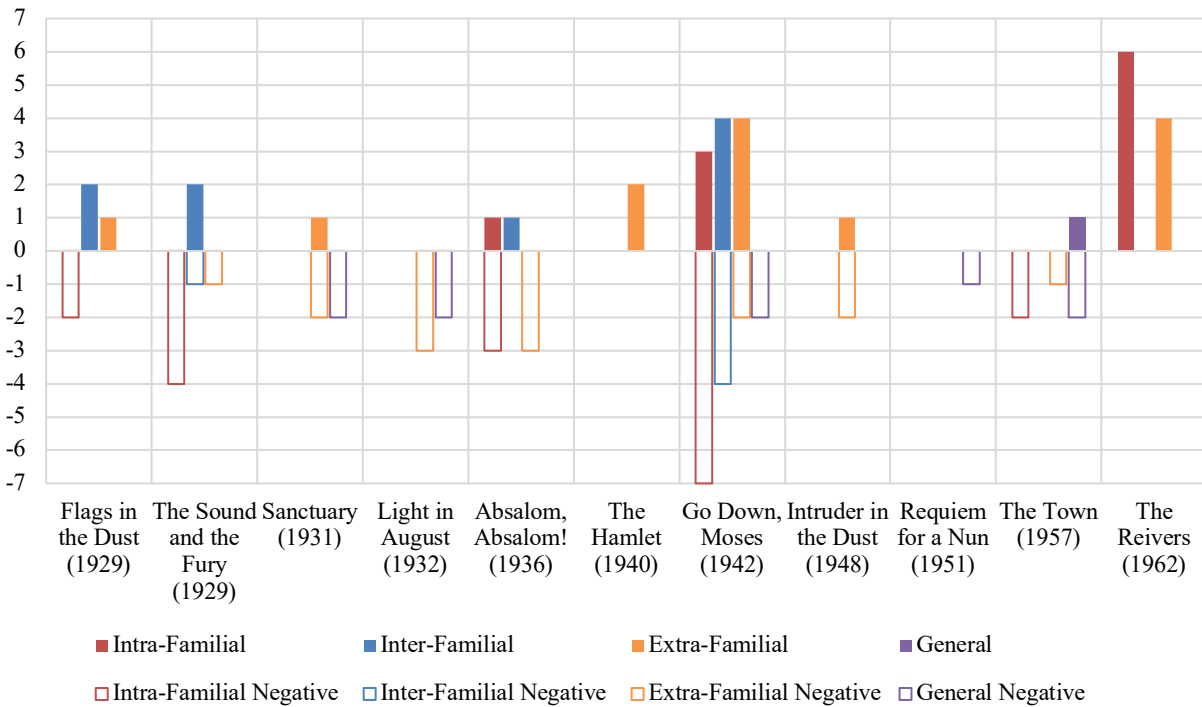
Figure 3. Race and Gender Co-occurrence by Novel. Each text was analyzed for intra-familial, inter-familial, extra-familial and general population co-occurrence by race and gender. All texts without a result were excluded.

Running from left to right, the series indicate a fluctuating pattern of cross-racial and cross-gender interactions. The novels that have the most diverse co-occurrences are *Go Down, Moses* and *The Reivers.* For *Go Down, Moses*, this is not entirely surprising, given its position in the Faulkner canon as one of the texts that most fully addresses the South's complex racial legacy. *The Reivers* is, perhaps, the more unanticipated result. A less canonical work, Faulkner's last novel is about a ragtag crew led in alternating fashion by Ned William McCaslin, Lucius Priest I, and Boon Hogganbeck. Ned is the grandson of an enslaved woman and the McCaslin dynasty patriarch, Lucius Quintin Carothers. Lucius Priest I is a young white scion of the McCaslin planter aristocracy, and great-great-grandson to the same Lucius Quintin Carothers. Finally, Boon is the Priest-family factotum with possible Native American ancestry. The three mismatched characters spend their time getting into various misadventures in and around Memphis, visiting brothels, race tracks, and aristocratic mansions, and cavorting with an appropriately motley cast of characters along the way.[50]

The concentration of diverse connections in only two novels paints quite a different picture of race across Faulkner's body of work than current scholarship suggests.

Undoubtedly, race is a thematic concern in other works, but co-occurrences between diverse characters are rare, and when they do occur, they are usually the result of family relationships. This very limited interaction pattern between different character types, speaks more to white anxieties about race than it does about the possible effects of racism on non-whites, who are largely excluded from the equation. In fact, in only one text positive cross-racial co-occurrences happen outside of a major family connection. In *The Town,* an African-American woman and Anglo-American woman appear in eight events together. The majority of these are the co-occurrences between the African-American wet nurse Old Het and Mannie Hait, a white woman, who are the principle characters in a vignette modelled after the short story "Mule in the Yard." Finally, the high number of negative intra-familial co-occurrences also bears some explaining. Because of the way the data is set up, all events in which at least two family members occur are measured. It might happen that a non-family member is present in these events. As these characters appear less frequently when either of the two family members are present, these are counted as negative co-occurrences. As insightful as co-occurrence analysis is, it runs into trouble when dealing with smaller data sets. Nor can it reveal anything about the modality of the co-occurrence. Mapping these co-occurrences onto their locations reveals the effect of space on racial interactions.

Experimental Demography: Jim Crow Spaces and Family

Mapping demography in the American South during the Jim Crow era is by no means a simple matter. In the North, housing segregation could be observed in residence patterns, with less-desirable locations being occupied by African-Americans. In the South, segregation was social, but not necessarily spatial. In place was a type of "backyard" segregation, where domestic servants would live on or near the property of their enslaver and later employer.[51] The spatial and demographic data required to represent the vast economic and social chasm between the alley and the street-front has to be incredibly detailed. Even then, it only captures part of the social dynamic, as "Negroes" during this period were by no means a static demographic category. Instead, some African-Americans could change their racial status based on their ability to pass, and changes in their economic circumstances.[52] Faulkner, to a certain extent, acknowledges both of these dynamics in his fiction, but the maps he

made and the population statistics he attributed to Yoknapatawpha are obviously nowhere near detailed enough for traditional demographic methods.

DY has made a lot of headway in supplying interpretive data to start establishing a demography of Yoknapatawpha, but there remain intractable problems with location position and consistency. Nevertheless, with some modifications, a feature density map or, "heatmap," creates a visual approximation of character distribution. Such a map breaks down the distribution of events by character attributes. Heatmaps effectively create a proxy for population distribution, as certain types of characters appear at certain locations more or fewer times. Any inconsistencies between maps will also be elided through this process because fixed places are ignored, and the frequency of different types of people at events becomes predominant. Since the events are time-dependent, this frequency map does not capture one singular moment in time, like a census, but instead the layering of all moments in time onto one space. It may be that some of the events happen well over two hundred years apart from each other. Through temporal accretion, the maps make it clear that certain kinds of people are more likely to occur in some areas as opposed to others.

To generate the map below, several interpretive moves were made (see Figure 4). First, some of the maps in *DY* have insets that either zoom into a location within Yoknapatawpha, such as Frenchmen's Bend in *The Hamlet,* or detail maps of places outside of Yoknapatawpha, such as Memphis in *Sanctuary.*[53] These insets had to be consolidated back into the main map by translating and reducing the coordinates to an area on the original map, or moving them into one point onto the "Region," "Nation," and "World" insets to the right of the Yoknapatawpha map. These re-mappings reduce the precision of the data, but for the purposes of showing event density, which operates at high frequencies over generalized areas, this is not consequential. Second, the density function parameters were set so that each heatmap layer was on the same scale. It also excluded any location where only one character occurred, which causes stippling across the map, and gives these locations more visual importance than they merit. To increase clarity, the number of races was reduced from eight to four by re-classifying anyone with uncertain or multiple ancestries as "mixed ancestry."

The map represents eight significant areas of activity. The plantations of Compson, McCaslin, Sutpen, and Sartoris surrounding the town, the old Chickasaw lands of Ikkemotubbe to the far north of town, the town of Jefferson proper, and to the far southeast of town, the agrarian hamlet of Frenchman's Bend. A notable ninth space of Yoknapatawpha is Memphis, which both conceptually, and in terms of plot, frequently functions as a repository of Jefferson's vices. Since the Memphis insets were re-mapped onto the regional map for this map, their data does not come to full expression here.
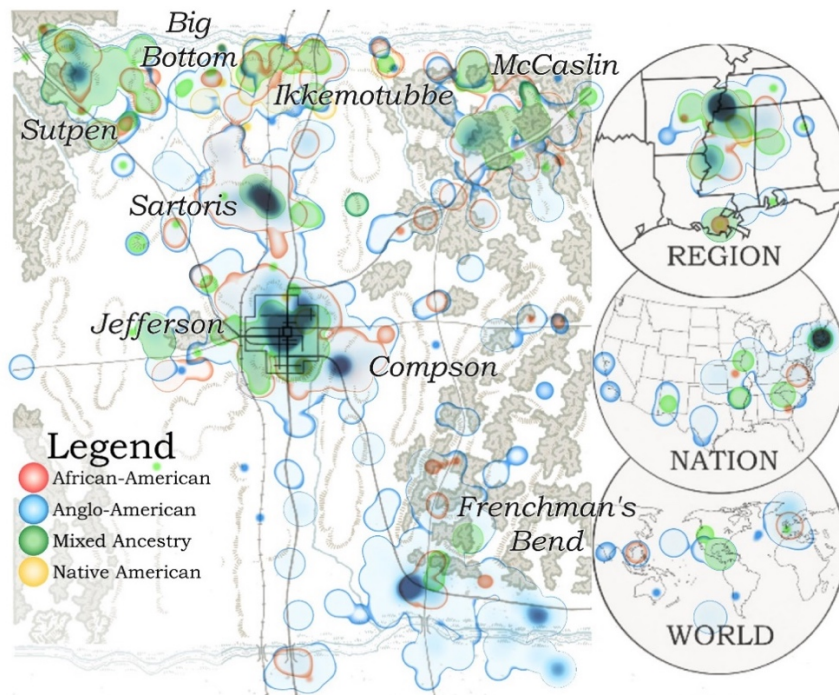

Figure 4. Characters at Events by Race. The following heatmap was created by superimposing four different layers. All characters who were not Anglo-American, African-American, or Native American, were consolidated into one "Mixed Ancestry" category. The map very clearly shows how Faulkner conceived of space as racialized.

The map reveals several key points. The first is that space in Yoknapatawpha is racialized. The events that take place in Frenchman's Bend to the southeast are dominated by Anglo-Americans. There are very few non-whites who enter this area, and as the opening of *The Hamlet* so forebodingly states, there is not "a negro landowner in the entire section" and "[s]trange negroes would absolutely refuse to pass through it after dark."[54] African-Americans stayed away for a good reason. Meanwhile, the most consistent overlap between African-American and Anglo-Americans occurs on the plantations. In some cases, the separation is stark, as with

the Compson family, there are only African-American and Anglo-American characters here. Moreover, on the Sartoris plantation there is one mixed ancestry character, Elnora Strother, whose father is the planter Colonel John Sartoris and whose mother was an enslaved woman, Euphrony Strother. In other cases, the segregation between the families is less marked, as evidenced by the McCaslin and the Sutpen plantations where mixed ancestry characters are far more present. These areas are also farthest removed from the homogenous area of Frenchman's Bend. In fact, the social lines between the races become far blurrier once the panoptic surveillance of the town of Jefferson is at a greater distance. This is not to suggest that the Sutpen and McCaslin plantations are sites of racial comity. Instead, the majority of mixed ancestry characters in these places were conceived through violent coercion. The only place on the map that is suggestive of racial harmony are the former lands of Sutpen's Hundred that are revisited and retraced during the hunting parties of *Go Down, Moses*. Here, men from different walks of life come together to practice the common ritual of the hunt. To all appearances, homosocial spaces reduced to a state of nature have a profound effect on Faulkner's treatment of racial relationships. While it establishes a bond between Sam Fathers and Ike McCaslin in *Go Down, Moses*, it lays bare Thomas Sutpen's subjugation of enslaved Africans in its most brutal terms. In a recurring scene in *Absalom, Absalom!* Sutpen is imagined "stark naked" caked in mud along with twenty Africans he enslaved.[55] The societal decorum that naturalizes one person's domination over another is not present here. Another remarkable feature of the map is just how isolated the Native Americans are. For the majority of the corpus they only appear in one small area to the north of town, and are slowly encroached upon, until there are more Anglo-American characters than Native American ones. Finally, it should be noted how devoid the wider world is of non-white characters, even in places where they are the majority population.

Animating Characters: Demographic Distribution across Time

By sequencing the narrative dates and chronological order of all the events, it is possible to construct an animation of where each type of character occurs across the narrative time of the corpus (See Video 1: "Racial Distribution by Event Chronology").[56] Animating the chronology of Yoknapatawpha confirms that the

racial distribution is the result of consistent patterning on Faulkner's part. While it would be inaccurate to say that Faulkner had a total vision, the animation shows that he was deliberate about who he placed where. The first events of Yoknapatawpha at the beginning of the nineteenth century, take place concurrently on the Chickasaw lands and at the McCaslin plantation. This is followed by the creation of Sutpen's Hundred in the northwest. Around the time of the Civil War, the Sartoris family comes into focus. It is not until the turn of the century that events occur at the Compson place, at the same time Frenchman's Bend sees more interactions. By the end of the nineteenth century, the rate of incursions on Native American land also starts to rise more rapidly. By the early twentieth century, many of the different racial boundaries have already been established in Yoknapatawpha. From here on, two important things change: America as an imagined space becomes larger, and the fantasy of racial purity in Yoknapatawpha starts to collapse. The green mixed ancestry shading becomes more pronounced in Jefferson, and at Sartoris and McCaslin throughout the twenties and thirties. Meanwhile, the areas noted on the national inset become more expansive, with the US stretching from New England to the Midwest, and occasional events in Texas and California. The world becomes more global too, with events in Europe, Australia, Asia, the Caribbean, and Mexico. Of course, any chronological telling of Faulkner's fiction is an idealized version of the timeline. It also elides the thematic differences from text to text.

Playing the timeline by publication date demonstrates Faulkner's creative process (see Video 2: "Racial Distribution by Publication Date").[57] The publication animation, working with the exact same data, arrives at the same result but through different patterns. Faulkner has most of his county established fairly early on, and even with just *Flags in the Dust,* the glimmers of the major areas of the fiction emerge. By *Absalom, Absalom!* all of the major areas of the map have seen at least some events. *Go Down, Moses* in 1942 represents a major inflection point in terms of the amount of mixed ancestry events across the country, underscoring its centrality to Faulkner's treatment of race. The later works revolve largely around Frenchman's Bend and the town of Jefferson.

These two animations are a baseline for a more granular analysis of race, family, and space. Of interest is the intersection of non-white members of major families into spaces that would otherwise be exclusively white. These spaces can be understood

as "white spaces" among the general population. For example, in *Flags in the Dust*, the African-American coachman for the Sartoris family, Simon Strother, drops off Jenny Du Pre at the telegraph office. There are no other non-white characters who appear here throughout the novel, and Simon is there because he is doing work for his employer.[58] Events where a non-white person enters an exclusively white space by virtue of a family relationship are statistically rare, and in *Flags in the Dust* this is the only time such an event takes place.

To render "white spaces among general population" all of the events involving any of the families were removed, then the locations were subsetted into those where only white characters appear across the entire text. As these remaining white characters are not part of a major family, they constitute the general population. Then the intersection of non-white characters from major families with these white spaces was calculated, and broken down by the gender. Each feature was scaled based on the total number of times a non-white character appears within the white spaces over the course of the narrative (see Figure 5).
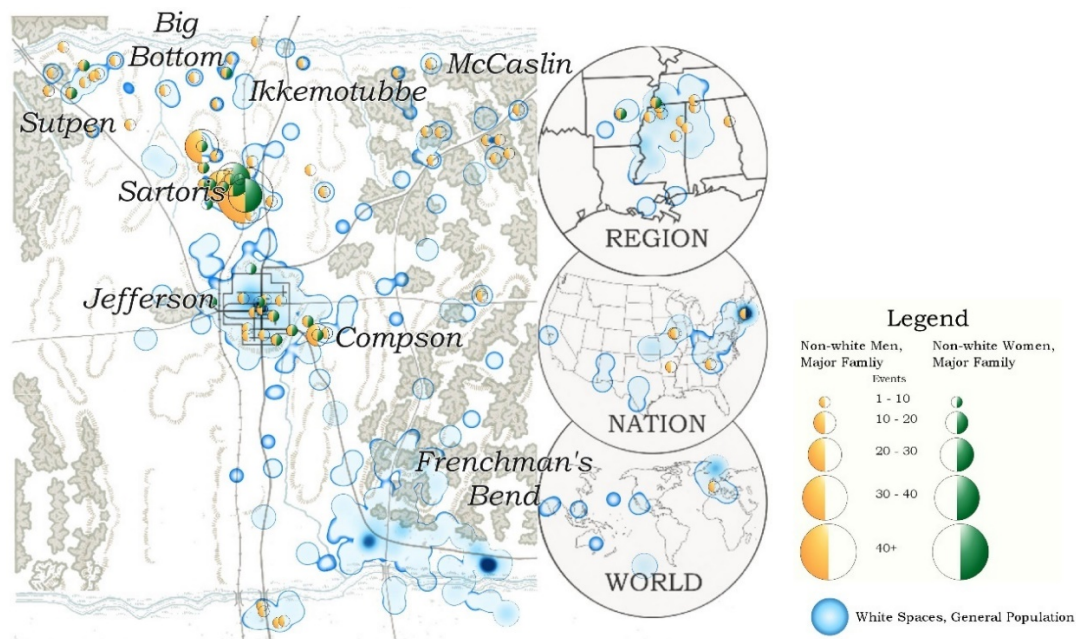

Figure 4. Non-white Characters from Major Families in White Spaces. The majority of non-white characters who travel outside of a domestic setting are men.

To provide a useful point of contrast, a layer with all the instances of non-white characters not from major families was also created (see Figure 6). This gives

some indication of where non-white characters go in the county, and also how frequently non-white characters without a family connection appear in the corpus.
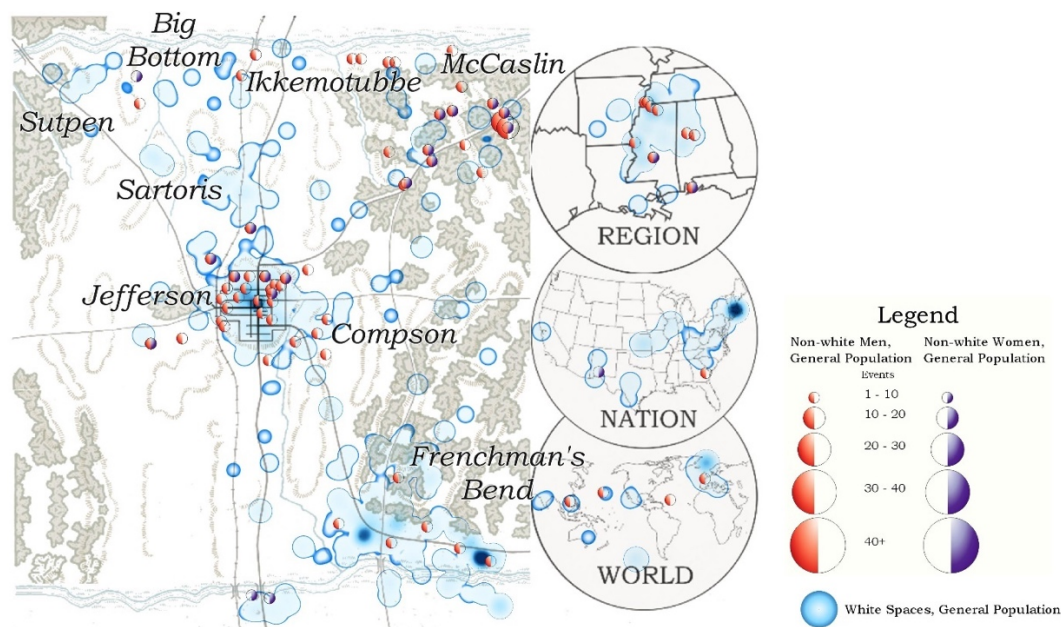

Figure 5. Non-white Characters without Family Connection among General Population. The number of times non-white characters not from major families enter "white spaces" is far less frequent.

Combining the layers and iterating it by publication date creates a compelling if complex animation (see Video 3: "Non-white Characters in White Spaces").[59] The first notable feature is the ubiquity of white spaces across the corpus. A full 52% percent of locations only ever feature white characters when characters connected to major families are removed from the data. To put this in slightly different terms, the majority of places that characters pass into and out of are exclusively white. Not only are these white spaces dominant spatially, they are also very numerous across Faulkner's career. Over 40% of the stories do not involve a non-white person if they are not connected to a family. Second, the movement of non-white characters from non-white families is very much constrained to places around the home of their enslaver or employer. Thus, while it might be true that certain African-American characters do have some level of agency, as Davis and Thyssen point out, this agency does not translate into any real mobility. In the case of women, there is only one instance of a non-white woman appearing in a "white space" outside of a domestic scenario: when Molly Worhsam Beauchamp visits the law office of Gavin Stevens to solicit help for her grandson in *Go Down, Moses*.[60] Among non-white men there

24

is more freedom of movement outside of the home, and even some on a national and global scale. Yet, for them too, entry into white space is highly qualified. The characters who are able to temporarily transcend racial boundaries are Ringo and Joby from *The Unvanquished*, Charles Bon in *Absalom, Absalom!* and Sam Fathers in *Go Down, Moses*. None of these characters could be mistaken for privileged. Ringo and Joby are both enslaved, and Ringo, an African-American child, is, implausibly, helping Bayard Sartoris fight the Union during the Civil War. Charles Bon has more freedom of movement because he is able to pass as white, but is murdered. Sam Fathers has white, African-American, and Native American roots, but is viewed as a "negro" in town, and therefore goes to live in the wilderness where white hunters respect his Native American ancestry. In short, even when non-white characters with a family connection appear to have marginally more access, this access is highly restricted. Among non-white characters not from a major family, space is even more limited, and events are far fewer. In cases where a character through his or her spatial positioning appears to have some form of agency, the narrative reality undercuts any possibility of independence. Rider, who moves around the northeast corner of the map in "Pantaloon in Black," is haunted by his wife's ghost, kills a man, and is lynched. Will Mayes from "Dry September" is also lynched. The African-American body servant to Issetibbeha who travels at the north of the map during "Red Leaves" is forced to sacrifice himself. In "Elly," Paul de Montigny is killed by his race obsessed lover, Elly, in a car crash. In fact, a full 12% of non-white men without a family connection meet a violent end.

At first blush, the general contours of these findings contradict the idea that Faulkner, a moderate white Southerner, was subverting Jim Crow era racial politics through his highly experimental aesthetics. In essence, the co-occurrence tables and animations present substantial evidence that a) Faulkner's world was highly segregated; b) non-white families who were enslaved by or servants for white families occur more often than the general population of non-whites; c) these major non-white families were slightly more mobile spatially because they were able to enter spaces that would otherwise be exclusively white; d) this effect was stronger for men. Temporarily, family connection could supersede racial segregation, even if it likely underscored the asymmetry of the relationship. While not surprising, the disparity between white and non-white is quite dramatic. Even when non-white characters are vehicles for underscoring the problem of the system of white

supremacy, in Faulkner's world they appear to have little existence outside of family connections, and when they do, they tend to serve as the victims of spectacular and voyeuristic violence.

Conclusions to New Beginnings

It is difficult to reconcile these insights with scholarship on race and Faulkner. Scholars generally accept that Faulkner's upbringing and the pressures of small-town Oxford limited the range of opinions he could publicly express about segregation, while also acknowledging that he had a spotty record of making reactionary comments about race when afforded the opportunity. Nevertheless, his writing, in its sheer imaginative richness, appears to treat race with nuance and complexity. When an aggregate view is taken, this subtlety, by necessity, falls away, and what is left is a world entirely constructed from the perspective of white society. There is little room for non-whites generally, and African-American characters specifically, to exist without being implicated in the South's racial problem. While Faulkner uses this perspective to show the inhumanity with which whites treat others, this, perversely, dehumanizes those very others by granting them no existence outside of exemplifying white supremacy's moral degeneracy. More bluntly, when Faulkner writes about race, it is more often about the anxieties of whiteness, and not the humanity of African-Americans.

This harsh conclusion is necessarily imperfect, and relies on reading the aggregated social world as more representative than a few select scenes. This raises significant philosophical issues about the function of minor characters in fiction. Echoing Alex Woloch's work on minor characters, one salient question is whether Faulkner's minor characters are part of the social world of Yoknapatawpha, or merely vehicles that facilitate the plot of the central character? [61] Relatedly, are characters who are only ever present in an event even part of the plot, or are they merely ornamental? In order to get a better understanding of how minor characters function in literature, it would be necessary to establish a comparative baseline with other corpora. Given the time investment it took to create *DY*, any similar database for another author is likely some time away.

Nevertheless, the methodologies deployed in this paper are highly suggestive, and find patterns between entities that are not possible to understand at a local level, and hard to capture using other techniques. Crucially, they assume that the organism they are observing, both human and non-human, are acting with some form of agency. The same cannot be said for characters in text. They do not share an ontology with the living world, and do not have agency in a similar sense. They are stuck between the unknowable intentions of the author, and the idiosyncratic interpretive strategies of the reader. The character patterns available through the *DY* database should, therefore, not be misconstrued as the collection of hopes, desires, and feelings of the people of Yoknapatawpha, but rather as the grains of sand that sediment Faulkner's native soil. Viewed from this perspective, character analysis at scale creates a special version of the sorites paradox. Each individual character action contributes to the building of an aggregate demography that is invisible to the builder. It has taken *DY* eight years to make this demography searchable and mappable. With phase II of the project, we will introduce keywords for events, which, hopefully, makes Faulkner's world just a little more knowable too.

Acknowledgments

Notes

[1] iath.faulkner.virginia.edu

[2] The data is available upon request. Please contact: Stephen Railton <sfr@virginia.edu> or Johannes Burgers <johannes.burgers@ashoka.edu.in> for more information.

[3] This essay was composed using a data grab from October, 2019. Since then, there have been minor updates to character, location, and event entries being used for analysis.

[4] For a good overview of recent work that has been done see: Niels Dekker, Tobias Kuhn, and Marieke van Erp, "Evaluating Named Entity Recognition Tools for Extracting Social Networks from Novels," *PeerJ Computer Science* 5 (2019). There have been various approaches to extracting character entities from literary fiction. A

particularly powerful one is that of Bamman, Underwood, and Smith who use a Bayesian mixed effects model to detect latent character types while also accounting for differences in authorial style, David Bamman, Ted Underwood, and Noah A. Smith. 2014. "A Bayesian Mixed Effects Model of Literary Character." (In *52nd Annual Meeting of the Association for Computational Linguistics, Baltimore, MA, June 2014*). Baltimore, MA: Association for Computational Linguistics. Meanwhile, John Markus et al., developed a tool that uses a combination of sentiment analysis and social network mapping with which users can explore character relationships across a narrative, Markus John et al. 2019. "A Visual Approach for the Comparative Analysis of Character Networks in Narrative Texts." (paper presented at the *IEEE Pacific Visualization Symposium, Bangkok, Thailand, 23-26 April 2019*), 10.1109/PacificVis.2019.00037. In a similar vein, Fischer et al., are working on a model to show the relationship between character networks and plot, Frank Fischer et al. 2017. "Network Dynamics, Plot Analysis: Approaching the Progressive Structuration of Literary Texts." (In *Digital Humanities 2017, Montréal, 8-11 August*). Montréal: McGill University. In the fifth chapter of his work, *Enumerations*, Andrew Piper deploys a multipronged character analysis approach to demonstrate, among other things, that female characters written by women in the early nineteenth century, have a higher tendency towards interiority and reflection, a feature of characterization that is appropriated by male authors writing about male characters in the twentieth century. Andrew Piper, *Enumerations: Data and Literary Study* (Chicago: The University of Chicago Press, 2018), 118-46. For work on conversational-networks see: David K. Elson and Kathleen R. McKeown. 2010. "Automatic Attribution of Quoted Speech in Literary Narrative." (In *Twenty-Fourth AAAI Conference on Artificial Intelligence, Atlanta, July 2010*). Association for the Advancement of Artificial Intelligence; Nathanael Chambers and Dan Jurafsky. 2008. "Unsupervised Learning of Narrative Event Chains." (In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics, Columbus, OH, 2008*). Association for Computational Linguistics; Hua He, Denilson Barbosa, and Grzegorz Kondrak. 2013. "Identification of Speakers in Novels." (In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, Sofia, Bulgaria, 2013*). Association for Computational Linguistics. Work on co-occurrence networks has been done by: Matt Fernandez, Michael Peterson, and Ben Ulmer, "Extracting Social Network from Literature to Predict Antagonist and Protagonist," *Technical Report* (2015); Mariona Coll Ardanuy and Caroline Sporleder. 2014. "Structure-based Clustering of Novels." (In *Proceedings of the 3rd Workshop on Computational Linguistics for Literature (CLFL), Gothenburg, Sweden*. Sweden Association for Computational Linguistics.

[5] There have been a number of studies related to the representation of character gender in fiction. Jonathan Cheng's recent contribution gives a good overview of what is possible, and also what remains elusive. In particular, current parsers cannot pick up on the gender of generic nouns. Jonathan Cheng, "Fleshing Out Models of Gender in English-Language Novels (1850 – 2000)," *Journal of Cultural Analytics* (2020). There have been attempts in other disciplines to gather demographic data through NLP techniques, but the unstructured data sets need to be complemented by either structured demographic data or explicitly defined communication networks. See for example: Jane R. Grafton et al., "Identifying Race/Ethnicity Data via Natural Language Processing among Women in a Uterine Fibroid Cohort Study," *Journal of Patient-Centered Research and Reviews* 3, no. 3 (2016); M. Saravanan, "Determining Ethnicity of Immigrants using Twitter Data," in *Proceedings of the 4th Multidisciplinary International Social Networks Conference* (Bangkok, Thailand: Association for Computing Machinery, 2017); Adi V. Gundlapalli et al., "Combining Natural Language Processing of Electronic Medical Notes with Administrative Data to Determine Racial/Ethnic Differences in the Disclosure and Documentation of Military Sexual Trauma in Veterans," *Medical Care* 57 (2019).

[6] Piper, *Enumerations: Data and Literary Study*, 121.

[7] Theresa M. Towner, *Faulkner on the Color Line: The Later Novels* (Jackson: University Press of Mississippi, 2000), 9.

[8] John N. Duvall, *Race and White Identity in Southern Fiction: From Faulkner to Morrison* (New York: Palgrave Macmillan, 2008), 60. Michael Wainwright, *Darwin and Faulkner's Novels: Evolution and Southern Fiction* (New York: Palgrave Macmillan, 2008), 19-21. Dorothy Stringer, *"Not even Past": Race, Historical Trauma, and Subjectivity in Faulkner, Larsen, and Van Vechten* (New York: Fordham University Press, 2010), 39. Leigh Ann Duck, "The World of Jim Crow," in *William Faulkner in Context*, ed. John T. Matthews (New York: Cambridge University Press, 2015), 139-40. Keith Cartwright, "What Was Africa to Faulkner?," ibid., 59. Kristin Fujie, "'Through a Piece of Colored Glass': Faulkner, Race, and Mediation," *Modern Fiction Studies* 65, no. 3 (2019): 413.

[9] Noel Polk, "Man in the Middle: Faulkner and the Southern White Moderate," in *Faulkner and Race*, ed. Doreen Fowler and Ann J. Abadie (Jackson: University Press of Mississippi, 1987), 142.

[10] Thadious M. Davis, *Games of Property: Law, Race, Gender, and Faulkner's Go down, Moses* (Durham: Duke University Press, 2003), 7-8.

[11] Duvall, *Race and White Identity in Southern Fiction: From Faulkner to Morrison*, 60-61.

[12] Christina Thyssen, "'Ah kin pass wid anything': Blackness as Figural Excess in Faulkner's *Go Down, Moses*," *Faulkner Journal* 29, no. 2 (2015): 106.

[13] Don Harrison Doyle, *Faulkner's County: The Historical Roots of Yoknapatawpha* (Chapel Hill: University of North Carolina Press, 2001), 303.

[14] Stephen Railton and Christopher Rieger, "Faulkner Mapping|Mapping Faulkner," University of Virginia, http://faulkner.iath.virginia.edu/media/resources/DISPLAYS/FaulknerMapsHP.html.

[15] This includes characters whose racial identity is unknown, even to themselves like Joe Christmas in *Light in August*.

[16] William Faulkner, *Requiem for a Nun* (New York: Vintage Books, 2011), 169-73. *Absalom, Absalom!* (New York: Vintage International, 1990), 41.

[17] Piper, *Enumerations: Data and Literary Study*, 122.

[18] For a more general overview of the data limits and possibilities see Johannes Burgers, "Yoknapatawpha at 30,000 Feet: Locations, Characters, and Events in the *DY* Database," in *From Yoknapatawpha to the Virtual World: The Shape of Faulkner's Digital County*, ed. Theresa M. Towner (Charlottesville: University of Virginia Press, Forthcoming).

[19] Stephen Railton. "Instructions." *Digital Yoknapatawpha,* University of Virginia. http://faulkner.drupal.shanti.virginia.edu/content/instructions-0

[20] Jay Watson, "Introduction: Situating Whiteness in Faulkner Studies, Situating Faulkner in Whiteness Studies," in *Faulkner and Whiteness*, ed. Jay Watson (Jackson: University press of Mississippi, 2011), viii - ix.

[21] William Faulkner, "Spotted Horses," in *Uncollected Stories* (New York: Vintage, 1997), 166.

[22] *Sanctuary: The Corrected Text* (New York: Vintage Books, 1987), 180.

[23] "A Name for a City," *Harper's Magazine*, October 1950, 201.

[24] "Lion," 192.

[25] Railton and Rieger, "Faulkner Mapping|Mapping Faulkner".

[26] David J. Bodenhamer, "Making the Invisible Visible: Place, Spatial Theories, and Deep Maps," in *Literary Mapping in the Digital Age*, ed. David Cooper, Christopher Donaldson, and Patricia Murrieta-Flores (New York: Routledge, 2016), 212; Anne-Kathrin Reuschel and Lorenz Hurni, "Mapping Literature: Visualisation of Spatial Uncertainty in Fiction," *The Cartographic Journal* 48, no. 4 (2011): 294-95.

[27] Raphael Alvarado, "Characters and Locations in Force Directed Graph," http://faulkner.iath.virginia.edu/characters-force.html?text_na=FD.

[28] John et al., "A Visual Approach for the Comparative Analysis of Character Networks in Narrative Texts," 247-48.

[29] Joseph A. Veech, "A Probabilistic Model for Analysing Species Co-Occurrence," *Global Ecology and Biogeography* 22, no. 2 (2013): 252.

[30] Ibid., 254.

[31] Daniel Griffith, Joseph Veech, and Charles Marsh, "cooccur: Probabilistic Species Co-Occurrence Analysis in R," *Journal of Statistical Software* 69 (2016).

[32] Veech, "A Probabilistic Model for Analysing Species Co-Occurrence," 252.

[33] Doyle, *Faulkner's County: The Historical Roots of Yoknapatawpha*, 106-08.

[34] Wainwright, *Darwin and Faulkner's Novels: Evolution and Southern Fiction*, 103.

[35] Richard H. King, *A Southern Renaissance: The Cultural Awakening of the American South, 1930-1955* (New York: Oxford University Press, 1980), 27; Valérie Loichot, *Orphan Narratives: The Postplantation Literature of Faulkner, Glissant, Morrison, and Saint-John Perse* (Charlottesville: University of Virginia Press, 2007), 15-16.

[36] Watson, "Genealogies of White Deviance: The Eugenic Family Studies, Buck v. Bell, and William Faulkner, 1926–1931." Jay Watson makes the compelling argument that the growing eugenics movement also spurred an interest in genealogy and family history. This likely had an effect on Faulkner's early novels.

[37] Mary Paniccia Carden, "Fatherless Children and Post-Patrilineal Futures in William Faulkner's *Light in August*, *Absalom, Absalom!*, and *Go Down, Moses*," *The Faulkner Journal* 27 (2013): 72-73; Heidi Kathleen Kim, "The Foreigner in Yoknapatawpha: Rethinking Race in Faulkner's 'Global South'," *Philological Quarterly* 90, no. 2/3 (2011): 201; Jenna Sciuto, "'For Fear of a Scandal': Sexual Policing and the Preservation of Colonial Relations in William Faulkner and Marie Vieux-Chauvet," in *Faulkner and the Black Literatures of the Americas*, ed. Jay Watson and James G. Thomas (University Press of Mississippi, 2016), 185-87.

[38] For the purposes of this paper, the Bundren and De Spain characters have also been coded with their family designation. These characters were already in the database, but had not been identified by family yet.

[39] All analysis going forward relies on a *P*-value of .05 or less to test whether two characters co-occur positively or negatively.

[40] The controlled vocabularies evolved over the course of the project, and with each expansion of the terms list retroactive revisions were made to character attributes. Not all the terms that were created were utilized regularly. This is especially true in the case of "Yeoman," a distinction used by an older generation of Faulkner scholars to describe simple landholding farmers, Cleanth Brooks, *William Faulkner: The Yoknapatawpha Country* (Baton Rouge, LA: Louisiana State University Press, 1991), 10-28. Across the database Yeoman is only used to describe 28 out of 4,659 characters, and they only occur in 84 events. Their occurrence is so sparse that they were folded into lower class. A critic like Brooks might bristle at this move, because he saw them as a separate class, but in the context of *DY* they are spatially and economically similar to their other country folks. In a similarly reductive move, Indian Tribal Chief and Indian Tribal Leader all became Indian Tribal Members. There are five Asian characters in the corpus, but their appearance is so fractional that they do not register in any of the analysis. Finally, as stated earlier, all events with multiracial, multiclass, and multigender groups were removed, because these undifferentiated groups fundamentally change the context of an event that could affect co-occurrence analysis. As there is no way to ascertain the magnitude of the effect since groups are non-countable, the more conservative approach is to leave them out entirely.

[41] One reason not all combinations are possible is because the database divides class into three categories based on race, as discussed earlier.

[42] William Faulkner, *Light in August: The Corrected Text* (New York: Vintage Books, 1987), 471.

[43] Kevin Railey, *Natural Aristocracy: History, Ideology, and the Production of William Faulkner* (Tuscaloosa: University of Alabama Press, 1999), 7-8.

[44] Micki McElya, *Clinging to Mammy: The Faithful Slave in Twentieth-Century America* (Cambridge: Harvard University Press, 2007), 13; Duvall, *Race and White Identity in Southern Fiction: From Faulkner to Morrison*, 51.

[45] Rebecca Sharpless, *Cooking in Other Women's Kitchens: Domestic Workers in the South, 1865-1960* (Chapel Hill: University of North Carolina Press, 2010), 110.

[46] Railey, *Natural Aristocracy: History, Ideology, and the Production of William Faulkner*, 137-41.

[47] Malcolm Cowley, *The Faulkner-Cowley File; Letters and Memories, 1944-1962* (New York: Viking Press, 1966), 389.

[48] This appears to confirm Melanie Benson Taylor's thesis that while Native Americans are conspicuously absent from the more central history of Yoknapatawpha, they are an uncanny presence throughout the saga, Melanie Benson Taylor, "Doom and Deliverance: Faulkner's Dialectical Indians," in *Faulkner and the Native South* (University Press of Mississippi, 2019).

[49] This included all the short stories and novels such as *As I Lay Dying* and *The Mansion*, which are about relatively homogenous white communities with very few non-white characters. Curiously, it also included *The Unvanquished,* a novel about Bayard an upper class white male, and Ringo his enslaved companion. This had many co-occurrences between African-American and Anglo-American characters. As there are only four character types in the novel, the number of times they co-occur is virtually the same as would be expected at random. Consequently, there are simply not enough character types to make a meaningful differentiation. The result is clearly counterintuitive, and suggests that this co-occurrence methodology for fiction works well across the corpus, but produces contradictory results for novels and short stories.

[50] Interestingly, Boon, who is not a family member, appears in the intra-familial results because he is usually present when Lucius and Ned are.

[51] Angelina Grigoryeva and Martin Ruef, "The Historical Demography of Racial Segregation," *American Sociological Review* 80, no. 4 (2015): 815.

[52] Aliya Saperstein and Aaron Gullickson, "A 'Mulatto Escape Hatch' in the United States? Examining Evidence of Racial and Social Mobility During the Jim Crow Era," *Demography* 50, no. 5 (2013): 1935.

[53] Johannes Burgers, John Corrigan, and James Carothers. "Faulkner's *The Hamlet*." *Digital Yoknapatawpha*, University of Virginia, http://faulkner.iath.virginia.edu/index.html?text=H. Stephen Railton, Ben Robbins, and Theresa M. Towner. "Faulkner's *Sanctuary*." *Digital Yoknapatawpha*, University of Virginia, http://faulkner.iath.virginia.edu/index.html?text=SY. See the individual stories for the insets.

[54] William Faulkner, *The Hamlet: The Corrected Text* (New York: Vintage Books, 1991), 5.

[55] *Absalom, Absalom!: The Corrected Text* (New York: Vintage Books, 1990), 20.

[56] Johannes Burgers. "Racial Distribution by Event Chronology." *Digital Yoknapatawpha*, University of Virginia. http://faulkner.iath.virginia.edu/demo/Racial_Distribution_by_Event_Chronology1.mp4.

[57] "Racial Distribution by Publication Date." ibid. http://faulkner.iath.virginia.edu/demo/Racial_Distribution_by_Publication_Date1.mp4.

[58] *Flags in the Dust* (New York: Vintage International, 2012), 391-92.

[59] Johannes Burgers. "Non-white Characters in White Spaces," *Digital Yoknapatawpha*, University of Virginia. http://faulkner.iath.virginia.edu/demo/Major_Families_General_Population_White_Space.mp4.

[60] *Go Down, Moses* (New York: Vintage Books, 1990), 352; "Go Down, Moses," 257.

[61] Alex Woloch, *The One vs. the Many: Minor Characters and the Space of the Protagonist in the Novel* (Princeton University Press, 2009), 14.