The Rise of Jim Crow Rhetoric in Republican Economic Speeches*

Chris Becker

December 21, 2021

Link to Most Updated Version

Abstract

I use computational text analysis methods to quantitatively show that the Republican Party has taken the rhetoric that was used by segregationists to defend Jim Crow racial segregation in the Mid-20th Century and incorporated it into their modern-day speeches about economic issues. I analyze speeches from the Congressional Record and use machine learning methods to quantitatively identify phrases that were (i) disproportionately used in explicitly racial speeches related to Jim Crow from 1947-1967 and (ii) strongly associated with pro-segregation politicians. I then track the use of these pro-segregation phrases in speeches about economic issues into the present day. I find that over the course of the late-20th and early-21st Century, segregationist language and ideology from the Jim Crow era has become increasingly associated with the Republican Party in speeches about economic issues. However, Republicans have brought back only the abstract, legalistic language from the Jim Crow era—such as language related to "states' rights"—rather than explicitly racist language. This evidence suggests the importance of dog-whistles for signaling racial conservatism in Republican economic rhetoric.

*For exceptional advising I would like to thank Ran Abramitzky. Andy Hall and Gavin Wright also offered fantastic guidance in the completion of this paper. I also benefited tremendously from feedback by Boaz Abramson, Jaime Arellano-Bover, Travis Baseler, Eirik Berger, Sam Bufe, Alvaro Calderon, Brad DeLong, Barry Eichengreen, Dan Fetter, Vicky Fouka, Matt Gentzkow, Thomas Ginn, Nicholas Halterman, Yiming He, Florencia Hnilo, Saumitra Jha, Josh Kim, Helen Kissel, Jenna Kowalski, Ilyana Kuziemko, Alejandro Martínez-Marquina, Tamri Matiashvili, Jonas Mueller-Gastell, Joel Mokyr, Alain Pineda, Sebsatsian Puerta, Nancy Qian, Myera Rashid, Christina Romer, Wilson Sinclair, Hugh Wu, Tom Zohar and members of the Democracy and Polarization Lab at Stanford University. I also received financial support from the Dixon and Carol Doll Graduate Fellowship in completing this paper.
1. Introduction

In the Mid-20\textsuperscript{th} Century United States, Southern members of Congress were committed to an explicit ideology of White Supremacy. Their primary focus in national politics was protecting the Jim Crow system of formal racial segregation that existed in the South and offered privileged access to jobs, schools, public accommodations and voting rights to White Southerners over Black Southerners (Key 1949). But through the 1950s and 1960s all of the central pillars of Jim Crow were outlawed by the federal government in spite of near-universal opposition from the South. Slowly over time, most Southern politicians downplayed defense of explicit, formalized segregation and White Supremacy and embraced right-wing economic policies while joining the economically conservative Republican Party over the Democratic Party that had historically been associated with defense of White Supremacy in the South.

In this paper, I show that despite this decline in explicit support for racial segregation, the language and ideology of Jim Crow has survived into modern-day political debates. I use computational text analysis methods to quantitatively show that the modern-day Republican Party has taken the language that was used by segregationists to defend Jim Crow and incorporated it into their speeches about economic issues.

The quantitative evidence in my paper is consistent with the historical hypothesis of the 'Southern Strategy' which contends that the Republican Party introduced racist messaging on top of its traditional support for conservative economic policy in an effort to win over Southern Whites disenchanted with racially progressive policies put forward by the national Democratic Party. Lee Atwater, a top campaign adviser for the Ronald Reagan and George H.W. Bush presidential campaigns vividly laid out the role language played in the Southern Strategy:

\begin{quote}
You start out in 1954 by saying, '[n-word], [n-word], [n-word]'. By 1968 you can’t say [n-word]— that hurts you. Backfires. So you say stuff like forced busing, states’ rights and all that stuff. You’re getting so abstract now you’re talking about cutting taxes, and all these things you’re talking about are totally economic things and a byproduct of them is blacks get hurt worse than whites — Lee Atwater (Lamis 1999)
\end{quote}

Quotes like this one from Republican strategists provide striking anecdotal evidence that there was an attempt by the Republican Party to racialize economic policy debates. But they do not tell us how widespread the adoption of this language was within the Republican Party. Was Atwater an outlier? Or representative of a broader pattern? I overcome this
limitation by using machine learning text analysis methods to analyze millions of speeches given on the floor of Congress. I statistically identify the phrases used disproportionately by segregationists to defend Jim Crow as a way to pinpoint the language that was historically associated with White Supremacist ideology. An advantage of this method is that it will pick up on dog-whistle phrases like 'states’ rights' that are not explicitly racist but still signal commitment to racially conservative ideology, a key ingredient for the success of the Southern Strategy as outlined by Atwater. The approach is also data-driven— relying on the observed language choices of explicitly racist politicians— rather than relying on the discretion of the researcher to identify dog-whistles.

I measure all of this empirically in 4 steps. First I use automated topic modeling methods from the natural language processing literature to identify speeches from 1947-1967 that are about explicitly racial issues surrounding Jim Crow segregation and related debates. I then statistically identify two-word phrases that were disproportionately used in speeches about racial segregation relative to other topics to focus in on the language that was specifically chosen to defend White Supremacy as opposed to more generic conservative or liberal language that would show up for any topic. Then, I use machine-learning regression methods to identify which of these phrases were used by pro-segregation vs. anti-segregation politicians — as measured by conservatism of a congressperson’s voting record on explicitly racial issues— during the Jim Crow Era. Finally, I track the net usage of these pro-segregation phrases in economic speeches over time.¹

I find that over the course of the 20th Century and continuing into the present, this segregationist language from the Jim crow era has become increasingly associated with the Republican Party in speeches about economic issues as compared to the Democratic Party. Moreover, despite changes in topics and evolution of language over time, this trend has only accelerated in recent decades with the Republican Party increasing its usage of Jim Crow-era segregationist language in economic speeches since the 1980s. The language of Jim Crow did not disappear. Instead, the rhetoric of Jim Crow was repackaged into an economic agenda that emphasized right-wing economic positions like lower taxes, opposition to redistribution, and limited government intervention in the economy. The ideology of Jim Crow has been merged with the modern-day economically conservative ideology of the Republican Party.

Consistent with the Atwater quote, the qualitative content of the language I identify suggests the importance of dog-whistles in modern Republican language. I find that the language used by segregationists included both explicitly racist language (e.g. 'own race' and 'forced inte-

¹From this point forward I will use 'pro-segregation', 'racially conservative' and 'Jim Crow' language interchangeably to refer to language used to support Jim Crow segregation during the Civil Rights Era.
gration") but also more legalistic and non-explicitly racial phrases (e.g. 'federal bureaucrats' and 'usurping legislation'). An especially important theme is states' rights language that focused on challenging the legal authority of the federal government to override local and state policies regarding race. I find that the Republican Party brought back only this second category of non-explicitly racial phrases in economic debates. This suggests an important role for dog whistles. Republicans used language that was strongly associated with Jim Crow and signaled support for racial conservatism to the right audience while not being explicitly racist, offering plausible deniability in an era in which explicit racism has become increasingly taboo.

Next, I provide evidence on the mechanisms behind adoption of Jim Crow language by looking at heterogeneity across topics. Consistent with the hypothesis that Jim Crow language was used to signal racial conservatism, I find that Republicans today use more Jim Crow language in speeches about racially charged issues like crime, education, drugs, and anti-poverty programs. However, throughout the 20th Century, they actually use less Jim Crow language in speeches where they explicitly talk about race. This is consistent with a sociology and political science literature on “color-blind” racism (Williams, 1999; Spriggs and Williams, 1999; Gilens, 1996; Valentino, Hutchings, and White, 2002; Cook, 2015; Carter, 1999; Hall, 2005; Bonilla-Silva, 2015). Explicit racism is politically costly, so Republicans mimic progressive racial language when explicitly talking about race. Meanwhile, they maintain use of Jim Crow language as a signal of racial conservatism for issues like crime or poverty that are not explicitly about race, which nonetheless have racially disparate impacts. However in recent years, I find that the use of Jim Crow language in explicitly racial speeches has increased, converging toward the level in economic speeches. This suggests that increasingly Republicans are willing to signal racial conservatism even in more explicitly racial contexts.

I am able to rule out a number of alternative hypotheses to my claim that economic policy has been racialized by the introduction of Jim Crow ideology. One concern is that much of the language I identify that is associated with defense of Jim Crow such as suspicion of centralized federal power would always be associated with the Republican Party and its opposition to progressive economic policy regardless of its association with racist ideology. However, I show that initially this language is actually more strongly associated with the Democratic Party especially within the South, and even outside of the South it is not positively associated with the Republican Party in economic speeches. It is only since the mid-20th Century and the rise of federal intervention into Southern race relations that Jim Crow language has become associated with Republican economic speeches.

I can also conclude that the partisan polarization of Jim Crow language is not a function of
broader polarization of all language as identified in Gentzkow, Shapiro, and Taddy (2018); I find that polarization of Jim Crow language in economic speeches occurs in the 1960s, nearly three decades before language polarizes across-the-board. I can also rule out that the adoption of Jim Crow language by Republicans is driven by partisan realignment of Southern politicians toward the Republican Party; I find the same result when excluding Southern politicians. Consistent with the interpretation that this language is being used to signal racial conservatism, I find that use of Jim Crow language is less heavily used in less obviously racially-charged issues like energy policy and international trade.

Compared to much of the existing literature, I offer a different perspective on the relationship between economic and racial politics within the Republican Party. Both historical debates over Southern realignment in the 20th Century and more modern debates over the rise of Trumpism in the 21st Century are framed as a horse-race between economic determinants vs. racial determinants of political change. In the realignment debate, some papers advance the narrative of a white backlash against federal civil rights policies that drove southerners from the Democratic Party while others focus on economic changes in the South. Meanwhile, modern debates revolve around whether Trump voters are motivated by racial resentment or economic anxiety. My findings suggest that this framing may be misleading. Increasingly, it is actually economic issues that have become the vehicle for racially conservative ideology. The racial ideology of the 50s and 60s has been fused with the defense of a right-wing economic policy so that the two issues are not so easily separated.

Secondly, this paper suggests a growing interdependence between right-wing economic policy and racism in the United States. During the Jim Crow Era, racial hierarchy was enforced through a formal racial caste system that directly barred Black people from participation in the political and economic system. These particular modes of legalized segregation and discrimination on the basis of race were outlawed by legislation and court decisions in the 50s and 60s. However, the elimination of Jim Crow did not erase all the structural economic disadvantages facing Black Americans. They still disproportionately lived in poorer areas with under-funded schools, had significantly lower wealth and income, had worse access to good jobs, and have experienced much higher rates of incarceration and police violence, among other systemic inequities. This means that failure by the government to proactively intervene to reduce economic inequality would also tend to perpetuate racial inequality.\(^2\) This

\(^2\)For more on structural racialized economic inequalities see, for example, Darity and Hamilton, 2017; Darity and Mason, 1998; Darity, 2005; Davis, 2019; Hardy, Logan, and Parman, 2018; . On the general point about adaptability of White Supremacy, see Alexander (2010) and Cox (2014) in the context of the criminal justice system. Specific examples of structural inequalities include the racial wealth gap which ensures inter-generational transmission of economic disadvantage (Darity and Hamilton, 2017; Zewde, 2019;
aligned Southern racial priorities with the economic priorities of the Republican Party and business elite who were focused on rolling back the welfare state. I highlight the role that language played in uniting this coalition of economic and racial conservatives with historical White Supremacist ideology serving as a bridge between support for formalized Jim Crow Segregation and modern opposition to the welfare state.

2. Race, Economics, and Realignment in Historical Context

I next outline the broad patterns of political realignment along dimensions of economic and racial policy through the 20th Century. In the middle of the 20th Century, from approximately the 1930s through the 1960s, economic conservatism and racial conservatism were represented by two distinct factions. On economic issues, ideology fell largely along partisan lines. The Republican Party—which was almost completely nonexistent in the Southern States—was the primary opponent of high taxes, government spending, and the New Deal welfare state. Southern Democrats were more economically conservative than Northern Democrats but offered strong support for much of the New Deal agenda in the 1930s and still substantial but more mixed support for the War on Poverty by the 1960s. Moreover, the electoral dominance of Democrats in the "Solid South" helped to sustain long-term Democratic majorities in Congress and elect economically liberal Democratic presidents. It should be noted that while Southern politicians did support many welfare state policies, they used their influence

Jones-DeWeever; Hamilton; Rockeymoore; Tippett; Daritry, Jr., 2014), de facto residential segregation and market-based exploitation in housing markets leading to poorer access to good schools for Black students (Quick and Kahlenberg, 2019; Phillips, 2010; Bayer and Charles, 2018, Darling-Hammond, Hyler, Williamson-Lott, 2012; Darling-Hammond, 2004, Taylor, 2012; Hackett and King, 2019), worse access to employment networks (Pedulah and Pager 2019), and structural sources of health inequality (Gee and Ford 2011; Logan and Parman 2018

3A mostly qualitative, literature on the politics of this ideological merging exists in sociology, history and American political development (Dawson and Ming Francis, 2016; Bobo, Kluegel, and Smith, 1997; Dowd Hall, 2005; Inwood, 2015; Kruse, 2007; Cook, 2015; Lowndes, 2009). For a more general overview of the relationship between race and the development of the American State see Johnson (2016).

4An important exception to Southern loyalty to the New Deal is on labor policy (Bateman, Katzenelson, Lapinski, 2015; Wright, 1986; Katzenelson, Geiger, and Kryder, 1993). On this issue Republicans and Southern Democrats united to push back against pro-union New Deal legislation through the Taft-Hartley Act. See Glickman (2019) for the history of the "free enterprise" movement that emerged as a backlash against the New Deal by business elites. Labor policy was not independent of racial policy, however. Unionization especially within broad-based unions like the CIO, was viewed as a threat to the Jim Crow racial order by Southern elites.

5See Fraser and Gerstle (1989) for details on the rise and decline of this "New Deal Coalition" that supported these progressive policies.
to exclude Black Southerners from much of the benefits, so Southern conservatives never favored a racially inclusive welfare state.

Racial voting records, on the other hand, fell largely along geographic lines. National debates over race in the 1950s and 1960s revolved almost entirely around the distinctively Southern system of Jim Crow segregation. It was Southern Democrats who were the most fierce defenders of this system, whereas Northern Democrats and Republicans tended to be moderates or liberals on racial issues. A coalition of Northern Democrats and many Republicans overcame Southerners and their substantial institutional power in Congress to pass landmark Civil Rights legislation outlawing most of Jim Crow.

In the aftermath of federal intervention into Southern race relations, the New Deal coalition that supported progressive economic policies fell apart largely due to Southern ideological and partisan realignment. From the mid-20th Century into the beginning of the 21st Century Southern politicians on net defected to the Republican Party and moved to the right on economic issues. Leading modern-day conservative Southern politicians have largely built their platform around right-wing stances on issues like taxes, regulation, and redistribution. On racial issues, however, Southerners gradually gave up their defense of overt, formal racial segregation under the Jim Crow system. In the middle of the 20th Century, Southern legislative priorities revolved almost entirely around maintenance of Jim Crow racial segregation. Southern elections, where competitive, turned mostly on which candidate could position themselves as the most ardent supporter of segregation (Feldman 2011). Over time they shifted their focus to preventing proactive measures to integrate society such as busing and affirmative action. Northern Republicans, for their part, were willing to sign onto this more moderated form of racial conservatism. Northern and Southern Republicans unified around a common conservative agenda focused on right-wing economic policy but a substantially more moderate conservative platform on explicit racial discrimination as compared to segregationists of the past. Even hardcore segregationists like George Wallace distanced themselves from

---

6See e.g. Johnson (2011) for a discussion of agricultural policy and Derononcourt and Montialoux (2021) on how Black workers were initially excluded from the Fair Labor Standards act mandating a minimum wage

7Primary pillars of Jim Crow system included segregated schools and public accommodations (e.g. hotels, restaurants, stores), segregated occupations and industries with widespread discrimination in wages, and the denial of voting rights to Black Southerners all of which was supported through lynching and terrorism which largely went unpunished by the state (Epperly, Witko, Strickler, White, 2016; Cook, Logan, and Parman, 2018).

8Southerners had disproportionate power due to high-ranking positions on committees that could block Civil Rights legislation from reaching a vote. The filibuster in the Senate also prevented Civil Rights legislation from passing without widespread Northern support

9The main exception are districts with high Black populations that usually elect liberal Democrats since the Voting Rights Act dramatically increased voting access for Black Southerners

10This can be seen in the shift in political platforms over time. The third party Dixiecrats who had electoral
their segregationists pasts. In later years, Wallace amended his famous statement that he supported "segregation now, segregation tomorrow, segregation forever" claiming that what he should have said was "states’ rights now, states’ rights tomorrow, states’ rights forever" (Carter, 1999). See Appendix A for quantitative measurement of these policy shifts.

The literature on this political realignment has often treated race and economics as distinct, competing explanations of political change. One set of explanations places racial issues at the heart of realignment (Sears and Valentino, 2005; Washington and Kuziemko, 2018; Carmines and Stimson, 1989; McKee, 2017; Schreckhise and Shields, 2003). The basic idea is that because of the historical legacy of the Civil War and Reconstruction, the South was dominated by the Democratic Party. So long as Northern Democrats refrained from interfering in Southern race relations, Southerners went along with most of the liberal economic agenda of the national Democratic Party (Black and Black, 2003). But in the 1940s-1960s the Northern Democratic Party moved aggressively in favor of expanded economic and political rights for Black people. This culminated in the prohibition of all of the central pillars of Jim Crow which led to a political backlash by White Southerners against the Democratic Party. These Southerners then increasingly adopted the economic positions of the national Republican party. An alternative explanation suggests that ideological and partisan realignment was not driven by racial backlash, but instead by the growth of a middle class and business elite that cared more about economic issues than race (Johnson and Shafer, 2009; Knotts, Abramowitz, and Saunders, 2005; Brewer and Stonecash, 2001). By this argument, shifts in demographics and the Southern economy are the fundamental driving force behind Southern realignment rather than White Backlash.

success in Southern States in 1948, adopted a platform revolving almost entirely around the protection of Jim Crow. On the other hand, the landmark "Contract With America" written by Southern Congressmen Newt Gingrich and Dick Armey which laid the groundwork for the Republican legislative agenda for several years in the 1990s included 6 proposals related to tax policy and welfare reform, but none related to civil rights issues.

---

11 Another famous example is Strom Thurmond. In 1965 he was one of the staunchest opponents of the Voting Rights Act in the Senate. But by 1982, the famous segregationist supported renewal (Roberts 1982).

12 These details on ideological realignment mirror a point made in the political science literature on polarization (see e.g. Poole and Rosenthal 2007). In the middle of the 20th Century American politics was "two-dimensional" in the sense that economic ideology largely fell along party lines while racial ideology fell along North-South lines. But, the ideological realignment discussed above along with the partisan realignment of the South from the Democratic to the Republican Party meant that the full spectrum of political issues—economic, racial, and more generally—now fall along party lines.

13 Most studies focus on partisan realignment: the movement of Southern politicians and voters away from the Democratic Party toward the Republican Party. But in many ways ideological realignment preceded partisan realignment. Early opponents of the New Deal from the South were all Democratic politicians. It would be several decades before economically conservative Southerners actually identified as Republican. Some studies such as Kruse (2007) and Lassiter (2006) provide historical accounts of the interactions between economics and race. Lowndes (2009), like me, focuses on the merging of economic and racial conservatism and what brought together Northern Republicans and Southerners through qualitative case studies.
I argue that this bifurcation between economic vs. racial policy and ideology is misleading because increasingly it is right-wing economic policy that has become the vehicle for racially conservative language. On the surface, Republicans are emphasizing explicitly economic issues like taxes or labor that are not explicitly about race. But through the use of dog-whistle language historically associated with defense of explicit segregation, they have infused racial ideology into their messaging around right-wing economic policies.

This is not to say that economic policy was not at all racialized prior to the Civil Rights Movement. Economic policy and racial hierarchies were always interconnected. Even the Jim Crow system itself was a system of economic exclusion of Black Southerners as much as cultural exclusion (Cook, Jones, Rosé, and Logan, 2018; Wright, 2013). What I am showing is that increasingly the ideology of Jim Crow has become tied to the Republican Party and defense of right-wing economic policy more generally rather than defense of more explicit formal segregation and racism.14

3. Quantitatively Identifying Jim Crow Language from Congressional Speeches

3.1 Text Analysis Methods and Empirical Specification

To measure the diffusion of segregationist rhetoric from speeches about Jim Crow into speeches about economic issues, I analyze language choices by politicians on the floor of Congress, using data compiled by Gentzkow, Shapiro, and Taddy (2019). This includes all speeches given on the floor of the House or Senate that could be matched to a single politician. To reduce noise, I take a number of standard pre-processing steps to simplify the language under consideration.15 The goal is to identify language used to defend Jim Crow during the Civil Rights era and then track the use of this language in economic speeches going forward. I proceed in four steps. (i) I identify speeches from 1947-1967 (this roughly corresponds to the height of the Civil Rights Movement and debate over Jim Crow) about explicitly racial issues surrounding Jim Crow segregation. (ii) Next, I filter down to phrases that are disproportionately used in racial speeches during this period relative to other topics to focus

---

14 For more qualitative work on the interest alignment of economic conservatives and racial conservatives and the growth of a cohesive political coalition between these two factions see Dawson and Ming Francis (2016), Brewer (2013), Hall (2005), Inwood (2015), Bobo, Kluegel, and Smith (1997); Robinson (2016) and Cook (2015)

15 See Appendix B for details.
in on language that was specifically chosen when talking about race. (iii) To identify pro-Jim Crow language, I use machine learning regression methods to identify two-word phrases that were disproportionately used by pro-segregation politicians vs. anti-segregation politicians in these debates over race. And finally (iv) I track the net usage of the pro-segregation phrases (weighted by regression coefficient to account for the intensity of the association with segregationists) in speeches about economic issues to measure the diffusion of Jim Crow language into economic speeches. I outline each step below.

In order to identify language used in speeches about segregation and track the usage of this language in economic speeches, I must distinguish speeches about racial and economic issues from other topics. Given that there are millions of speeches in the Congressional Record, this necessitates an automated method of topic classification for which I use Latent Dirichlet Allocation (Blei, Ng, and Jordan 2017). The model is completely unsupervised and infers topics from the speech data without any guidance from the researcher about the content of the topics. The intuition of this method is that it identifies keywords using patterns of co-occurrence in text data. So, for example, the model will recognize that words such as 'budget', 'deficit', and 'appropriation' are often used together and infer that these words constitute a topic. Ex-post the researcher could look at this list of keywords and realize that this is a topic related to the federal budget. I estimate a topic model for the period from 1947-1967 to identify speeches about racial issues at the height of the Civil Rights Movement for the purpose of identifying the language of Jim Crow and then again for the period from 1947-2011 to identify a longer panel of speeches related to economic issues to study long-term trends in the use of Jim Crow language. In both cases, the LDA model clearly identifies topics related to racial and economic issues respectively.16

The content of the topics identified as being about racial issues from the Jim Crow era are listed in Table 1. For each topic I display 2-word phrases that are disproportionately used in speeches about each topic compared to all other speeches.17 The first topic largely captures the content of policy debates surrounding racial issues with explicit references to race and racial issues ("Negro Vote", "Civil Right", "Race Color"). The second largely captures the

---

16 Technical notes: Most of this process is automated. But the number of topics must be user-specified. I choose 100 for each model. The model also does not select a single topic for each speech, rather each speech is modeled as a mixture of various topics. Unless stated otherwise, I count a speech as being about a topic if a non-zero proportion is assigned to that topic (proportions tend to be sparse; so usually only a small number of topics is assigned for each speech). Additionally, many topics are related to procedural issues related to the rules and formalities of Congress. Because I want to focus in on the most substantive speeches, I label each topic as substantive or procedural based on the content of the keywords and only include speeches for analysis that are at least 40% substantive.

17 The model estimates topics based on single words, not two-word phrases. But two-word phrases are often easier to interpret.
legal debates surrounding Civil Rights issues. It includes many words related to race ("Civil Right", "Race Color"). But many words are about law ("Federal Court", "Due Process", "Equal Right"). This association between racial keywords and legal keywords suggests that legal issues were central to congressional debates over race which is consistent with the historical record. Main points of contention included the role of the courts in adjudicating and enforcing integration as well as the balance of power between the federal and state governments in governing race relations (see more below). Therefore, in understanding Jim Crow ideology it is important not just to consider explicit racism, but also the more abstract, legalistic framework that was used to defend White Supremacy.\textsuperscript{18}

Next, I restrict attention to phrases that are disproportionately used in racial speeches relative to other topics. This has a few advantages. The first is computational. In the regressions performed below, there are many variables (distinct two-word phrases used by politicians) relative to observations (speeches) leading to potential for overfitting. Machine learning can partially account for this high-dimensionality, but pruning the number of phrases in the analysis can still reduce noise. More importantly, there is a substantive advantage. I want to focus on phrases that were specific to racial conservatism. If I include all phrases when comparing language between racial liberals and conservatives, I am more likely to pick up on phrases that are associated with conservatism more generically. For example, it could be that all types of conservatism (economic, racial and otherwise) caution against sweeping societal change in which case this conservative language would not be distinctive to Jim Crow ideology specifically. By filtering down to phrases that were disproportionately chosen for racial speeches, I can identify conservative language that was distinctive to racial ideology in the Mid-20\textsuperscript{th} Century.

Formally, I follow Gentzkow and Shapiro (2010), by calculating a chi-squared statistic which tells me how disproportionately each phrase is used in racial speeches versus speeches generally. For each two-word phrase, p, I calculate $C_{p,r}$ and $C_{p,nr}$ which indicated the number of times phrase p is used in racial and non-racial speeches respectively. I also calculate $C_{\sim p,r}$ and $C_{\sim p,nr}$ which indicate the number of times any phrase besides p is used in racial and non-racial speeches respectively. I then calculate:

\textsuperscript{18}We can also see that there is likely some error in these methods. For the first topic 'nuclear test' shows up as a word associated with race. This is likely due to the model correctly identifying 'test' as a word associated with racial debates due to the focus on literacy tests as a method for disenfranchising Black voters. In column 2 'antitrust law' shows up as an associated phrase. This is somewhat less far-fetched, but still plausibly the result of error. Despite this noise, a key advantage of automated topic classification is that I do not have to use my own discretion to decide what counts as being relevant to race and instead rely on patterns within the data to guide the classification.
\[ X_p^2 = \frac{(C_{p,r}C_{\sim p, nr} - C_{p, nr}C_{\sim p, r})^2}{(C_{p,r} + C_{p, nr})(C_{p,r} + C_{\sim p, r})(C_{p, nr} + C_{\sim p, nr})(C_{\sim p, r} + C_{\sim p, nr})} \times \text{sign}(C_{p,r}C_{\sim p, nr} - C_{p, nr}C_{\sim p, r}) \]

where \( \text{sign}(C_{p,r}C_{\sim p, nr} - C_{p, nr}C_{\sim p, r}) \) is added to indicate whether phrase \( p \) is more associated with racial speeches or non-racial speeches. I take the 10,000 phrases with the highest chi-squared statistic (indicating more disproportionate use in racial speeches) separately for all politicians, politicians above the mean value for conservatism of racial voting record in their chamber, and politicians below the mean level (to ensure I include phrases used by both sides). I then take the union of the phrases identified for all 3 groups. This procedure yields 14,229 phrases, about 7% of the total vocabulary post-processing. So, I am including only a small portion of phrases that were most associated with racial debates.

Once I have identified a set of speeches related to racial debates from the 1940s-1960s and phrases that were strongly associated with these debates relative to other topics, I next want to figure out which phrases were used by proponents vs. opponents of racial segregation. I estimate the following regression where observations are at the speech level:

\[ R_i = \beta X_{ik}^{\text{race}} + \epsilon_{ik} \]

where \( R_i \) is the racial conservatism of legislator \( i \) who gave speech \( k \) based on their voting record on Civil Rights issues as measured by the Nominate scaling procedure described in Appendix A. A more positive score indicates a more extreme voting record in favor of racial segregation. \( X_{ik}^{\text{race}} \) is a vector indicating the language choices used by politician \( i \) in racial speech \( k \). In particular, each element \( X_{ikp}^{\text{race}} \) indicates the proportion that politician \( i \) uses phrase \( p \) out of all phrases used by politician \( i \) in racial speech \( k \). Most entries are zero because most two-word phrases are not used in any given speech. A more positive number indicates greater use of that phrase in speech \( k \). Then, each estimated coefficient \( \hat{\beta}_p^{\text{race}} \) can be interpreted as the extent to which phrase \( p \) is used disproportionately by pro-segregation politicians, with more positive numbers indicating a more positive association with the pro-segregation side and negative numbers indicating greater use by the anti-segregation side.

Because there are so many phrases, OLS estimation is susceptible to overfitting, a problem that can be overcome with machine learning methods. I estimate the regression using ridge minimization. Formally, \( \beta \) is chosen to minimize \( \sum_{k=1}^{N}(R_i - \beta X_{ik}^{\text{race}})^2 + \lambda \sum_{j=1}^{P} \beta_j^2 \). The term \( \lambda \sum_{j=1}^{P} \beta_j^2 \) penalizes putting higher weight on regression coefficients, shrinking all toward zero.
\( \lambda \) is chosen to minimize the sum of squared errors using cross-validation, settling on a \( \lambda \) of 0.1. The ridge specification (as opposed to Lasso) ensures that smaller weights will be spread across all phrases rather than setting most coefficients equal to zero and focusing on a few phrases.\(^{19}\) I do this so my measure of racial language does not rely too heavily on just a small set of words and to give a richer sense of segregationist ideology. The estimated regression achieves an \( R^2 \) of 0.209.

Finally, to track the use of this language in economic speeches, I calculate on net how often (weighted by regression coefficient) these pro-segregation phrases identified in 2 are used in speeches about economic issues. In particular, I calculate:

\[
Net \ Jim \ - \ Crow \ Language \ in \ Economic \ Speech_k = \hat{\beta}^{race} \ x_{ik}^{economic}
\]

where \( X_{ik}^{economic} \) is the vector of phrase proportions used by politician \( i \) in economic speech \( k \). A higher number for this index indicates more intensive use of pro-segregation phrases (relative to anti-segregation phrases) in economic speech \( k \). Stated differently, this is just the predicted value of regression 2 which identified pro-segregation language in racial speeches when language choices in economic speeches are substituted. So, the calculated value captures how similar a politician’s language choices in an economic speech are to the language used by pro-segregation politicians in the Mid-20\(^{th}\) Century to defend Jim Crow. This will be my main outcome of interest.

### 3.2 Qualitative Content of Pro-Segregation Language

Table 2 shows some selected phrases from the 50 phrases most associated with the pro- and anti-segregation side estimated from regression 2. For both sides, there are two types of language that pop out, each indicated by different colors. First, there are phrases that are explicitly about race. Focusing on this language in blue, we can see that racial conservatives were more likely to use some explicitly racist language ("own race", "race riot"). Whereas anti-Jim Crow politicians were more likely to call out injustices of racism ("racial discrimin", "discrimiantori practic").

But much of the language focuses on more abstract, legalistic concepts. The anti-segregation language seems very focused on issues of human rights ("human digniti", "peace justic", "civil liberti"). This generally makes sense; appeals for civil rights were often defended on moral

\(^{19}\)Main results are robust to choosing a Lasso specification instead.
grounds with appeals to concepts of justice, fairness, and equality. On the other hand, the pro-segregation language is largely focused on the proper scope of government power and the balance of power between national and state governments ("central govern", "constitut govern", "federal bureaucrat", "usurp legis"). This makes it clear that even in defending explicitly racist policies of segregation, politicians often did not use explicitly racist rhetoric. Instead they defended racism on the basis of abstract, legalistic objections. However, this strand of rhetoric has deep historical roots in the defense of White Supremacy in the United States. In defending both slavery in the 1800s and segregation in the 1900s the concepts of states’ rights and federal overreach were often central philosophical underpinnings of racist ideology. The political logic was simple. So long as Southern States were permitted to set their own laws on race relations, they would overwhelmingly choose to maintain White Supremacy. It was the federal government and expansion of its powers that primarily threatened (and ultimately dismantled) Jim Crow.²⁰

4. Results

I now examine the diffusion of this pro-Jim Crow language into economic speeches. Figure 1 shows the average use of pro-Jim Crow language in speeches about economic issues by Republicans vs. Democrats over time. Initially, in the first half of the 20th Century, pro-segregation language is actually associated with the Democratic Party not the Republican Party. But beginning in the 1960s, at the peak of the Civil Rights Movement, the use of pro-Jim Crow language becomes more associated with the Republican Party instead. Figures C1 - C10 show trends for a variety of other economic issues. There are slight differences across them, but in all cases the story remains the same with pro-segregation language becoming more strongly associated with the Republican Party over time.

This trend has only accelerated in recent decades. Since the 1980s, the use of pro-Jim Crow language has become increasingly polarized and the Republican Party has actually increased its usage of segregationist language from the past even in absolute terms. This result is strengthened by the potential for attenuation bias resulting from broader societal language changes. Over time, there will be turnover in the particular phrases that are used by English speakers. For example, whereas mid-Century politicians would have used the term "colored people", today politicians would say "Black people." I identify language from a particular historical era. This means that many phrases will fall out of use due to broader language changes.

²⁰For more on the historical connection between the American federalist system of government and White Supremacy, see Mulcare, 2008; Einhorn, 2000; King, 2011; Weir, 2015; Lieberman and Lapinski, 2001.
change rather than changing political incentives. Conversely, new racialized language has been developed over time ("welfare queen"), but my method will not incorporate these newer phrases. This makes comparing absolute magnitudes over time more difficult and may partially explain why use of Jim Crow language is smaller in absolute terms than in the past. However, gaps between parties should still be meaningful because both parties would be subject to any broader language trends. Overall, this language churn should result in attenuation bias due to my method relying on a shrinking set of potential vocabulary. Despite this, I find growing polarization and increased usage by Republicans in recent decades.

Next, I turn to analysis of the qualitative content of the segregationist language being brought back by the Republican Party in its economic speeches. To get at this, I compare the pro-segregation phrases used by segregationists in racial speeches during the Jim Crow Era to the pro-segregation phrases used by Republicans in economic speeches in the modern period. I quantify the contribution of individual phrases to the gap in pro-segregation language between Southerners vs. Non-Southerners in racial debates between 1947-1967 and Republicans vs. Democrats in economic speeches between 1967-2017. Considering the comparison of Republicans to Democrats, for each two-word phrase \( k \) I calculate:

\[
Contribution\ Phrase_k = (freq_{rep}^k - freq_{dem}^k) \hat{\beta}_{race}^k
\]

where \( freq_{rep}^k \) and \( freq_{dem}^k \) indicate the proportion that phrase \( k \) is used by Republicans and Democrats across all speeches about economic issues from 1967-2017 and \( \hat{\beta}_{race}^k \) is the coefficient from equation 2 that indicates the association of phrase \( k \) with pro-Jim Crow politicians. So, this accounts for how strongly each phrase is associated with Jim Crow as well as how much more frequently it is used by Republicans in economic speeches. I then do this same calculation for racial debates between 1947-1967 but comparing Southern vs. Non-Southerners politicians.

Table 3 compares the 20 highest-contributing phrases across the 2 eras and demonstrates the shift toward more abstract, non-explicitly racist language. For segregationists in racial speeches during the Jim Crow era, there is a mixture of explicitly racial language ("white peopl", 'color peopl", 'poll tax") and more abstract language ('feder govern', 'central govern', 'constitut unit"). Though even in this period, a majority of the most influential words are of the more abstract variety even for these explicit segregationists. However, for modern-day economic speeches, the pro-segregation language distinctive to the Republican Party is all abstract and philosophical covering themes related to states’ rights and government authority ('feder govern', 'feder control', 'govern power'), private property ('private properti", 'properti
right") and references to the founding of the country ("found father", "thoma jefferson"). This is consistent with the idea that the Republican Party is using dog whistles; language that signals a commitment to racial conservatism to the right audience but is not explicitly racist in order to insulate the Republican Party from charges of racism.

This finding is further confirmed by looking at changes in the use of these individual pro-segregation phrases in economic speeches over time. Figure 2 shows trends in the relative frequency of individual words from Table 3 in economic speeches over time. Initially, phrases explicitly related to race or key Civil Rights debates such as "white peopl", "color peopl", and "poll tax' are relatively common in economic speeches. But over time these fade away almost completely. On the other hand, phrases related to states’ rights ("feder bureaucrat", 'local level") and references to "found father" and "private properti" increase or maintain their usage in economic speeches. It is the more abstract phrases from the Jim Crow era which have survived into modern-day economic speeches.

Beyond the usefulness of states’ rights language as a dog-whistle for signaling racial conservatism, there is reason to believe delegation of power over economic policy to local and state governments disproportionately harms Black Americans. Evidence for the substantive role that federalism plays in maintaining racial hierarchy is especially well-established for welfare policy. Welfare reform enacted in 1996 gave substantial power to states in the administration of welfare policy. There is empirical evidence that this gave states with large Black populations the ability to invest less in antipoverty programs in general and to use discretion to favor White applicants over Black applicants (Hardy, Samudra, and Davis, 2019; Soss, Schram, Vartanian, and O’Brien, 2001; Schram, Soss, Fording, and Houser, 2009; Davis 2007; Michener, 2018).

Comparing the relative use of Jim Crow language across topics supports the interpretation that Jim Crow language was infused in economic policy debates as an effort to racialize these discussions. My goal is to compare the use of Jim Crow language across topics that vary in the extent to which they are racially-charged with the hypothesis that Republicans would use more Jim Crow language in these more racially-charged debates. I first look at 4 topics that have been major flashpoints in public debates over racism: crime, education, drugs, and anti-poverty programs.

The issues with the most obvious connections to racism in public debates are issues related to crime and drugs (although these issues are related, my topic model separates these 2 topics). Especially since the 1990s— which saw the rise of mass incarceration and the war on drugs as well as renewed fear-mongering about immigration— issues of crime and drugs
have become focal points for debates over the role of racism in American society with social movements like Black Lives Matter and pro-immigrant groups seeking to challenge racial injustices in these areas. Education policy was central to past racialized debates over issues like desegregation of public schools and busing while even to this day persistent gaps in educational attainment are often viewed as a major obstacle to racial economic inequality. Anti-poverty programs have also received attention as being particularly racialized with opponents invoking racially-charged tropes to denigrate welfare recipients and academic work focused on bias in provision of benefits (Davis 2007; Gilens 1996; Bradley, Samudra, Davis 2019; Schram, Soss, Fording and Houser 2009; Valentino, Hutchings, and White 2002).

Conversely, I focus on energy policy and international trade as issues that are potentially less racialized in public debate. While almost all issues have some connection to racism meaning that there is no perfect placebo topic, it is less obvious that issues like international trade or energy have been a focal point of broader debates over race in American political discourse. To validate that my ordering of these topics in terms of connections to race/racism corresponds to public perceptions, Figure C15 plots the use of keywords related to each topic after the word "race" in Google Books data (Lin, Michel, Aiden, Orwant, Brockman and Petrov 2012) over time. Both "crime" and "school" are used in conjunction with race much more frequently than the other topics, and "drug" is more often than "trade" or "energy."

Figure 3 focuses on just Republican speeches since the late-1960s the turning point I identify after which Jim Crow language became more associated with the Republicans. For each figure, the line in grey indicates the use of Jim Crow language in all economic speeches aggregated together as in Figure 1. The line with color indicates the use of Jim Crow language in speeches identified as the corresponding topic. For the more racially-charged topics— drugs, education, crime, and anti-poverty programs— generally the Republican Party uses more Jim Crow language than for typical economic speeches (although all 4 topics see a decline relative to economic speeches in the 1980s followed by a resurgence). On the other hand, for less racialized topics— trade and energy— generally much less Jim Crow language is used by Republicans, even largely falling below zero indicating that the language is more similar to the anti-segregation side.

The story is very different for when Republicans talk explicitly about race, however. Figure 4 compares the use of pro-Jim Crow language in speeches about economic issues vs. speeches explicitly about race for Southern politicians at the height of the Jim Crow era and

---

21 All words chose are also top keywords identified for each corresponding topic in my topic model.
22 This is not the same topic as either of the topics used to identify Jim Crow language in racial debates discussed above. Instead, this topic emerges out of the separate topic model used to identify economic
for Republicans since the late-1960s. Southerners during the Jim crow era (largely by construction) use pro-Jim Crow language much more often when talking explicitly about race compared to when talking about economic issues. But for modern-day Republicans this pattern is flipped. Republicans through most of the 20th Century actually use more Jim Crow language in speeches about economic issues compared to speeches where they explicitly talk about race.

This is consistent with a shift toward "color-blind" racism (Williams, 1999; Spriggs and Williams, 1999; Gilens, 1996; Valentino, Hutchings, and White, 2002; Cook, 2015; Carter, 1999; Hall, 2005; Bonilla-Silva, 2015). Explicit racism is politically costly, so when explicitly talking about race, Republican politicians have an incentive to mimic progressive anti-racist values. But in less explicitly racial contexts—as when discussing crime, drugs, education, and anti-poverty programs—it is safer to signal commitment to segregationist values. However, in recent years the use of pro-Jim Crow language in racial speeches by Republicans has converged to the level used in economic speeches. This suggests that increasingly the Republican Party is willing to use racialized language and ideology even in more explicitly racial contexts. Altogether, this suggests the importance of looking at racialized policy impacts of ostensibly 'race-neutral' policies for understanding race in American political economy.

5. Ruling Out Alternative Hypotheses and Robustness

I now turn to the robustness of my findings and attempt to rule out alternative hypotheses to my claim of increased racialization of economic policy. One potential concern is that these results could be driven by partisan realignment in the South. As Southern politicians switch to the Republican Party, they may maintain their use of segregationist language and drive a positive correlation between Republican identification and use of segregationist language. But this would just be a composition effect driven by the switching of party labels. To get at this, Figure C11 shows the trend in pro-segregation language in economic speeches when Southern congresspeople are excluded. Unlike before, in the first half of the 20th Century use of pro-segregation language is not used more by the Democratic Party (this was driven by Southern Democrats). But it is also still not more strongly used by the Republican Party. However, the main finding—that in the mid-20th Century use of pro-segregation language in economic speeches becomes increasingly associated with the Republican Party—still holds up.
So, the use of pro-segregation language by the Republican Party is not driven by Southerners alone or party realignment.

Another potential concern is that language supporting states’ rights, property rights, and constitutionality like that found in Table 2 would naturally be associated with Republicans and critics of government intervention in the economy under all circumstances regardless of any connections to race. But this is clearly not the case; I find that initially this language is associated with the Democratic Party overall and not positively associated with the Republican Party even outside of the South. It is only in the wake of federal intervention against the Jim Crow system that the trend toward Jim Crow language in Republican speeches occurs.

The same picture emerges when I look at use of Jim Crow language for economically conservative vs. liberal politicians based on their voting records. It is important to look at voting records separately from party because economic voting records have become more polarized by party over time especially within the South. Figure C12 shows the correlation at the speech level between the use of pro-segregation language and the conservatism of a speaker’s economic voting record. Initially there is no correlation between conservatism of economic voting record and use of pro-segregation language in economic speeches. But over time a positive correlation develops continuing into the present day. This suggests that segregationist language has not always been tied to right-wing economic positions. This has only been true since the Civil Rights era began.

It is also clear that this result is not a function of broader trends in polarization of political language. Gentzkow, Shapiro and Taddy (2019) analyze the same Congressional Record data and show that language generally across all issues polarized by party in the early 1990s. However, I show that polarization of Jim Crow language in economic speeches precedes that shift by 3 decades.

6. Conclusion

In 1967, just a few years removed from the Civil Rights Act and Voting Rights Act that marked the greatest legislative victories for the Civil Rights movement, Martin Luther King Jr. set his sights on a "new phase" of the Civil Rights Movement. He acknowledged the substantial progress that had been made by the 'breakdown of legal segregation' in the South. But he recognized that true equality for Black Americans depended not just on legal equality, but also necessitated achieving economic equality. He argued that the Civil Rights Movement
must call on the federal government "to get rid of poverty, to get rid of slums, to make quality integrated education a reality." He explicitly made the connection between the 'evils' of racism and poverty. But he noted that this struggle for economic equality would more difficult than the first. He noted that 'the gains [against legal segregation] were obtained from the power structure at bargain rates; it didn’t cost the nation anything to integrate lunch counters. It didn’t cost the nation anything to integrate hotels and motels. It didn’t cost the nation a penny to guarantee the right to vote. Now we are in a period where it will cost the nation billions of dollars [to achieve economic equality]' (King, 2018)

King’s words today seem prophetic. Indeed, demands for economic equality riled up enormous resistance from the economic elite represented by the Republican Party. Unlike demands for legal segregation in the South, the power and profits of the Northern business elite were directly threatened by the egalitarian policies demanded by King. But they too saw the connection between racism and economic inequality and along with it a political opportunity. By appealing to the racism of Southern whites and politicians, the Republican Party could expand support for its right-wing economic agenda and beat back demands for economic justice.

This paper demonstrates one of the ways that the Republican Party appealed to these Southern White Supremacists. It took the language and ideology of Jim Crow and incorporated it into their arguments for low taxes, limited government spending and opposition to redistribution. It turned these economic issues into racial ones and made substantial progress in moving American economic policy to the right. Instead of slowly fading over time, this racialization of economic issues still lives with us through a modern Republican Party that has revitalized Jim Crow rhetoric. And this coalition between economic racial conservatives has proven durable. The limited progress on economic and racial equality that we see today is largely the result of this successful gambit by the Republican Party.

This paper also speaks to the importance of looking at non-explicitly racial issues for understanding the economics and politics of racial inequality. I show quantitatively that Republican downplayed segregationist rhetoric in speeches explicitly about race, but injected it into issues like drug, crime, education and anti-poverty policy that have important implications for racial inequities. This suggests the importance of looking beyond surface-level rhetoric and focusing only on explicit racial discrimination in understanding the racialized implications of American policy and politics.
7. References


Fox, C. 1989. 'A Stop List for General Text,' *SIGIR*.


23


Roberts, Steven V. 1982. “Voting Rights Act Renewed In Senate By Margin Of 85-8.” The


### Tables and Figures

Table 1: Selected Top 40 Key Phrases for Topics Related to Jim Crow Debates

<table>
<thead>
<tr>
<th>Topic 1: Race Relations</th>
<th>Topic 2: Legal Issues Civil Rights Policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negro Vote</td>
<td>Supreme Court</td>
</tr>
<tr>
<td>Civil Right</td>
<td>Civil Right</td>
</tr>
<tr>
<td>Color People</td>
<td>Federal Court</td>
</tr>
<tr>
<td>Race Color</td>
<td>Race Color</td>
</tr>
<tr>
<td>White People</td>
<td>Constitutional Right</td>
</tr>
<tr>
<td>Vote Right</td>
<td>Vote Right</td>
</tr>
<tr>
<td>Negro Citizen</td>
<td>Due Process</td>
</tr>
<tr>
<td>Poll Tax</td>
<td>Poll Tax</td>
</tr>
<tr>
<td>Regardless Race</td>
<td>Equal Protect</td>
</tr>
<tr>
<td>Nuclear Test</td>
<td>Antitrust Law</td>
</tr>
</tbody>
</table>

Note: These are a selection of 2-word phrases most disproportionately associated with the two topics identified by the topic model described in Section 3 that I have labeled as being related to Civil Rights and Jim Crow (association of each 2-word phrase with each topic is calculated the same way as in Equation 1 but separately for each topic). The model is run for Congressional speeches from 1947-1967 and 100 topics are identified automatically. I count a topic as being about race if it puts high weight on keywords directly related to race or racial debates from 1947-1967 (e.g. "civil" and "right", "white", "black" etc.) See discussion in Section 3. Each speech is assigned multiple topics, each of which is given a proportion/probability. I assign each speech as being about race if any proportion is assigned to either of the two topics listed above (the model is sparse, so that most speeches are only assigned a small subset of the topics), Equation 2 is then estimated for this sample of speeches.
Table 2: Selected Top 50 Most Pro-Jim Crow vs. Anti-Jim Crow Phrases

<table>
<thead>
<tr>
<th>Pro-Jim Crow Phrases</th>
<th>Anti-Jim Crow Phrases</th>
</tr>
</thead>
<tbody>
<tr>
<td>own race</td>
<td>fellow american</td>
</tr>
<tr>
<td>forc integr</td>
<td>world peac</td>
</tr>
<tr>
<td>subcommitte constitut</td>
<td>human digniti</td>
</tr>
<tr>
<td>qualif voter</td>
<td>civil liberti</td>
</tr>
<tr>
<td>school integr</td>
<td>school desegreg</td>
</tr>
<tr>
<td>control elect</td>
<td>color skin</td>
</tr>
<tr>
<td>central govern</td>
<td>social econom</td>
</tr>
<tr>
<td>race south</td>
<td>discriminatori practic</td>
</tr>
<tr>
<td>feder bureaucrat</td>
<td>equal represent</td>
</tr>
<tr>
<td>own properti</td>
<td>negro american</td>
</tr>
<tr>
<td>racial integr</td>
<td>discrimin american</td>
</tr>
<tr>
<td>imparti juri</td>
<td>fellow citizen</td>
</tr>
<tr>
<td>white color</td>
<td>fair procedur</td>
</tr>
<tr>
<td>racial equal</td>
<td>local ordin</td>
</tr>
<tr>
<td>race riot</td>
<td>human valu</td>
</tr>
<tr>
<td>integr school</td>
<td>peac justic</td>
</tr>
<tr>
<td>voter qualif</td>
<td>racial discrimin</td>
</tr>
<tr>
<td>white peopl</td>
<td>captiv nation</td>
</tr>
<tr>
<td>constitut govern</td>
<td>citizen regardless</td>
</tr>
<tr>
<td>usurp legisl</td>
<td>vote discrimin</td>
</tr>
</tbody>
</table>

Note: This is a selection of phrases that are most disproportionately associated with pro-segregation or anti-segregation politicians in speeches about topics related to Jim Crow. The association of each 2-word phrase with pro- and anti-segregation politicians is estimated according to equation 2. On the left-hand side I include a selection of phrases from the top 50 phrases with the highest positive coefficients (indicating the strongest association with pro-segregation politicians). On the right-hand side I include a selection of phrases from the top 50 phrases with the most negative coefficients (indicating the strongest association with anti-segregation politicians). Each phrase is classified as being explicitly racial or abstract/legalistic ex-post.
Figure 1: Net Use of Pro-Jim Crow Language (Relative to Anti-Jim Crow Language) in Speeches about Economic Issues: Republicans vs. Democrats

Note: This figure measures the net usage of pro-segregation phrases (relative to anti-segregation phrases) in speeches about economic issues (as identified through the topic model described in section 3 where a speech is included if at least 25% of the topic proportions assigned by the topic model are related to economic issues; e.g. keywords related to taxes, labor, health care etc.). An observation is at the speech level. For each speech the level of pro-segregation rhetoric is calculated according to equation 3. For each 2-word phrase, the frequency of that phrase in the speech is calculated (as a proportion of total words) and then multiplied by the coefficient that measures its association with pro-segregation politicians in Jim Crow era racial speeches as estimated in equation 2. Then, I sum across all phrases in the speech. A more positive number indications more intense usage of pro-segregation phrases (net of anti-segregation phrases). For each 2-year term of congress I calculate the average pro-segregation score of all speeches by Democrats and Republicans and plot these scores over time. Standard error bars for a 5% level of statistical significance are included.
Table 3: Individual Phrases that Contribute the Most to Gaps in Pro-Jim Crow Language

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>trial juri</td>
<td>feder govern</td>
</tr>
<tr>
<td>socal civil</td>
<td>privat properti</td>
</tr>
<tr>
<td>north carolina</td>
<td>north carolina</td>
</tr>
<tr>
<td>south carolina</td>
<td>feder bureaucrat</td>
</tr>
<tr>
<td>constitut unit</td>
<td>found father</td>
</tr>
<tr>
<td>qualif voter</td>
<td>labor union</td>
</tr>
<tr>
<td>poll tax</td>
<td>pass law</td>
</tr>
<tr>
<td>feder govern</td>
<td>south carolina</td>
</tr>
<tr>
<td>white peopl</td>
<td>govern govern</td>
</tr>
<tr>
<td>numer branch</td>
<td>feder control</td>
</tr>
<tr>
<td>color peopl</td>
<td>properti right</td>
</tr>
<tr>
<td>suprem court</td>
<td>illeg immigr</td>
</tr>
<tr>
<td>central govern</td>
<td>central govern</td>
</tr>
<tr>
<td>right commiss</td>
<td>attorney fee</td>
</tr>
<tr>
<td>court unit</td>
<td>thoma jefferson</td>
</tr>
<tr>
<td>peopl south</td>
<td>district columbia</td>
</tr>
<tr>
<td>district columbia</td>
<td>local level</td>
</tr>
<tr>
<td>crimin contempt</td>
<td>school board</td>
</tr>
<tr>
<td>system govern</td>
<td>govern power</td>
</tr>
</tbody>
</table>

Note: This figure compares the individual phrases that contribute the most to the gap in pro-segregation language for Southerns vs. Non-Southerners in racial debates between 1947-1967 and the ones that contribute the most to the gap in pro-segregation language for Republicans vs. Democrats in economic debates between 1967-2017. See details of calculation in Section 4.
Figure 2: Trends in Use of Individual Pro-Segregation Phrases Over Time

Note: For selected pro-segregation phrases listed in Table 3, this figure plots the relative use of each phrase in economic speeches over time (normalized to be proportions that sum to 1 for each term of congress).
Figure 3: Net Use of Pro-Jim Crow Language (Relative to Anti-Jim Crow Language) by Republicans Across Topics (Compared to all Economic Speeches)

Note: This figure measures the net usage of pro-segregation phrases (relative to anti-segregation phrases) in speeches across a variety of topics (as identified through the topic model described in section 3) for Republican speakers. For each figure, the line in grey tracks the use of pro-segregation language in economic speeches aggregated as described in Figure 1. The other line in color indicates the trend in pro-segregation language for the corresponding topic. An observation is at the speech level. For each speech the level of pro-segregation rhetoric is calculated according to equation 3. For each 2-word phrase, the frequency of that phrase in the speech is calculated (as a proportion of total words) and then multiplied by the coefficient that measures its association with pro-segregation politicians in Jim Crow era racial speeches as estimated in equation 3. Then, I sum across all phrases in the speech. A more positive number indications more intense usage of pro-segregation phrases (net of anti-segregation phrases). For each 2-year term of congress I calculate the average pro-segregation score of all speeches and plot these scores over time. Standard error bars for a 5% level of statistical significance are included.
Figure 4: Net Use of Pro-Jim Crow Language (Relative to Anti-Jim Crow Language) in Speeches about Economic Issues vs. Explicitly Racial Issues by Southerners, 1947-1967 vs. Republicans, 1967-2017

Note: This figure measures the usage of pro-segregation phrases (net of anti-segregation phrases) in speeches about economic issues (as described in Figure 1) vs. explicitly racial debates (as identified through the same topic model) for Southern members of Congress form 1947-1967 and for Republican members of Congress from 1967-2017. An observation is at the speech level. For each speech the level of pro-segregation rhetoric is calculated according to equation 3. For each 2-word phrase, the frequency of that phrase in the speech is calculated (as a proportion of total words) and then multiplied by the coefficient that measures its association with pro-segregation politicians in Jim Crow era racial speeches as estimated in equation 2. Then, I sum across all phrases in the speech. A more positive number indications more intense usage of pro-segregation phrases (net of anti-segregation phrases). For each 2-year term of congress I calculate the average pro-segregation score of all speeches and plot these scores over time. Standard error bars for a 5% level of statistical significance are included.
9. Appendix

9.1 Appendix A: Quantifying Voting Record and Realignment

Change in conservatism of voting records by politicians over time can be measured quantitatively through the analysis of roll call votes. I use the Nominate scaling method. This method places legislators on a left-right ideological spectrum based on their roll call votes in Congress\(^{23}\) where a position farther to the right on the spectrum indicates a more conservative voting record. The model assumes each member of congress has an ideal point along an ideological spectrum. For each bill they vote on, the model also assumes that each outcome (‘yes’ vs. ‘no’ on passing the bill) also has a position in that ideological space to measure how conservative or liberal each bill is. The model then uses iterative maximum likelihood estimation to estimate ideal points for legislators and bills in order to rationalize the observed pattern of roll call votes.\(^{24}\) I estimate conservatism of voting record separately for explicitly racial issues (civil rights, desegregation, busing, affirmative action, voting rights) and economic issues. For compactness, I show voting record changes over time for just the House, but trends in the Senate are very similar.

Figures C13 and C14 show the evolution of economic and racial conservatism respectively for Northern Republicans, Northern Democrats, and Southerners (of either party) in Congress. Prior to the 1930s, Southerners are just as progressive on economic issues as Northern Democrats. But over the course of the Century, they gradually move toward the economic conservatism favored by Northern Republicans. Meanwhile, the story is largely reversed for racial ideology. Southerners are by far the most racially conservative politicians in the early decades, but over time they moderated their racial voting record considerably. On the other hand, Northern Republicans move in a more racially conservative direction on explicitly racial issues, converging toward the Southern position.\(^{25}\) Still, both Southerners and Northern Republicans in the 21st Century are less racially conservative than the arch-segregationists of the Jim Crow Era. The overall picture is that conservatism generally has become more extreme on economic issues and less extreme on explicitly racial issues. And Southern politicians in particular have largely abandoned explicitly racist politics in favor of economic

\(^{23}\)Roll call vote data comes from the Voteview database (Lewis, Poole, Rosenthal, Boche, Rudkin, and Sonne 2018) which includes the universe of all roll call votes on the floor of Congress classified by issue area https://voteview.com/

\(^{24}\)For more detail on the estimation of ideology see Poole (2005).

\(^{25}\)For both economic and racial ideology, the gap between Southerners and Northern Republicans in later years reflects the fact that some districts in the South (mostly with large black populations or Northern liberal transplants) elect Democrats with liberal voting records on economic and racial issues.
conservatism.

9.2 Appendix B: Language Pre-Processing

I get rid of common "stop" words like 'the', 'and' and 'with' that communicate little substance but substantially increase the number of phrases based on the list provided by Fox (1989). I reduce words to their stems (so, for example, "runs", "running", and 'run' are all reduced to 'run") based on the algorithm of Porter (1980). I convert all letters to lowercase. Finally, I want to identify language that is used in the modern era and in the past, so I include only phrases that are used at least 5 times on average per term of congress from 1947-1967 and 1967-2011 and by at least 2 speakers on average per term of Congress in both periods. This gets rid of typos and very uncommon words, as well as words that are really only used in the past or present.
9.3 Appendix C: Additional Tables and Figures

Figure C1: Net Use of Pro-Jim Crow Language (Relative to Anti-Jim Crow Language) for Speeches about Tax and Budget

Note: This figure measures the usage of pro-segregation phrases (net of anti-segregation phrases) in speeches about taxes and budget issues (as identified through the topic model described in section 3). An observation is at the speech level. For each speech the level of pro-segregation rhetoric is calculated according to equation 3. For each 2-word phrase, the frequency of that phrase in the speech is calculated (as a proportion of total words) and then multiplied by the coefficient that measures its association with pro-segregation politicians in Jim Crow era racial speeches as estimated in equation 2. Then, I sum across all phrases in the speech. A more positive number indicates more intense usage of pro-segregation phrases (net of anti-segregation phrases). For each 2-year term of congress I calculate the average pro-segregation score of all speeches about taxes and budget by Democrats and Republicans and plot these scores over time. Standard error bars for a 5% level of statistical significance are included.
Figure C2: Net Use of Pro-Jim Crow Language (Relative to Anti-Jim Crow Language) for Speeches about Pensions/Retirement

Note: This figure measures the usage of pro-segregation phrases (net of anti-segregation phrases) in speeches about retirement issues (as identified through the topic model described in section 3). An observation is at the speech level. For each speech the level of pro-segregation rhetoric is calculated according to equation 3. For each 2-word phrase, the frequency of that phrase in the speech is calculated (as a proportion of total words) and then multiplied by the coefficient that measures its association with pro-segregation politicians in Jim Crow era racial speeches as estimated in equation 2. Then, I sum across all phrases in the speech. A more positive number indications more intense usage of pro-segregation phrases (net of anti-segregation phrases). For each 2-year term of congress I calculate the average pro-segregation score of all speeches about entitlements by Democrats and Republicans and plot these scores over time. Standard error bars for a 5% level of statistical significance are included.
Figure C3: Net Use of Pro-Jim Crow Language (Relative to Anti-Jim Crow Language) for Speeches about Health Care

Note: This figure measures the usage of pro-segregation phrases (net of anti-segregation phrases) in speeches about healthcare issues (as identified through the topic model described in section 3). An observation is at the speech level. For each speech the level of pro-segregation rhetoric is calculated according to equation 3. For each 2-word phrase, the frequency of that phrase in the speech is calculated (as a proportion of total words) and then multiplied by the coefficient that measures its association with pro-segregation politicians in Jim Crow era racial speeches as estimated in equation 2. Then, I sum across all phrases in the speech. A more positive number indications more intense usage of pro-segregation phrases (net of anti-segregation phrases). For each 2-year term of congress I calculate the average pro-segregation score of all speeches about healthcare by Democrats and Republicans and plot these scores over time. Standard error bars for a 5% level of statistical significance are included.
Figure C4: Net Use of Pro-Jim Crow Language (Relative to Anti-Jim Crow Language) for Speeches about Labor

Note: This figure measures the usage of pro-segregation phrases (net of anti-segregation phrases) in speeches about labor issues (as identified through the topic model described in section 3). An observation is at the speech level. For each speech the level of pro-segregation rhetoric is calculated according to equation 3. For each 2-word phrase, the frequency of that phrase in the speech is calculated (as a proportion of total words) and then multiplied by the coefficient that measures its association with pro-segregation politicians in Jim Crow era racial speeches as estimated in equation 2. Then, I sum across all phrases in the speech. A more positive number indications more intense usage of pro-segregation phrases (net of anti-segregation phrases). For each 2-year term of congress I calculate the average pro-segregation score of all speeches about labor by Democrats and Republicans and plot these scores over time. Standard error bars for a 5% level of statistical significance are included.
Figure C5: Net Use of Pro-Jim Crow Language (Relative to Anti-Jim Crow Language) for Speeches about Crime

Note: This figure measures the usage of pro-segregation phrases (net of anti-segregation phrases) in speeches about crime issues (as identified through the topic model described in section 3). An observation is at the speech level. For each speech the level of pro-segregation rhetoric is calculated according to equation 3. For each 2-word phrase, the frequency of that phrase in the speech is calculated (as a proportion of total words) and then multiplied by the coefficient that measures its association with pro-segregation politicians in Jim Crow era racial speeches as estimated in equation 2. Then, I sum across all phrases in the speech. A more positive number indications more intense usage of pro-segregation phrases (net of anti-segregation phrases). For each 2-year term of congress I calculate the average pro-segregation score of all speeches by Democrats and Republicans and plot these scores over time. Standard error bars for a 5% level of statistical significance are included.
Figure C6: Net Use of Pro-Jim Crow Language (Relative to Anti-Jim Crow Language) for Speeches about Education

Note: This figure measures the usage of pro-segregation phrases (net of anti-segregation phrases) in speeches about education issues (as identified through the topic model described in section 3). An observation is at the speech level. For each speech the level of pro-segregation rhetoric is calculated according to equation 3. For each 2-word phrase, the frequency of that phrase in the speech is calculated (as a proportion of total words) and then multiplied by the coefficient that measures its association with pro-segregation politicians in Jim Crow era racial speeches as estimated in equation 2. Then, I sum across all phrases in the speech. A more positive number indications more intense usage of pro-segregation phrases (net of anti-segregation phrases). For each 2-year term of congress I calculate the average pro-segregation score of all speeches by Democrats and Republicans and plot these scores over time. Standard error bars for a 5% level of statistical significance are included.
Figure C7: Net Use of Pro-Jim Crow Language (Relative to Anti-Jim Crow Language) for Speeches about Drugs

Note: This figure measures the usage of pro-segregation phrases (net of anti-segregation phrases) in speeches about education issues (as identified through the topic model described in section 3). An observation is at the speech level. For each speech the level of pro-segregation rhetoric is calculated according to equation 3. For each 2-word phrase, the frequency of that phrase in the speech is calculated (as a proportion of total words) and then multiplied by the coefficient that measures its association with pro-segregation politicians in Jim Crow era racial speeches as estimated in equation 2. Then, I sum across all phrases in the speech. A more positive number indications more intense usage of pro-segregation phrases (net of anti-segregation phrases). For each 2-year term of congress I calculate the average pro-segregation score of all speeches by Democrats and Republicans and plot these scores over time. Standard error bars for a 5% level of statistical significance are included.
Figure C8: Net Use of Pro-Jim Crow Language (Relative to Anti-Jim Crow Language) for Speeches about Anti-Poverty Programs

Note: This figure measures the usage of pro-segregation phrases (net of anti-segregation phrases) in speeches about issues related to anti-poverty program and some mentions of housing policy (as identified through the topic model described in section 3). An observation is at the speech level. For each speech the level of pro-segregation rhetoric is calculated according to equation 3. For each 2-word phrase, the frequency of that phrase in the speech is calculated (as a proportion of total words) and then multiplied by the coefficient that measures its association with pro-segregation politicians in Jim Crow era racial speeches as estimated in equation 2. Then, I sum across all phrases in the speech. A more positive number indications more intense usage of pro-segregation phrases (net of anti-segregation phrases). For each 2-year term of congress I calculate the average pro-segregation score of all speeches by Democrats and Republicans and plot these scores over time. Standard error bars for a 5% level of statistical significance are included.
Figure C9: Net Use of Pro-Jim Crow Language (Relative to Anti-Jim Crow Language) for Speeches about Energy

Note: This figure measures the usage of pro-segregation phrases (net of anti-segregation phrases) in speeches about education issues (as identified through the topic model described in section 3). An observation is at the speech level. For each speech the level of pro-segregation rhetoric is calculated according to equation 3. For each 2-word phrase, the frequency of that phrase in the speech is calculated (as a proportion of total words) and then multiplied by the coefficient that measures its association with pro-segregation politicians in Jim Crow era racial speeches as estimated in equation 2. Then, I sum across all phrases in the speech. A more positive number indications more intense usage of pro-segregation phrases (net of anti-segregation phrases). For each 2-year term of congress I calculate the average pro-segregation score of all speeches by Democrats and Republicans and plot these scores over time. Standard error bars for a 5% level of statistical significance are included.
Figure C10: Net Use of Pro-Jim Crow Language (Relative to Anti-Jim Crow Language) for Speeches about International Trade

Note: This figure measures the usage of pro-segregation phrases (net of anti-segregation phrases) in speeches about education issues (as identified through the topic model described in section 3). An observation is at the speech level. For each speech the level of pro-segregation rhetoric is calculated according to equation 3. For each 2-word phrase, the frequency of that phrase in the speech is calculated (as a proportion of total words) and then multiplied by the coefficient that measures its association with pro-segregation politicians in Jim Crow era racial speeches as estimated in equation 2. Then, I sum across all phrases in the speech. A more positive number indicates more intense usage of pro-segregation phrases (net of anti-segregation phrases). For each 2-year term of congress I calculate the average pro-segregation score of all speeches by Democrats and Republicans and plot these scores over time. Standard error bars for a 5% level of statistical significance are included.
Figure C11: Net Use of Pro-Jim Crow Language (Relative to Anti-Jim Crow Language) for Speeches about Economic Issues, Excluding the South

Note: This figure measures the usage of pro-segregation phrases (net of anti-segregation phrases) in speeches about economic issues (as identified through the topic model described in section 3 where a speech is counted as about "economic" issues if at least 25% of the topic proportions assigned by the topic model are counted as economic in nature; e.g. keywords related to taxes, labor, health care etc.) for only non-Southern politicians (where 'Southern' is defined as congresspeople from the former Confederacy). An observation is at the speech level. For each speech the level of pro-segregation rhetoric is calculated according to equation 3. For each 2-word phrase, the frequency of that phrase in the speech is calculated (as a proportion of total words) and then multiplied by the coefficient that measures its association with pro-segregation politicians in Jim Crow era racial speeches as estimated in equation 2. Then, I sum across all phrases in the speech. A more positive number indications more intense usage of pro-segregation phrases (net of anti-segregation phrases). For each 2-year term of congress I calculate the average pro-segregation score of all speeches by Democrats and Republicans and plot these scores over time. Standard error bars for a 5% level of statistical significance are included.
Figure C12: Correlation between Pro-Jim Crow Language in Speeches about Economic Issues and Economic Conservatism of Voting Record

Note: This figure measures the correlation at the speech level between the use of pro-segregation language and the conservatism of a speaker's economic voting record as measured by W-Nominate scores estimated on economic roll call votes. A correlation is calculated for all speeches in a term of Congress and then plotted over time.
Figure C13: Trend In Economic Ideology

Note: this figure measures changes in the average economic ideology of politicians across three groups of politicians: Northern Republican, Northern Democrats, and Southern politicians of either party. Ideology is measured based on voting behavior over roll call votes. Economic ideology is estimated across all roll call votes related to taxes, budgets and appropriations, and redistribution. See Section 2 and Poole (2005) for more detail. The unit of observation is a politician over the course of their career. So, changes over time reflect changes in the cohort of politicians.
Note: this figure measures changes in the average racial ideology of politicians across three groups of politicians: Northern Republican, Northern Democrats, and Southern politicians of either party. Ideology is measured based on voting behavior over roll call votes. Racial ideology is estimated across all roll call votes related to segregation, voting rights for Black people, busing, and affirmative action. See Section 2 and Poole (2005) for more detail. The unit of observation is a politician over the course of their career. So, changes over time reflect changes in the cohort of politicians.
Note: this figure measures the frequency of two-word phrases in Google Books data of the form 'race X' where X is a word corresponding to the topics listed in Figure 3 and discussed in Section 4 as being more or less racialized.