

# Firm turnover and the return of racial establishment segregation

John-Paul Ferguson  
Stanford University  
jpferg@stanford.edu

Rembrand Koning  
Harvard Business School  
rem@hbs.edu

January 8, 2018

## Abstract

Racial segregation between American workplaces is greater today than it was a generation ago. This increase has happened alongside the declines in within-establishment occupational segregation on which most prior research has focused. We examine more than 40 years of longitudinal data on the racial employment composition of every large private-sector workplace in the United States to calculate between-establishment and within-establishment trends in racial employment segregation over time. We demonstrate that the return of racial establishment segregation owes little to within-establishment processes but rather stems from differences in the turnover rates of more- and less-homogeneous workplaces. Present research on employment segregation focuses intently on within-firm processes. By doing so, we may be overstating what progress has been made on employment integration and ignoring other avenues of intervention that may give greater leverage for further integrating firms.

## Introduction

Racial segregation between American workplaces is greater today than it was a generation ago. Over that same generation, the American workforce has steadily grown more racially diverse. Our goals in this paper are to document this trend, which on its face seems paradoxical; describe some of the organizational processes which have produced it; and discuss its implications for research on stratification and for attempts to further integrate workplaces.

Racial segregation in the American workplace declined, at shifting and uneven rates, from the 1960s through the mid-1980s (Tomaskovic-Devey et al., 2006). The largest declines happened in the 1970s; progress then slowed or stalled in many industries (Cohen et al., 2009). Social scientists have spent years trying to understand why integration seemed to hit a wall after an early period of substantial (though by no means easy) progress, and evaluating policies that might push it farther along (Reskin, 2000; Bielby, 2000; Kaufman, 2010).

Employment segregation has several dimensions. Most research concentrates on occupational segregation or on the diversity of managerial personnel. There are empirical and methodological reasons for such focus, but it also reflects an assumption that workplaces as a whole have become more racially diverse. If they have, then it makes sense to ask whether minority workers have truly become integrated or whether they instead are concentrated in entry-level, non-supervisory, or token positions (Kalev et al., 2006; Dezsó et al., 2016). We reaffirm here that establishment-level racial diversity has increased for a subset of workplaces, namely, those dominated by white workers. Yet the increased presence of visible minorities in once overwhelmingly white workplaces exists amidst decreased racial diversity across establishments as a whole.

This article advances scholarship on employment segregation by documenting changes in *between-establishment* segregation and by highlighting how *firm turnover* drives this increase in segregation even as incumbent workplaces have become increasingly diverse. We should clarify at the start that increases in workplace segregation, coming after years of decreases, do not simply mean that the integration of the 1970s and early 1980s was reversed—that workplaces started expelling minority workers in favor of whites. Rather, the differential birth and death rates of organizations, combined with within-workplace mechanisms like ethnic succession, appear to generate substantial between-workplace segregation. Indeed we think that a major contribution of this study is to emphasize that within-firm processes, such as hiring and promotion, are only one channel for affecting workforce composition (Petersen and Saporta, 2004; Fernandez and Fernandez-Mateo, 2006; Kalev et al., 2006; Castilla, 2008). The attention that that channel has received notwithstanding, over the past several decades, the turnover of organizations appears to have washed away progress on racial employment integration faster than such processes have built it up.

To measure workplace segregation, we leverage the population of EEO-1 establishment surveys gathered by the Equal Employment Opportunity Commission (EEOC) over more than four decades, from the early 1970s until recently. These surveys provide annual data on workforce racial composition for every large private-sector establishment in the U.S. economy (Robinson et al., 2005). We explore racial employment segregation through decomposition of the Theil statistic, which lets us use a multigroup rather than a dichotomous measure (Reardon and Firebaugh, 2002) and also permits straightforward decompositions to evaluate the relative contributions of different factors to overall levels of segregation. This type of decomposition is necessary to describe, explain, and evaluate countervailing trends like those we explore here.

We present our case as follows. First we review the state of knowledge on employment segregation. Our knowledge of trends in between-establishment racial segregation is particularly thin. We review possible mechanisms that could produce increasing between-establishment segregation, both mechanically and substantively. Next we discuss our measurement strategy, and the advantages of using a Theil statistic to track segregation. We present trends in the Theil statistic of employment segregation over time, and decompose the Theil to show the relative contribution of mechanisms like establishment turnover and within-establishment change. We demonstrate that the trends we observe are compatible with findings from earlier work, such as that occupational segregation by race has fallen drastically over the last forty years (Tomaskovic-Devey et al., 2006). This implies that the within-firm measures used in earlier work are necessary but insufficient for characterizing the progress of integration.

We demonstrate that racial segregation has increased between establishments; we also show that this increase has not been driven by increasing segregation *within* establishments. Within-cohort establishment segregation almost always decreases, albeit at decreasing rates. At the same time, though, establishments entering the population tend to be less diverse than exiting ones. The net increase contributed by turnover more than offsets within-firm decreases. And this is a between-*firm* story, not simply an establishment one: we show that increase in segregation between establishments has taken place between firms, rather than between establishments within larger firms.

We also confirm that these differences between new and existing workplaces do not simply reflect changes in the underlying demographics of the workforce. The share of workplaces with majority-minority employment has risen for years, roughly on track with minorities' growing share of the population. This increase in majority-minority establishments, combined with population change, contributes more to between-establishment segregation than the entry of minorities into once all- or heavily white establishments.

In short, we describe complex dynamics that have stopped or reversed a significant chunk of

the progress made on racial employment integration in the wake of the Civil Rights movement. But the complexity of these dynamics should not obscure the simplicity of the dilemma. We value integration as a social goal. The typical American adult spends a substantial chunk of their waking hours at work. If contact across racial and ethnic divides is a precondition for healthy civic life in a diverse society (Allport, 1954), then we must be concerned with the racial diversity of our workplaces. Furthermore, we know that between-firm differences within industries have been a major driver of the growth in wage inequality in the last generation (Song et al., 2015; Autor et al., 2017; Cobb and Lin, 2017). It behooves us to explore whether those inequalities reflect racial segregation. The extensive research on diversity policies attests to these broader concerns about the progress of integration. Yet our progress has been less than we often think it has. In particular, changes to the population of establishments, above the level and control of the individual establishment, can swamp integration efforts within establishments. Research on employment segregation should also train its sights on population-level process.

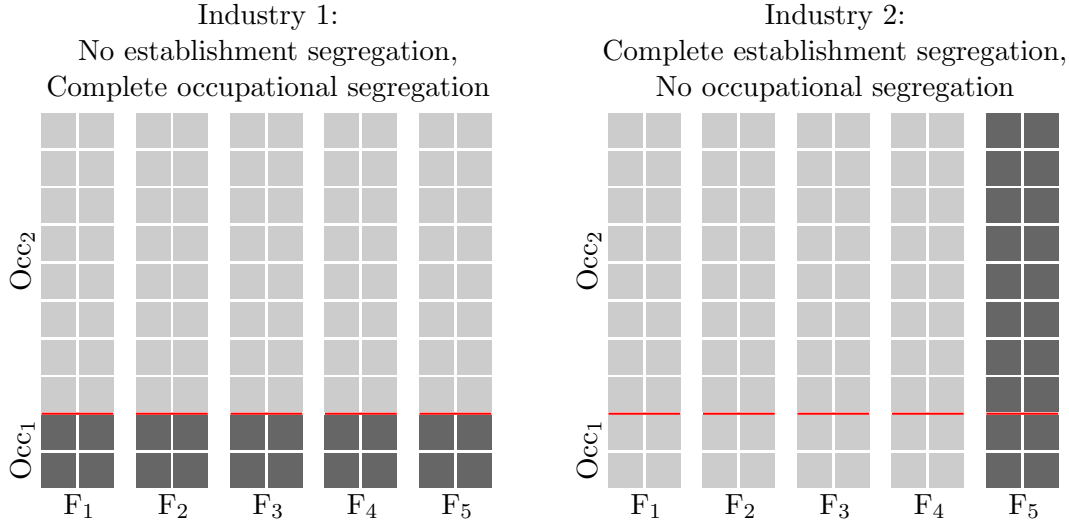
## Trends in employment segregation

Data used in employment-segregation research take two main forms. Most common are individual- or household-level surveys, such as the Current Population Survey or the American Community Survey. An advantage of these surveys is that they have detailed information on the individual respondents. This allows researchers to partial out features like education or language skills that may affect segregation yet covary with race (Reskin et al., 1999; Alonso-Villar et al., 2012). Their disadvantage is that respondents cannot be linked to the places where they work. This research therefore focuses on segregation by occupation, rather than by establishment (Cain, 1986; King, 1992; Hirsch and Macpherson, 2004). There are conceptual grounds for studying occupational segregation, not least that we would not call integrated a workplace where minorities were restricted to low-level jobs. Yet the disproportionate number of studies that examine occupational segregation also reflects the simple fact that establishment-level data are rare. As a result,

[Nearly] all our systematic knowledge about national *change* in the employment opportunities of women and race/ethnic minorities comes from surveys of individuals that describe employment in terms of occupations detached from their workplace context, the presumed site of discrimination and the target of antidiscrimination legislation (Tomaskovic-Devey et al., 2006, p. 565).

The only nationally representative, establishment-level data on workforce composition available come from the establishment surveys collected annually by the EEOC, which we describe below. Tomaskovic-Devey et al. (2006) used these data to build out the first national time series in segregation trends between 1966 and 2003. Yet that work maintains prior studies' focus on occupational segregation, albeit as a weighted average across establishments rather than across occupational aggregates. Their main finding was that occupational segregation by race had declined in the 1970s and essentially stalled in the 1980s. They note "some disturbing evidence of resegregation after 1995 in old economy sectors" (p. 584), but emphasize that "there has been real workplace-level desegregation since 1964" (p. 565). Stainback and Tomaskovic-Devey (2012) followed up on this work with more detailed analysis of said desegregation, noting that there was significant variance by industry and that increased diversity within workplaces was accompanied by increased over-representation of white males in the most high-status jobs.

A subset of studies, such as Bielby and Baron (1986), Carrington and Troske (1998), and Hellerstein and Neumark (2008), have constructed datasets that record both individual- and establishment-



**Figure 1:** Comparing industries with high (low) occupational and low (high) establishment segregation. Each hypothetical industry has the same composition: 80% majority (light gray) and 20% minority (dark gray).

level characteristics. In studying the sorting of workers into workplaces rather than into occupations, these works reach different conclusions about the progress of integration. Hellerstein et al. (2008) in particular provide “extremely useful initial evidence” that, across establishments, “segregation by race has increased substantially” (Bender et al., 2008, p. 6) between 1990 and 2000. Unfortunately, most of these studies rely on idiosyncratic settings or sampling methods that restrict the generalizations that can be made. Hellerstein et al. (2008) build a sample that is broadly representative of US employers, but their cross-sectional data cannot speak to trends over time.

Establishment and occupational segregation are both important; the two measures capture different ideas and can vary independently of one another. Each asks how many people we would have to transfer across sub-units (establishment or occupation) to make those sub-units representative of the larger labor market. In this light, consider the two stylized industries in figure 1. In each, the minority share of the workforce (dark gray) is 20 percent. In industry 1, the majority and minority are proportionately represented in establishments that have complete occupational segregation. In industry 2, the majority and minority are proportionately represented in occupations that are completely separated between establishments.<sup>1</sup> In one case, two races can be in frequent contact albeit within relationships of dominance and subordination; in the other, the two races can hold comparable positions in distinct hierarchies. Changing one of these measure does not necessarily change the other.

Although the two measures vary independently, when racial occupational segregation is calculated on establishment-level data, it can be treated as nested within between-establishment segregation. A measure of occupational segregation like the index of dissimilarity (Duncan and Duncan, 1955) indexes the reallocation of group members across sub-units within a larger whole; it takes the marginal distribution of groups across the whole for granted (Grusky and Charles, 1998). When calculated on occupations within establishments, such measures implicitly take the distribution of races *between* establishments as given. If we presume that establishment segregation is

<sup>1</sup>Scholars of residential segregation will recognize that these two patterns correspond to status and residential inequality, such as characterized southern and northern patterns of segregation after the Civil War (Grigoryeva and Ruef, 2015).

steadily declining then it makes sense to look at the distribution of races within occupations within each establishment. Given how little we know about trends in between-establishment segregation, though, it would be better to use a measure that can simultaneously and separately account for both between and within segregation.

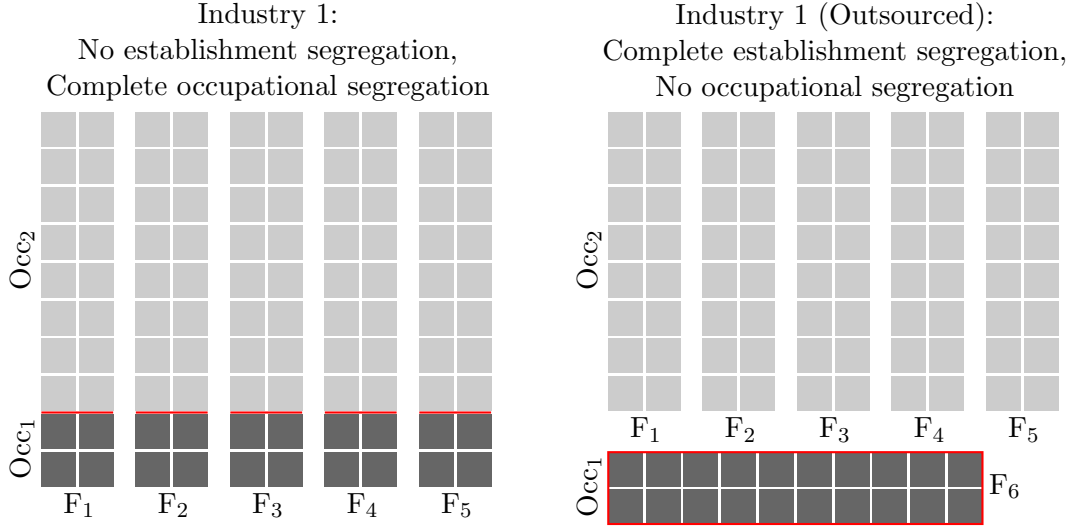
## Drivers of between-establishment segregation

Before we discuss the details of measurement, it is worth laying out some possible sources of between-establishment segregation (here and below, we will often refer to “establishment segregation” for brevity). This is important because while establishment segregation is measured at the organizational level, it is not necessarily produced there. This is a general issue in stratification research: studies of bias in hiring for example have long noted that hiring managers’ choices are constrained by the composition of their applicant pools, over which the firm has far less control (Fernandez and Fernandez-Mateo, 2006; Fernandez and Friedrich, 2011). We similarly want to consider more distant and proximate drivers of racial sorting between workplaces.

The first source to consider is whether establishment segregation is driven by changes within workplaces over time, or changes in the population of workplaces. Presume that establishment segregation has risen: it is possible that new workplaces are more diverse but that older workplaces have stopped making progress or backslid on integration. It is also possible that existing establishments have integrated but that their progress is more than offset by greater homogeneity in new establishments. (Of course, movements in the same direction are also possible.) It is important to know how much of any change in establishment segregation comes from establishment turnover, because some sources are invisible if we only follow changes within firms. For example, if newer firms start out more homogeneous and then diversify, a purely within-firm analysis will see diversification everywhere without capturing the changes in the base rates caused by changes in the population of firms. This distinction is also particularly important for designing interventions, because trends at the population level limit the effectiveness of policies at the workplace level, and vice versa. Thus, a first-order question is how much weight to place on change within versus turnover between organizations.

Whether establishment segregation trends are a within- or between-firm phenomenon, we want to know how much they owe to things that are feasibly under organizations’ control. Consider demographic changes like increasing racial diversity in the labor force. If this is unevenly distributed geographically then we can observe establishment segregation that really just reflects this distribution. For simplicity: if a firm operates establishments in two labor markets, one of which is all white and one which is all Hispanic, then accounting for the different allocation of races across labor-market areas would explain away the apparent establishment segregation. Such pre-organizational (Petersen and Saporta, 2004) factors can be important in their own right, as evidence of housing discrimination (Aaronson et al., 2017), spatial mismatch (Kain, 1992; Wilson, 1996), and more; but they may also be more appropriately addressed through housing and transportation policy than by focusing on the employing establishments (Fernandez, 2006).

Pre-organizational factors go beyond demography itself. They can include any characteristics that are valuable for job performance yet covary with race. If establishments employ such factors in different proportions, then they will induce a mechanical segregation between establishments by race. We know that, despite documented progress particularly in professional certifications (Stainback and Tomaskovic-Devey, 2012), most occupations have nothing like proportionate representation across racial groups. Inequality in educational attainment also biases the labor supply for many types of jobs (Alonso-Villar et al., 2012). Often many such factors can be correlated by



**Figure 2:** Comparing outsourcing in an industry with high initial occupational and low initial establishment segregation. The workforce composition is the same in each: 80% majority (light gray) and 20% minority (dark gray).

the industry in which an establishment operates, and we know that industries vary widely in their racial employment profiles. For much of the postwar era, durable manufacturing offered relatively plentiful, good jobs for black men (Nelson, 2001; Rosenfeld and Kleykamp, 2012), while contemporary service jobs tend to disproportionately employ non-white women (Catanzarite, 2000; Dwyer, 2013). Again, unequal representation like this may itself be problematic, but it may be ineffective to consider the workplace as critical place of intervention in such cases.

That said, including the industry in which a firm operates in a list of “pre-organizational” sources of establishment segregation hints at a problem with such a distinction. The industries in which a firm operates are to endogenous to strategic and other organizational decisions (Zuckerman, 2000; Kogut et al., 2002; Kim et al., 2013). The same goes for the occupations, and more specifically the jobs, that an establishment employs. This points to a more fine-grained source of establishment segregation: shifts in the boundary of the firm. Again, a stylized example is useful. The left panel of figure 2 reproduces industry 1 from figure 1. The right panel of figure 2 shows the same industry, only now with six firms instead of five. The original five firms have outsourced occupation 1 to firm 6, a specialized contractor. This outsourcing decision has several interesting implications for segregation. First, segregation within and between establishments can change absent changes in the workforce, if firm decisions redefine workplaces. Second, the employees in firm 6 will now likely show up in a different industry, but this was a firm choice; hence our saying that industry is not necessarily a pre-organizational source of establishment segregation. Third, the shift shown in figure 2 both raises segregation between establishments and lowers occupational segregation within them. Within-workplace measures of occupational segregation, considered alone, will pick up on the “improvement” in occupational segregation but not the decline in diversity. Fourth, this type of establishment segregation is potentially more pernicious than that depicted in the right panel of figure 1. Where in that case there was an implicit career ladder between occupations, albeit in segregated firms, in figure 2 racial segregation is combined with reduced opportunity for movement, at least via internal labor markets (Irwin (2017) offers a recent example, comparing the careers of janitors at Kodak in the 1980s versus those at the contractor Apple uses today).

We do not have the necessary data to evaluate all of these sources of establishment segregation.

We only have relatively broad occupational information, so we cannot track changes in the mix of jobs within establishments. Thus we cannot directly measure outsourcing with our data alone since even when a firm outsources jobs, it would be unlikely to outsource the entirety of one of the broad occupational categories in our data. We do not have detailed worker-level characteristics beyond race and sex; thus we cannot separately evaluate changes in the educational mix of jobs across establishments. Our data do let us test some of these sources though, and thus inform our analytic strategy here. First we introduce our metric for segregation, the Theil statistic, and explain how its decomposability lets us partial out and evaluate the relative contributions of factors like establishment turnover, demographic differences between labor markets, and the like to establishment segregation. Next we describe the data. Then we present results from a series of decompositions of the Theil statistic, to evaluate these sources. Our goal with these analyses is to document that racial segregation between establishments has risen; that that rise is due to the turnover of establishments rather than “re-segregation” within establishments; and that it does not mechanically reflect shifts in pre-organizational sources like labor-market demography or industry mix. Organizations play a role here; and we conclude by discussing ways that future work could explore that role in more detail.

## Measuring and decomposing segregation with the Theil statistic

Segregation is a multi-dimensional concept. In focusing on establishment segregation, we focus on the dimension that Massey and Denton (1988) called “evenness”: how proportionately members of non-overlapping groups are spread across units. Massey and Denton (1988) defined multiple dimensions to recognize that integration could be advanced in different ways. Groups could be proportionately allocated across units and still have no *exposure* to one another; groups could be more or less *concentrated* on a spatial dimension that is uncorrelated with the analytic units; and so on. When the units are physical establishments, as here, unevenness also reflects segregation on several other dimensions by construction. Unless we note otherwise, the groups in our analyses are races; the units, establishments.

Researchers have used many measures of inequality to study segregation. The most frequently deployed are the index of dissimilarity (Duncan and Duncan, 1955) and the Gini coefficient (Gini, 1921); others include the relative diversity index, the squared coefficient of variance, and the normalized exposure index. These measures have various strengths, but they all suffer from at least one of two major shortcomings. First, many of these measures work well for studying segregation between two groups but do not generalize well to a multigroup context. The most frequent issue has to do with “perverse transfers”: moving a group member from a sub-unit where the group is over-represented to one where it is under-represented should always lower segregation, yet some measures (including the index of dissimilarity) can be raised by such transfers. Second, most of these measures have no straightforward decomposition. This makes it difficult to impossible to discuss the relative contribution of different sub-units to the overall level of measured segregation. In their review and formal analysis of multigroup measures of segregation, Reardon and Firebaugh (2002) demonstrated that only the Theil (1972) statistic behaves properly in a multigroup context while being analytically decomposable. The Theil’s decomposability is useful for comparing the relative contribution of different clusters of establishments, such as industries or labor markets, to overall segregation. It is also decomposable by groups rather than by units (Reardon and Firebaugh, 2002). Because of its facility with handling America’s changing racial landscape and its analytic decomposability, the Theil statistic has drawn increasing interest among scholars of inequality (e.g., Chetty et al., 2014).

Theil’s statistic leverages an entropy index. Intuitively, it tells us how much new information we gain by knowing the sub-unit, relative to just knowing about the population as a whole. Picture an economy where half of the workers are white and half are black. If you were asked to guess the race of a randomly chosen worker, you might as well flip a coin. Now presume you were told where that worker worked. How would that affect your uncertainty? If there is no between-establishment segregation—if, in every workplace, 50 percent of the workers are white—then knowing the workplace gives no useful information. Your uncertainty would not be reduced at all; accordingly, a Theil statistic calculated on this economy would be zero. By contrast, if there were complete segregation between establishments by race, then learning the workplace completely eliminates your uncertainty about a given worker’s race. A Theil statistic calculated on this economy would be maximized. In this sense, the Theil statistic follows the same logic as contingency tables or  $\chi^2$  tests, where greater values imply that conditionals are more informative than marginals, and its decomposability offers the same analytic leverage that relative contributions to a  $\chi^2$  statistic do.

Say that we want to calculate a Theil statistic for racial segregation between  $J$  establishments within a labor market. We write this as  $H_j$ . Let  $\phi_r$  be race  $r$ ’s share of the population in the area. The entropy index for the area is  $E = \sum_r \phi_r \log_2 \frac{1}{\phi_r}$ ; <sup>2</sup> we define  $\phi_r \log_2 \frac{1}{\phi_r} = 0$  when  $\phi_r = 0$ . If  $j \in J$  is the establishment then its entropy index is  $E_j = \sum_r \phi_{rj} \log_2 \frac{1}{\phi_{rj}}$ . We can then sum across establishments to calculate the Theil statistic  $H_j$  for the area:

$$H_j = \sum_j \frac{w_j}{w} \frac{E - E_j}{E}, \quad (1)$$

where  $w_j$  and  $w$  represent the number of workers in establishment  $j$  and the total workforce in the area. Thus  $H_j$  is the weighted average of deviations of establishments’ entropies from the larger labor market’s entropy. A group of all-white firms, for example, should have an effective segregation of zero if they are based in an all-white labor market. <sup>3</sup>

The Theil statistic can be gracefully extended to measure segregation across multiple labor markets. Say we have  $A$  labor-market areas and  $J$  establishments in each labor market. Again, we would like to measure racial segregation across establishments. Unlike the example above, now there are two sources that contribute to overall segregation. The first source is the uneven distribution of races between labor markets. We want to account for this separately, because if we do not then we will confound differences in the pre-organizational geographic distribution of the labor force with organizational differences. <sup>4</sup> Building on the logic above, we can calculate a Theil statistic  $H_a$  for between-market segregation; this is the weighted average of deviations of labor markets’ entropies from the overall population’s entropy. The second source is the uneven distribution of races within labor markets, between establishments. For each labor market  $a$ , we can calculate this as we did  $H_j$  above; we denote these  $H_j^a$ . We then sum across the workforce-weighted contributions to calculate an overall labor-market/establishment Theil statistic,  $H_{aj}$ :

<sup>2</sup>In a homogeneous area, there is no uncertainty about a worker’s race and thus no entropy, and  $E = 0$ . Entropy increases as the number of races increases and as those races’ shares of the population become more equal. Conceptually, then, the entropy index works like a Blau or Herfindahl index. Its advantages inhere in how it can be formally decomposed. The use of base 2 for the logarithm is a legacy of the focus on binary signals by Shannon (1948). The choice of base affects the (arbitrary) units of the index but not the relative relationships among sub-units or trends over time.

<sup>3</sup>An all-white labor market, in a country as diverse as the United States, can imply segregation on different dimensions. We do not dispute the point but here take labor-market composition as given.

<sup>4</sup>If most Latinos live in the American southwest, for example, then we will see many Latino workers in southwestern organizations and few Latinos in organizations elsewhere. But since someone cannot really work in another part of the country without moving, it would be conceptually wrong to treat this as segregation between establishments rather than between regions.



$$\begin{aligned}
H_{aj} &= \underbrace{\sum_a \frac{w_a}{w} \frac{E - E_a}{E}}_{\text{Between area}} + \underbrace{\sum_a \frac{w_a}{w} \left( \sum_{j \in a} \frac{w_j}{w_a} \frac{(E_a - E_{aj})}{E_a} \right)}_{\text{Within area, between establishment}} \\
&= H_a + \sum_a \frac{w_a}{w} (H_j^a) \\
&= H_a + \bar{H}_j^a
\end{aligned} \tag{2}$$

where  $w_a$  and  $E_a$  are the workforce and entropy measures in each labor market and  $w$  and  $E$  are the workforce and entropy measures for the entire population. Notice that while  $\bar{H}_j^a$  is a weighted average across labor markets, it captures the average between-establishment segregation within labor markets. Each within-area part  $H_j^a$  shows how much total segregation would be reduced if segregation in labor market  $a$  were eliminated (i.e., if workers of different races were allocated to establishments  $j \in a$  proportionally to their total employment in  $a$ ) while leaving all the other labor markets unchanged. The between-area part  $H_a$  shows how much total segregation would be reduced by allocating workers to labor markets proportionally to their shares in the total population, while leaving establishment segregation in each area unchanged (Reardon and Firebaugh, 2002). From a policy perspective, if  $H_a$  is significantly smaller than  $\bar{H}_j^a$  then reassigning workers across establishments, but within the same labor market, could greatly reduce racial segregation in US workplaces; if  $\bar{H}_j^a$  is significantly smaller than  $H_a$  then between-establishment segregation mostly reflects geographic population differences rather than organizational outcomes.

This point can be generalized: the Theil statistic can decompose segregation into within and between components for an arbitrary number of nested groups (Cowell, 1985).<sup>5</sup> For example, one could calculate a Theil measure that accounts for segregation between labor markets, between industries within each labor market, between establishments within labor market/industries, and between occupations within labor market/industry/establishments. For an arbitrary number of nested groups  $x \supset y \supset z \dots$  we can calculate the Theil statistic as follows:

$$H_{xyz\dots} = H_x + \bar{H}_y^x + \bar{H}_z^{xy} + \dots \tag{3}$$

For any nested levels  $x \supset y$ ,

$$\begin{aligned}
H_y^x &= \sum_{y \in x} \frac{w_y}{w_x} \frac{E_x - E_y}{E_x} \\
\bar{H}_y^x &= \sum_x \frac{w_x}{w} H_y^x
\end{aligned} \tag{4}$$

Notice that the subscripts indicate the level of segregation that each term captures. Superscripts indicate the *within* level of segregation, subscripts the *between* level. Multiple superscripts show

---

<sup>5</sup>It is worth noting that the Theil statistic is not top-bounded. Theil argued that his index's not having a fixed upper bound but instead being influenced by population size and the level of granularity was conceptually desirable. If we have two societies where one person has all the income, but the first society has ten people and the second has 1,000, a statistic like the Gini coefficient would take the same maximum value. Yet most of us would argue that the latter is more unequal. Similarly, inequality should increase if we add in inequality within sub-units, within sub-sub-units *et cetera* (Conceição and Galbraith, 1998). Adding levels of decomposition—essentially, taking into account more possibilities for micro-segregation—will weakly increase the statistic, such that  $H_{xy} \geq H_x$ . Because the statistic is not top-bounded, the levels and trends of its additive components are invariant to which other levels are considered.

how levels are nested. A bar indicates that the statistic is the weighted average across the relevant units.

Aggregation is performed the same way. Thus we could define total segregation in the United States as the sum of racial segregation between labor market areas  $a$ , racial segregation within areas and between establishments  $j$ , and segregation within establishments and between occupations  $o$ . This is what allows us to examine trends in between-establishment segregation while also tracking the within-establishment occupational segregation that prior work has focused on. We can expand equation 3 to capture these relationships:

$$\begin{aligned}
 H_{ajo} &= H_a + \bar{H}_j^a + \bar{H}_o^{aj} \\
 &= \underbrace{\sum_a \frac{w_a}{w} \frac{E - E_a}{E}}_{\text{Between-area}} + \overbrace{\sum_a \frac{w_a}{w} \left( \sum_j \frac{w_{aj}}{w_a} \frac{E_a - E_{aj}}{E_a} \right)}^{\text{Within-area, between-establishment}} + \underbrace{\sum_j \frac{w_{aj}}{w_a} \left( \sum_o \frac{w_{ajo}}{w_{aj}} \frac{E_{aj} - E_{ajo}}{E_{aj}} \right)}_{\text{Within-establishment, between-occupation}} \quad (5)
 \end{aligned}$$

We go to such lengths to describe these mechanics of the Theil statistic because, below, we represent multiple decompositions of racial employment segregation in the United States. In some of our graphics the same trend line is associated with different values on the  $y$ -axis. This is because these are all additive components of an index that is just the sum of its components. This reflects a benefit of the Theil statistic: because it is the sum of its components, changes in it can be analyzed into the sum of changes in its components. This means that the *trends* in a component are unaffected by which other components are included. We demonstrate examples of this below.

## Data

We study patterns within the EEO-1 establishment surveys filed with the EEOC. This longitudinal data on the sexual and racial composition of large establishments has emerged as a gold standard in organizational research (Leonard, 1985; Robinson et al., 2005; Kalev et al., 2006; Ferguson, 2015). This is one of our motivations for using these data: doing so rules out the possibility that we find different trends than others because of some idiosyncrasy of our data source.

In 1966, to help monitor compliance with the Civil Rights Act, the EEOC began collecting data on the sexual and racial composition of workforces in establishments with 50 or more employees (Executive Order 11246 originally extended this reporting requirement to establishments with 25 or more employees, if those establishments have at least \$50,000 of federal contract work annually). The EEO-1 survey form gathers identifying information for the establishment, such as its location, industry, and (when relevant) its parent firm. The bulk of the form is a matrix of occupational classifications and race/sex combinations, into which employers enter counts of employees.<sup>6</sup> EEO-1 data are unavailable for the years 1967–1970, 1974, and 1976–1977. The long gap between the 1966 and 1971 datasets makes interpolating trends between them suspect, and the 1966 data has scanty information on firm locations below the state level. We therefore conduct analyses on the data between 1971 and 2014.<sup>7</sup> We omit the missing-data years when depicting time trends; when analyzing year-on-year changes, we use data from 1971/2 to 1972/3 and from 1978/9 forward.

<sup>6</sup>A sample EEO-1 form can be viewed at <https://www.eeoc.gov/employers/eeo1survey/upload/eeo1-2-2.pdf>.

<sup>7</sup>Access to the EEO-1 reports is obtained through an Intergovernmental Personnel Act agreement between the first author and the EEOC. Data for this project were first provided in 2015.

In 2007, the EEOC made several important changes to the EEO-1 form and reporting process. They began collecting data once again from establishments whose size is below the mandatory reporting threshold. Participation in this new program is not mandatory for small establishments, however, and so patterns seen in those data are not necessarily representative of national trends. We exclude these smaller establishments from our analyses. The Commission also pioneered an initiative that year to use Dun and Bradstreet data to contact establishments that should have been filing EEO-1 surveys but were not. As a result, the cohort of establishments that enter the data in 2007 is much larger and has a greater proportion of (relatively) smaller, single-establishment firms than other years. As we discuss in more detail in the appendix, the size and unusual composition of the 2007 data biases calculations of the Theil statistic then and in later years. When calculating or presenting statistics, we therefore do not report results for 2007. We also remove establishments that entered the data in 2007 from calculations for subsequent years.

The EEO-1 survey forms originally included five racial/ethnic categories: “White (not of Hispanic origin),” “Black (not of Hispanic origin),” “Hispanic,” “Asian or Pacific Islander,” and “American Indian or Alaskan Native.” Because an employee can only be counted once on the survey, this classification scheme effectively treats Hispanic as a racial category, insofar as it is mutually exclusive with other categories. In 2007 the Commission redesigned the survey form: employees were first to be classified by “Race/Ethnicity” as “Hispanic or Latino” or not. Within the “Not-Hispanic or Latino” category, six options are available: “White,” “Black or African American,” “Native Hawaiian or Other Pacific Islander,” “Asian,” “American Indian or Alaskan Native,” and “Two or more races.”<sup>8</sup> The EEOC also began encouraging employers to record race and ethnicity based on employees’ self-identification rather than the employer’s visual inspection.<sup>9</sup>

Because the survey still only allows employees to be counted in one column, this new scheme *also* treats Hispanic or Latino origin as mutually exclusive with racial categories like white and black. In this regard the EEOC’s classification system differs from, for example, that used by the United States census, where respondents can declare a race and also identify as being of Hispanic or Latino origin. Treating Hispanics as a *de facto* coherent racial group is unusual but, frankly, no more fraught than any other socially constructed racial scheme (Roediger, 1991). Because these are the primitives we have in the data, we therefore assign employees to one of four categories in our analyses: white, black, Hispanic, or other.

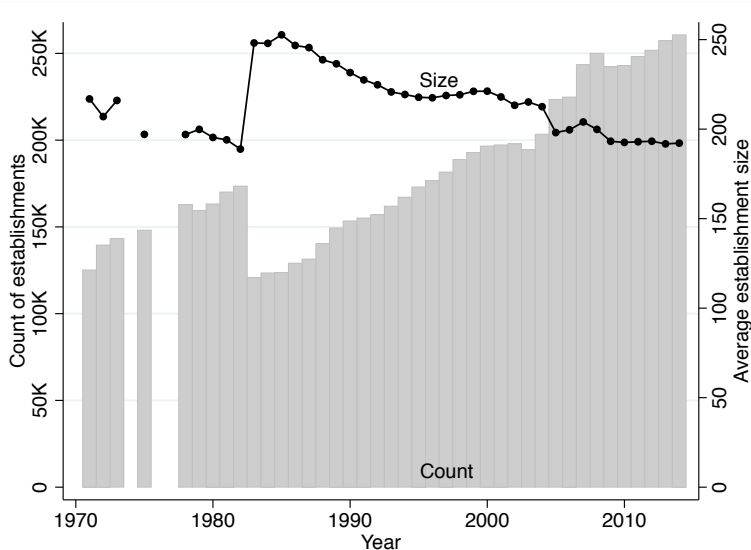
The data comprise 196,372 firms with 1,026,251 unique establishments, which together contribute 7,445,025 annual observations. The average establishment is observed for six years, though the range covers the length of the data series. Figure 3 presents counts of establishments and average firm size per year. The drop in observations after 1983 reflects an earlier change to the EEOC’s reporting requirements: in that year, they increased the minimum required reporting size to 100 employees (and 50 for federal contractors). We show in the appendix that excluding the smaller-establishment observations from before 1983 does not alter the pattern of our results.

These data give us several choices for defining the labor market used in our calculations of the Theil statistic. First we must geographically specify the labor market. For the results we present here, we have used the county. We get substantively similar trends, though, if we aggregate counties into Metropolitan Statistical Areas or Commuting Zones (Autor et al., 2017). We prefer to use counties because MSAs and CZs are not defined for non-metropolitan areas. Second we must demographically specify the labor market. Here we proxy the county workforce with the sum of large-establishment workforces in the county in our data. We compared the racial shares yielded by

---

<sup>8</sup>Examples of both versions of the survey form are included in the appendix. In the decade since the “Two or more races” category has been included, fewer than 3 percent of workers have opted for it as an identification.

<sup>9</sup>Instructions for employers can be found at <https://www.eeoc.gov/employers/eeo1survey/2007instructions.cfm>. Directions for assigning race can be found in appendix §4 therein.



**Figure 3:** Establishment counts (left) and average size (right) in the EEOC’s annual archives of EEO-1 establishment surveys.

this approach with estimates of county working-age populations from the decennial censuses and found they track each other very closely; details are in the appendix. We also found that county racial shares jump around more in the intercensal annual population estimates, partly because of the smaller sample size used in those surveys. We want to make sure that year-on-year changes in the Theil statistic are not being driven by noise in the estimation of the county population, so we prefer our aggregate-workforce measure. To the extent that the observed workforce underrepresents minority populations (for example, due to differences in employment or incarceration rates), our results will tend to understate the level and growth of segregation over time.

## Results

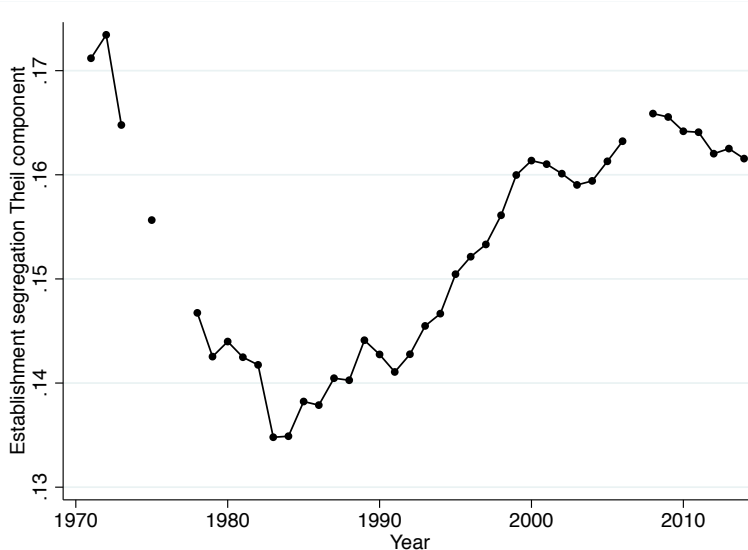
We begin with the overall trends. Figure 4 presents our index  $\bar{H}_j^a$  of within-area, between-establishment racial employment segregation, from 1971 through 2014. Such segregation declined sharply between the early 1970s and the early 1980s. After stalling for a few years in the mid-1980s, it began rising again, just as sharply, through the late 1990s. It then plateaued at levels comparable to the early 1970s through the Great Recession.<sup>10</sup> Establishment segregation in 2014 was comparable to forty years earlier.

This pattern is starkly different from that found for occupational segregation *within* establishments. Figure 5 plots  $\bar{H}_o^{aj}$  over the time period.<sup>11</sup> This is conceptually analogous to the indices of dissimilarity calculated across the same occupations within establishments by Tomaskovic-Devey et al. (2006), albeit as a multigroup rather than dichotomous measure.<sup>12</sup> This component follows

<sup>10</sup>Because the Theil statistic is nominally calculated on the full population of establishments, error bars are not explicitly defined for it. In the appendix we produce results with bootstrapped confidence intervals. These intervals are extremely small, so for simplicity we just present the calculated statistics here and elsewhere in the main text.

<sup>11</sup>For occupations, we use the nine specified on the EEO-1 survey form: officials and managers, professionals, technicians, sales workers, administrative support workers, craft workers, operatives, laborers and helpers, and service workers. Prior studies of occupational segregation using EEO-1 surveys have used the same.

<sup>12</sup>The index of dissimilarity does not explicitly account for unequal distribution of occupations across establishments,



**Figure 4:** Between-establishment segregation calculated from the Theil statistic, 1971–2014. See the text for information on adjustment for the 2007 cohort.

the same trajectory traced by indices of dissimilarity: rapid declines in occupational segregation during the 1970s and more gradual decline since 1980. If anything, this component shows larger declines than dichotomous measures of occupational segregation. Figure 5 suggests that, conditional on being present in an establishment, a worker’s race has become less predictive of their occupation over time. This is real progress. Yet figure 4 demonstrates that, during the same period, race has become more predictive of which establishment a worker is employed in.

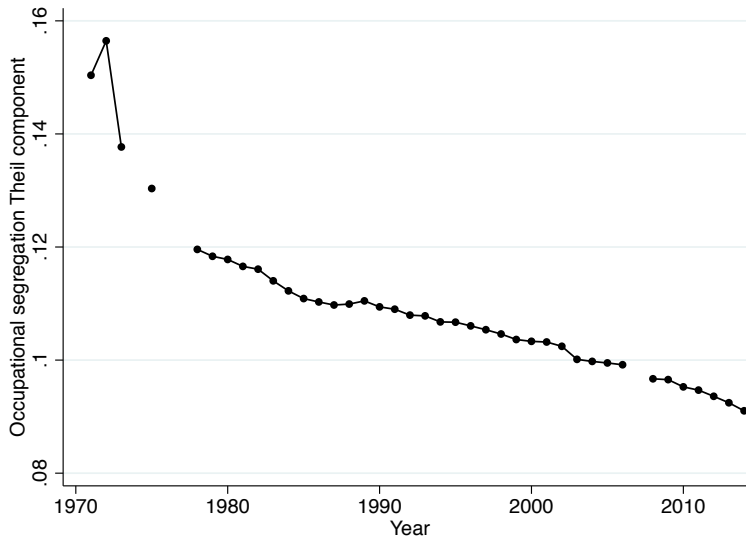
The relative levels of these components are also worth noting. In the early 1970s, the Theil statistic components for both establishment and occupational segregation were about .16. These comparable magnitudes imply that the two types of segregation were nearly equal contributors to total employment segregation in the United States. At that time, *either* reallocating workers across occupations within establishments *or* reallocating workers across establishments without altering occupational segregation would have had roughly the same impact on total segregation. By 2014, between-workplace segregation is a much larger contributor. Comparing figures 4 and 5 reveals that, in 2014, the contribution of between-establishment segregation is nearly twice as large as the contribution from occupational segregation in workplaces.

## Organizational change and turnover

As we discussed above, we first want to understand whether this increase in establishment segregation is due more to changes within organizations or to changes in the population of organizations (Stainback and Tomaskovic-Devey, 2012, p. 252). Some clarification of “organization” is necessary here. We have data on and have tried to focus our discussion around *establishments*, but of course multiple establishments can be parts of a single *firm*. It is worth asking how much of the increase in between-establishment segregation we have documented is really a between-firm story. Establishment turnover is much greater than firm turnover, for example, and firms might strategically segment their workforces between workplaces for a variety of reasons. The EEO-1 survey data include parent information for establishments. We thus decomposed  $H_{afj}$ , where  $f$  indexes

---

so it is most directly comparable to our within-county, within-establishment, between-occupation measure.



**Figure 5:** Within-establishment, between-occupation Theil statistic component, corresponding to occupational segregation, 1971–2014. See the text for information on adjustment for the 2007 cohort.

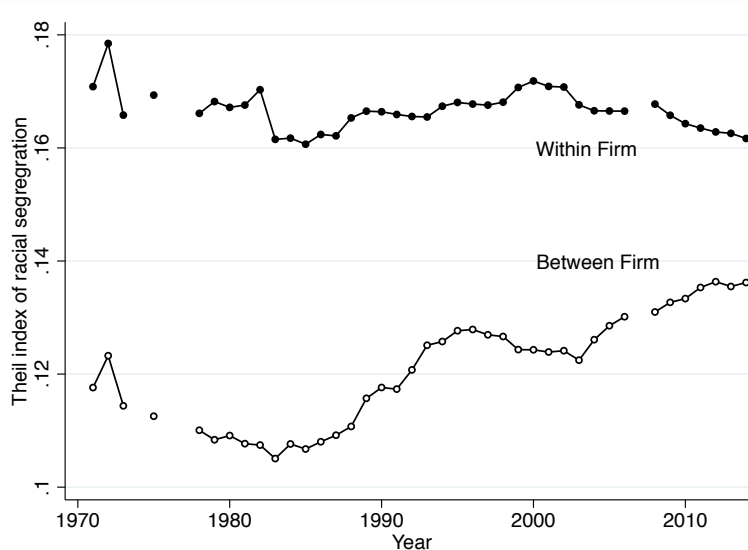
firms, to separate the two processes. Figure 6 shows the resulting components: for segregation *between* multi-establishment firms and for segregation between establishments *within* those firms. While the baseline level of racial segregation between establishments within firms is greater than between firms, all of the growth in between-establishment segregation among multi-establishment firms has been between firms, not between establishments within firms. We also show in the appendix that segregation has risen between *single*-establishment firms, where the distinction between firm and establishment is far less important. Thus, when we sometimes refer to “firms” rather than establishments here, we are not glossing over a major empirical distinction.

We can again decompose the Theil statistic to explore the relative contribution of establishment turnover to this growth in establishment segregation. In this case, we split each set of annual observations into  $t = 3$  “turnover groups”—establishments that enter the data in that year, ones that exit after that year, and ones that persist from the preceding through the succeeding year—and calculate  $H_{taj}$ .<sup>13</sup> We then calculate within-county, between-establishment  $\bar{H}_j^{ta}$  for each of those turnover groups. Figure 7 compares the values of  $\bar{H}_j^{ta}$  over time for entering and exiting establishments. In almost every year,<sup>14</sup> entering establishments are more segregated than exiting ones.<sup>15</sup> Segregation tends to be higher in smaller organizations (Bielby and Baron, 1984), partly for statistical reasons (Tomaskovic-Devey and Skaggs, 1999). Figure 7 suggests though that as newer establishments grow—up to the point where they reach 100 employees and come under regulatory oversight—their greater homogeneity persists. The persistent gap between entering and exiting workplaces means that, whatever changes may be happening within establishments, the turnover

<sup>13</sup>The year of an establishment’s entry into the EEOC’s data is not the founding date of the establishment. Establishments enter the first year that their employment crosses the threshold sizes described above, and exit when they drop below. Establishments that exit rarely reappear, suggesting that exit at least is correlated with establishment failure.

<sup>14</sup>The inversion in 1983 reflects the change in EEOC reporting requirements that year.

<sup>15</sup>One cannot conclude from figure 7 that diversity makes an establishment more likely to close—or indeed to exit the data in any other way. More diverse establishments have also on average been present in the EEOC’s data for longer, are probably older, and are more likely to be in shrinking industries. Disentangling any causal effect here requires considerably more data and research.



**Figure 6:** Between-firm and within-firm, between-establishment components of racial employment segregation, 1971–2014.

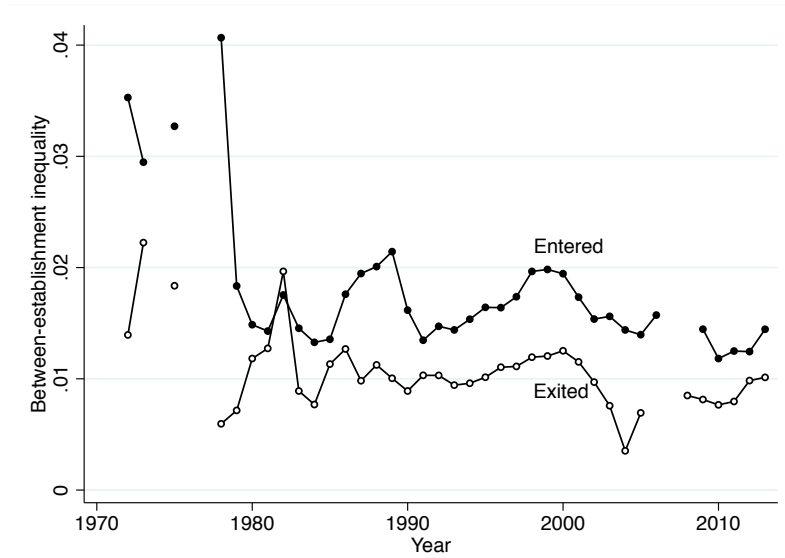
of establishments tends to exert an upward drag on segregation over time. Again, a simple but important implication should be noted here. Expecting newer workplaces to help drive integration is, at least under the current regulatory regime, a misplaced hope.

Table 1 presents all of the components of the Theil statistic when decomposed by entry, exit, and persistence. Columns 6 and 7 reproduce the data depicted in figure 7. Column 3 of that table shows that there is very little difference in the overall racial composition of entering, exiting, and persisting firms, in any given year or over time. Thus it is not that the overall workforce composition of entering establishments differs from that of exiting ones—Hispanic workers are not more likely to be in entering sites, for example, nor black workers in exiting ones. The higher levels of segregation among entering establishments represent greater actual segregation.

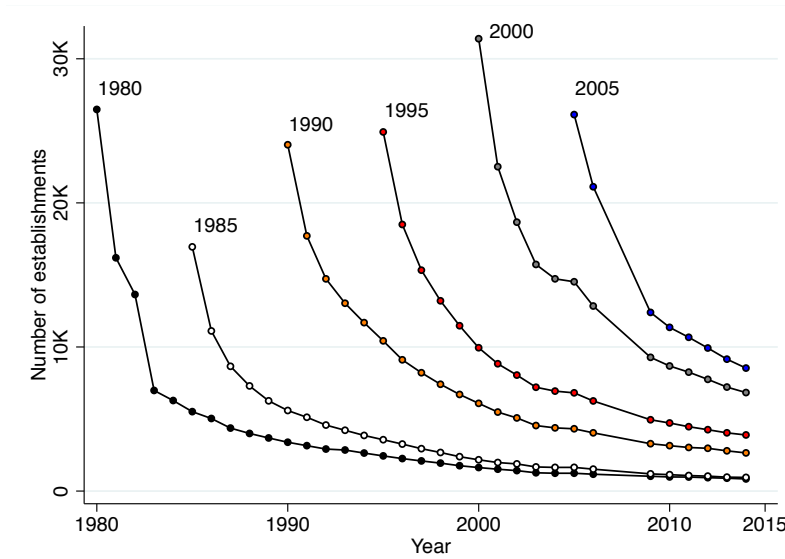
Column 5 of table 1 shows the Theil statistic component for persisting firms. This is much larger than the components for entering and exiting firms, and growing over time. This might seem to suggest that there is also increasing segregation within individual establishments over time; but care must be taken in interpreting this particular component. The cohort composition of these persisting establishments changes every year, as a result of entry and exit. Thus the persisting establishments in 1992 are those that were present in 1991, 1992, and 1993, while the persisting establishments in 1993 are those that were present in 1992, 1993, and 1994. Since we know that entering workplaces are less racially diverse than exiting ones, we could see an upward trend among persisting workplaces even if segregation were constant or declining at a slower rate than turnover were increasing it (See Gelman (2015) for a discussion of the same issue). And figure 8 demonstrates that turnover in this population is substantial.

To isolate within-establishment changes in the segregation index over time, we need to decompose the Theil statistic by establishment cohorts. The decomposition itself works the same way as when we split the data into entering, exiting, and persisting firms, save that this time we first partition segregation by establishments’ entry cohorts  $c$ , then by county. As before, the component of  $H_{caj}$  we are interested in is  $\bar{H}_j^{ca}$ , the within-cohort, within-county, between-establishment part of total segregation.

If one calculates this component, the results look almost identical to figure 8. This is because



**Figure 7:** Contributions to establishment segregation by entering and exiting establishments, 1971–2014. See the text for information on adjustment for the 2007 cohort.



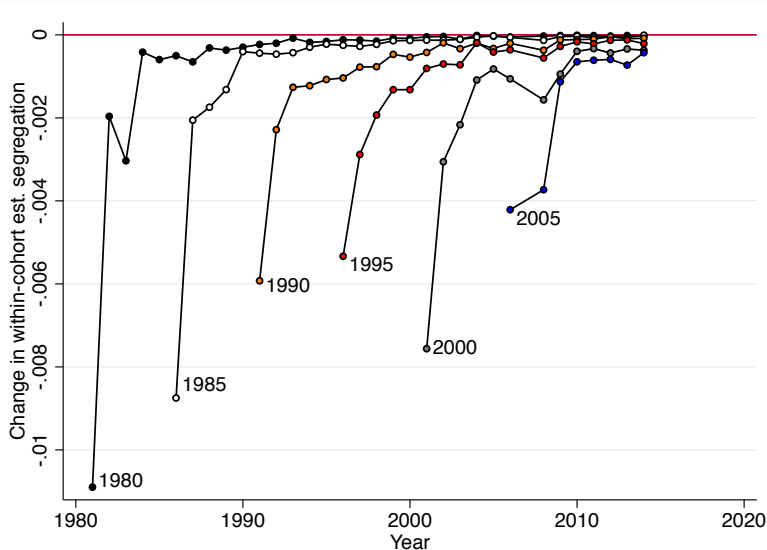
**Figure 8:** Counts of establishments present in each annual observation, grouped by entry cohort. For example, the “1990” curve shows the number of establishments that first entered the data in 1990 that are present in each year after 1990. The pattern is similar across all 40 available cohorts; we show six here for clarity.



**Table 1:** Decomposition of Theil statistic of racial segregation by firm's turnover status

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Year	$H$	Between status (Enter/ Persist/ Exit)	Within status, Between areas	Within area, Between worksite (Persist)	Within area, Between worksite (Enter)	Within area, Between worksite (Exit)
1972	.3125	.0008	.1385	.1275	.0328	.0129
1973	.3016	.0008	.1465	.1026	.0295	.0222
1974	.	.	.	.	.	.
1975	.2978	.0005	.1500	.0963	.0327	.0184
1976	.	.	.	.	.	.
1977	.	.	.	.	.	.
1978	.2930	.0006	.1529	.0928	.0407	.0060
1979	.2900	.0001	.1527	.1116	.0183	.0072
1980	.2904	.0002	.1519	.1115	.0149	.0118
1981	.2886	.0002	.1521	.1094	.0143	.0127
1982	.2884	.0001	.1527	.0984	.0175	.0197
1983	.2810	.0002	.1522	.1052	.0145	.0089
1984	.2835	.0001	.1538	.1086	.0133	.0077
1985	.2818	.0001	.1500	.1068	.0136	.0113
1986	.2828	.0001	.1510	.1015	.0176	.0127
1987	.2839	.0001	.1500	.1046	.0195	.0098
1988	.2882	.0003	.1534	.1032	.0201	.0112
1989	.2933	.0001	.1552	.1065	.0214	.0100
1990	.2952	.0001	.1570	.1129	.0162	.0089
1991	.2942	.0003	.1578	.1123	.0135	.0103
1992	.2963	.0005	.1590	.1118	.0147	.0103
1993	.3007	.0002	.1603	.1164	.0144	.0094
1994	.3022	.0002	.1605	.1165	.0154	.0096
1995	.3056	.0002	.1602	.1187	.0164	.0101
1996	.3053	.0002	.1581	.1196	.0164	.0110
1997	.3047	.0002	.1560	.1200	.0174	.0111
1998	.3048	.0003	.1538	.1191	.0197	.0119
1999	.3061	.0002	.1513	.1228	.0198	.0120
2000	.3059	.0002	.1500	.1238	.0194	.0125
2001	.3044	.0001	.1489	.1265	.0173	.0115
2002	.3043	.0003	.1494	.1296	.0154	.0097
2003	.3014	.0005	.1472	.1305	.0156	.0076
2004	.3035	.0003	.1476	.1378	.0144	.0035
2005	.3053	.0003	.1476	.1367	.0140	.0067
2006	.3065	.0006	.1478	.1327	.0157	.0096
2007	.	.	.	.	.	.
2008	.3065	.0006	.1478	.1327	.0157	.0096
2009	.3061	.0005	.1490	.1361	.0130	.0074
2010	.3042	.0004	.1480	.1383	.0105	.0070
2011	.3031	.0005	.1481	.1357	.0114	.0073
2012	.3018	.0006	.1495	.1318	.0111	.0089
2013	.2998	.0008	.1483	.1279	.0130	.0099

Columns 3–7 sum to column 2, which tracks the sum of  $H_a$  and  $H_j^a$  from figure 10. Columns 6 and 7 show the trends presented in figure 7. Column 3 shows no evidence of different racial composition *between* entering, persisting, and exiting firms.



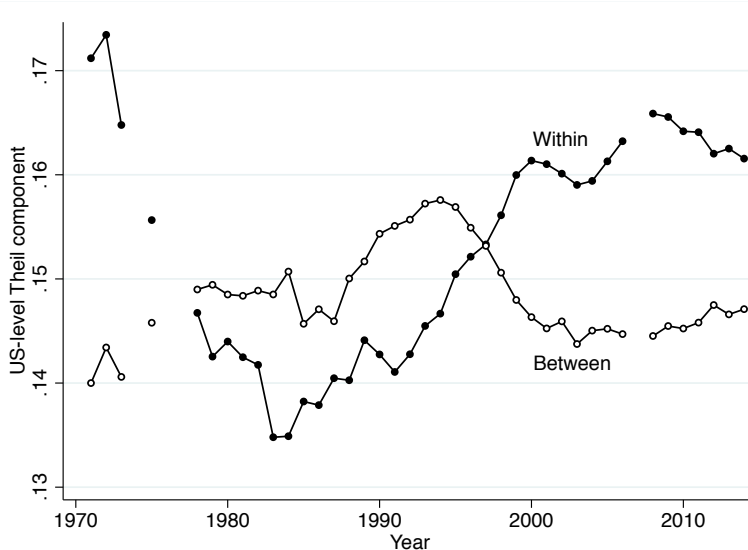
**Figure 9:** Changes in the pure-segregation element of the within-cohort, within-area, between-establishment of the Theil statistic. For example, the “1990” line shows year-on-year changes for establishments that first entered the data in 1990. The pattern is similar for all 40 available cohorts; we show six for clarity.

change in a component’s contribution to the total Theil statistic has two elements: the change in the “pure segregation” effect and the change of the size of the component relative to the total population (Conceição and Ferreira, 2000).<sup>16</sup> If  $H_t$  is the relevant component measure at time  $t$ ,  $S_t$  is the pure-segregation element and  $W_t$  is the size of the component, then  $H_2 - H_1 = (S_2 - S_1) \times \frac{W_2}{W_1}$ . We want to isolate  $\Delta S$  from the effects of  $\Delta W$ . We observe  $H_t$  and  $W_t$  and can therefore solve for  $\Delta S = \Delta H \frac{W_1}{W_2}$ .

Figure 9 plots  $\Delta S$  for selected cohorts. Two things should be highlighted. First, changes in the pure segregation effect are almost always negative. Overall, establishments *do* reduce their racial segregation over time. Second, the size of those reductions shrink over time. To some extent this reflects the nonlinear response of indices like the Theil. Replacing the same number of people in completely segregated establishment will have a larger effect than replacing the same number a second time, and so on (Conceição and Ferreira, 2000). Also to some extent such reductions *have* to be asymptotic. But to some extent this also likely reflects inertia within establishments (Stinchcombe et al., 1968). Conditional on survival, older establishments are less likely to substantially alter their workforces than younger ones.

When put together, table 1 and figures 8 and 9 help make sense of the increase in between-establishment segregation. It is not being driven by increasing segregation within establishments over time. There has been progress on integration *within* specific establishments. The size of that progress is modest though, such that the year-on-year changes within workplaces are smaller in aggregate than the differences in levels between new and existing workplaces. The typical establishment that exits the data, meanwhile, is more diverse than the population at large. Thus were we to observe the modal establishment over time, we might conclude that segregation is falling. When we observe the population though, we conclude the opposite.

<sup>16</sup>This is partly why the values in columns 6 and 7 of table 1 are so much smaller than those in column 5: there are considerably fewer entering and exiting than persisting establishments in the population at any given time.



**Figure 10:** Between-county and Within-county, between-establishment components of U.S.-level Theil statistic, 1971–2014. See the text for information on adjustment for the 2007 cohort.

### Compositional changes in demographics, occupations, and industries

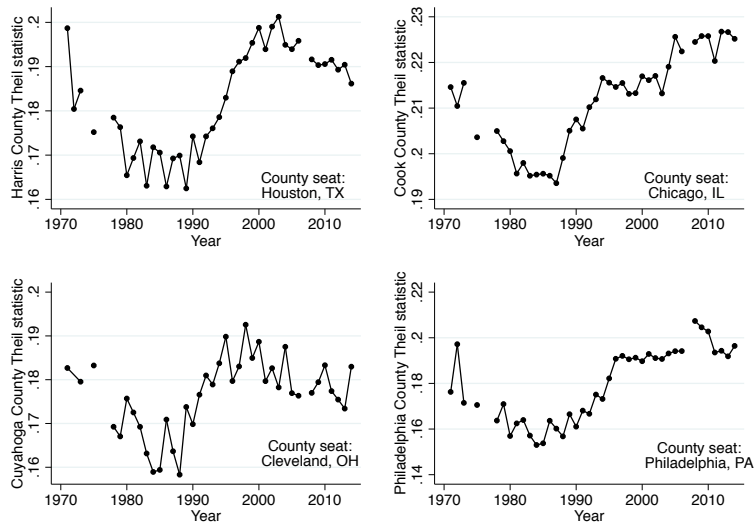
How might we account for increased racial sorting between workplaces? One possibility is that it simply reflects the changing demography of the American workforce. Perhaps the rise of immigration from the 1970s through the late 1990s was geographically concentrated so as to produce large cross-establishment differences that really just reflect very different labor markets. This is not the case, though. The within-county, between-establishment component of the Theil index shown in figure 4 measures the deviation of a workforce’s entropy from that of the combined workforces in its local labor market, not from the country as a whole.

We can formally compare the Theil statistic components for between- and within-area segregation. Figure 10 plots  $H_a$ , segregation due to the uneven distribution of workers across counties, and  $\bar{H}_j^a$ , our measure of within-county, between-establishment segregation (the series shown in figure 4). Between-county racial segregation did indeed increase starting in the late 1980s. Yet it declined just as sharply after the mid-1990s, even as within-county, between-establishment segregation continued to increase. Figure 10 underscores that the increase in racial segregation between establishments shown in figure 4 is independent of racial differences between county populations.

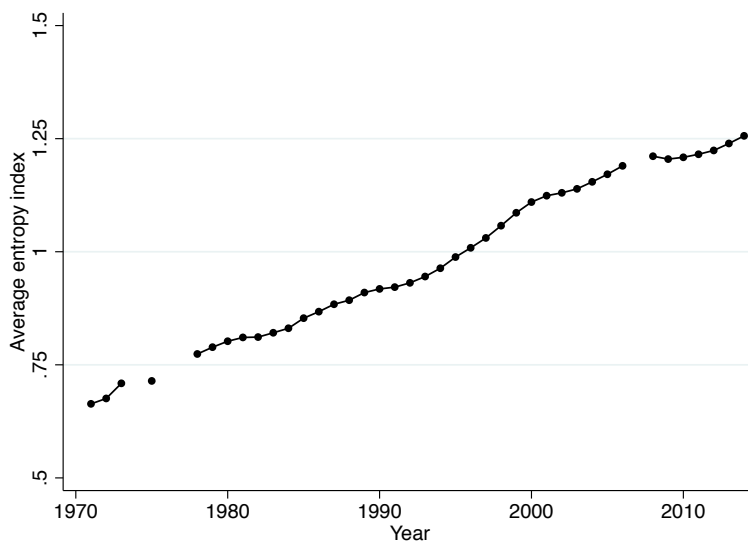
Furthermore, we observe these trends in between-establishment segregation not just across the United States as a whole but also within counties. Figure 11 plots the  $H_j^a$  components of  $\bar{H}_j^a$  for the counties of Houston, Chicago, Cleveland, and Philadelphia. The year-to-year variance is greater in individual counties because the total number of establishments is smaller, but the same basic pattern can be seen in each. (The trends in these four counties are broadly representative of those in the 100 largest counties over time; see the appendix for details.)

Nor are the component’s values driven by idiosyncrasies in the growth of overall workforce diversity. Figure 12 shows the trend in  $\bar{E}_a$ , the average of county-level entropies defined in equation 5. The racial diversity of the American workforce has monotonically increased during this period, even as establishment segregation first declined, then rose, then plateaued.

While between-county racial differences cannot explain between-establishment racial segregation, they can suggest forms that other explanations might take. Note that both the increase and



**Figure 11:** Between-establishment segregation indices for specific counties. Compare to the national trend in figure 4, which is a weighted average of all such county components.



**Figure 12:** Average of county-level entropy indices, weighted by county population size, 1971–2014.

the decrease in the between component in figure 10 reflect growth in the Hispanic population. That growth was first concentrated in the southwest and a few large northern cities, and that concentration tended to increase between-county segregation. Later, as Hispanics moved into more and more parts of the country (Stolberg, 2014), their concentration was reduced and that increase was offset. This raises a point whose implications we explore below: substantial population changes can be associated with comparable levels of segregation. None of which should undermine the main point: the increases we see in between-establishment segregation cannot be reduced to larger shifts in the population.

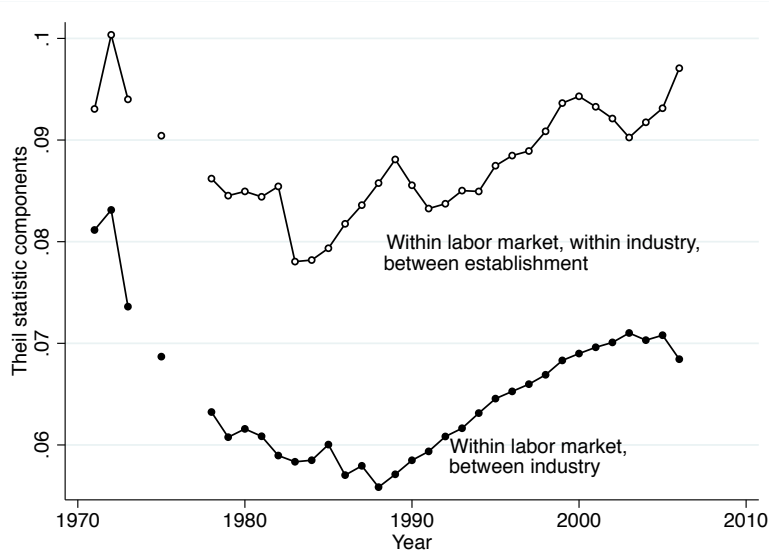
Nor can broad changes in the racial composition of occupations explain the patterns in figure 4. We depicted declining occupational segregation in figure 5. The decomposition from which that figure is taken treats occupations as nested within establishments, but we can also compute  $H_{aoj}$ , splitting establishments to nest them within occupations. Doing so produces virtually identical results for within-area, between-occupation racial segregation (the resulting figures are available in the appendix). This decomposition is useful for showing that the overall decline in occupational segregation is not simply due to the decreased racial diversity within establishments depicted for example in figure 2. However, it also rules out shifts in the mix of occupations between establishments as the source of rising establishment segregation: establishment segregation cannot be reduced to occupational segregation, because the latter is not rising.

Changes in the industrial structure of the economy could drive segregation in the same way that changes in the occupational structure could. If some industries tend to be more segregated than others, then their growth will increase observed segregation in a labor market. As discussed above, growing segregation between industries need not be a direct reflection of any racial preferences. If educational attainment for example is unevenly distributed and if some industries heavily rely on highly-educated workers, then the growth of such industries will raise observed racial segregation between establishments (Goldin and Katz, 2010). It makes sense therefore to study industry’s role.

There are some limitations on the industry data recorded in the EEO-1 surveys. The EEOC has used different versions of the Standard Industrial Classification system over the years, and has varied in how fine-grained its classification has been (that is, in some years 5-digit SIC codes are available, while in others only 2- or 3-digit codes are). As part of its larger updates in 2007, the EEOC began using the North American Industry Classification System rather than the SIC. The mapping from SIC to NAICS is many-to-many, meaning that we cannot SIC codes forward nor NAICS codes backward unless an individual establishment spans the transition. Thus when decomposing by industry we compute our time series through 2006. There are published crosswalks for the different editions of the SIC system that the EEOC used between 1971 and 2006. We standardize our observations to the 1987 SIC, using the rules outlined by Stainback and Tomaskovic-Devey (2012) in their methodological appendix. We conduct our analyses here using the two-digit SIC codes, which groups together industries with broadly similar labor-force skill and educational requirements.<sup>17</sup> Conceição et al. (2001) show that Theil statistics of income inequality built with increasingly fine-grained versions of the SIC codes produce qualitatively similar results as those from 2-digit codes; we find the same here.

Similar to the decompositions described above, we compute  $H_{aij}$ , where  $i$  indexes industries. We calculate  $\bar{H}_i^a$ , the racial segregation within labor markets but between industries, as well as  $\bar{H}_j^{ai}$ , within labor market/industries but between establishments. We display both in figure 13. We find that segregation between industries has been increasing since the late 1980s. This jibes with earlier work that has documented how the decline of manufacturing relative to service work has

<sup>17</sup>The EEOC includes 2-digit SIC codes in some years but its most complete coverage is at the 3-digit level. We construct 2-digit codes using the 3-digit series.

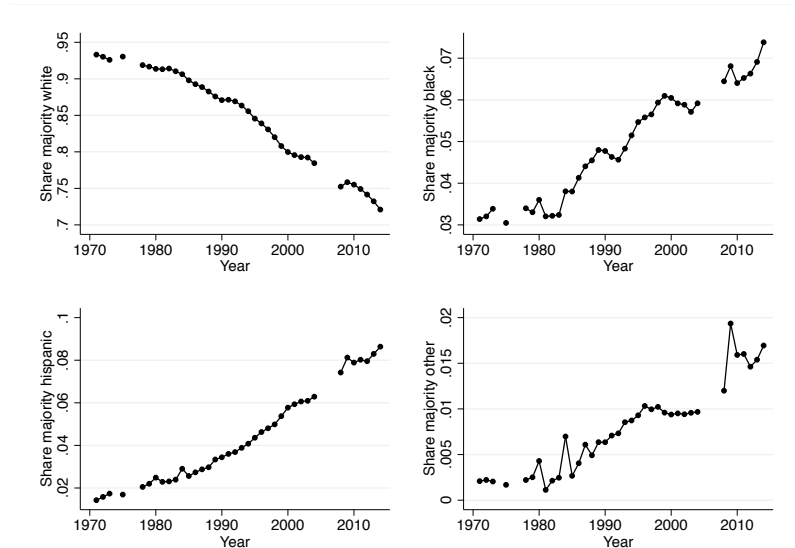


**Figure 13:** Racial employment segregation between industries and between establishments within industries, 1971–2006. The EEOC switched to the NAICS in 2007, which limits comparison; hence we stop this series in 2006.

come at the expense of larger, more integrated workplaces, i.e., production facilities (Sugrue, 1996; Dwyer, 2013). Yet accounting for this trend does not remove the upward trend in *within*-industry segregation between establishments. The comparison with occupations is informative here. Unlike the changes in the occupational structure, trends in racial sorting between industries can explain some of the increase in establishment segregation (Mouw and Kalleberg, 2010). Yet we cannot attribute all, nor even most, of the increase to those industrial changes.

Of course, two-digit industries are a coarse-grained disaggregation of the economy. Over the period we consider here, there has been considerable dissolution of the firm, as companies outsource more functions to specialized providers (Cappelli, 2001; Cobb and Stevens, 2016). If the establishments within an industry have grown increasingly specialized by the types of jobs they perform, then some of what we label increasing establishment segregation may actually reflect between-job segregation, combined with that specialization. This is the final source of establishment segregation that we outlined above. Here though we start to butt up against the limits of the EEO-1 data. The EEO-1 surveys record employment across nine broad occupational categories. While these categories are useful for some analyses, most notably the movement of different groups into managerial and supervisory roles (Kalev et al., 2006; Cohen et al., 2009; Skaggs, 2009), they are too coarse-grained to track specific shifts in sourcing arrangements. Furthermore, the EEOC’s occupational-classification rules have shifted somewhat over time (Robinson et al., 2005; Ferguson, 2016). Nailing down changes in the boundary of the firm as a motor of increasing establishment segregation must await more detailed data.

These analyses take us near the limits of what the EEO-1 survey data on their own can tell us about the evolution of establishment segregation by race. But focusing on those limits should not blind us to how much ground we have covered. Using these data, we can demonstrate that racial segregation in American workplaces has increased for more than a generation, such that it now stands at levels comparable to the early 1970s; that this rise has happened more through the turnover of specific workplaces, rather than through the turnover of personnel within workplaces upon which most prior work has focused; and that this rise cannot be explained away through shifts in pre-organizational sources of heterogeneity like population demographics, occupational change



**Figure 14:** Share of establishments with majority employment by specific races, 1971–2014. See the text for information on adjustment for the 2007 cohort.

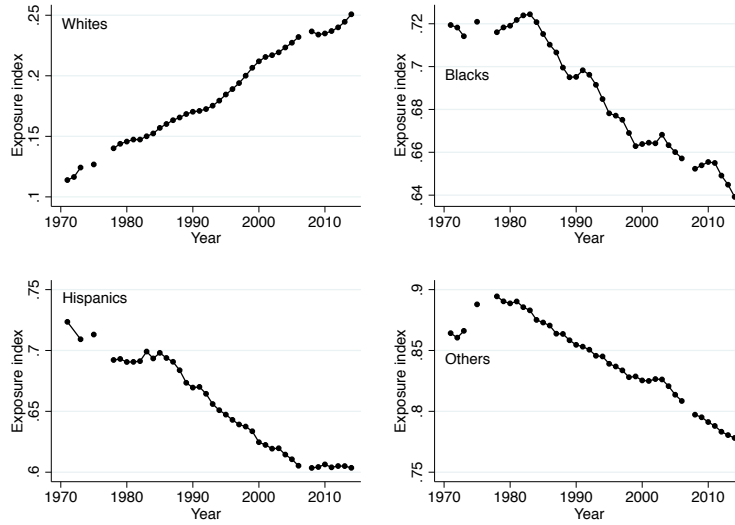
(within the EEOC’s broader categories), or industrial transformation. Organizations also play a role.

### The perception of segregation

The pattern of our results may feel surprising. It feels perverse to say that American workers are more segregated between workplaces by race today than they were in, say, the early 1980s. The United States has a long and sordid history of racial exclusion in employment (Olzak, 1989; Levy, 1994; Loewen, 2005). It also has a long and inspiring history of struggle against that exclusion (Stromquist and Bergman, 1997; Nelson, 2001). A signal trend of the last generation has been the opening up of once all-white, all-male workplaces to women and racial minorities Stainback and Tomaskovic-Devey (2012). The lived experience of many Americans has been one of increasing diversity in their workplaces. Are all of these impressions wrong?

They are not, but they are incomplete. To get a handle on why, we present one more set of findings. Figure 14 shows that the share of majority-white large establishments has fallen more or less consistently, from nearly 95 percent of large workplaces in the early 1970s to just over 70 percent in recent years. Meanwhile, the share of majority-minority establishments has grown. The share of establishments with no non-whites has also declined (Tomaskovic-Devey et al., 2006). At first glance this pattern is encouraging, but it should be hedged with caveats. First, whites still constitute a majority of the workforce in most U.S. counties, and a plurality in even more. If race were orthogonal to employment, we would expect more workplaces to have white majorities, and far fewer to have specific minority majorities. Second, figure 14 does not plot any particular trend in “minority businesses.” Employment need not indicate ownership, nor management composition.

Third, these different trends across races mean that the same level of segregation may be perceived very differently by different groups. We can formalize this with the exposure index, sometimes called the Lieberman (1980) index, which measures the likelihood that a member of one race is likely to interact with a member of a different race in some defined unit. It assesses the chance that a randomly chosen alter from ego’s unit will be of a different race than ego. Following



**Figure 15:** Lieberman’s exposure index, by race, 1971–2014. Each annual data point represents a population-weighted average of the index calculated across establishments at the county level. See the text for discussion of the post-2007 data adjustment.

the notation in Massey and Denton (1988): if  $x_j$  is the count of members of ego’s race in the unit,  $y_j$  is the count of members of different races,<sup>18</sup>  $w_j$  is the total membership of unit  $j$ , and  $X$  is the count of people of ego’s race in the larger area, then the index is defined as follows:

$${}_x P_y^* = \sum_{j=1}^J \frac{x_j}{X} \frac{y_j}{w_j} \quad (6)$$

We calculate equation 6 separately for each of our four racial categories in each county in each year. We can then calculate national exposure indices as averages of the county indices for each race, weighted by county-race size. We present the resulting four time series in figure 15. The difference in the trends should be obvious. For white workers, exposure to non-white co-workers has steadily increased for forty years—and most American workers are white. The lived experience of the modal worker is not incorrect. But it does not generalize.<sup>19</sup> For during the same period, exposure has tended to fall for nonwhite workers—fitfully for blacks, more steadily for Hispanics, Asians, and others. The decline in the exposure index is greatest for Hispanics, who were also the fastest-growing segment of the population during these years. It is tempting therefore to say that our results “just” pick up an immigration wave that, for reasons of language more than race, concentrated Hispanic workers in certain establishments. But this cannot explain the parallel decreases in exposure for black, Asian, and other workers.<sup>20</sup>

<sup>18</sup>The exposure index is a dichotomous measure. Thus we define  $x$  and  $y$  as white/nonwhite, black/nonblack and so forth.

<sup>19</sup>Stainback and Tomaskovic-Devey (2012) make a similar point about skewed perceptions resulting from labor queues that privilege white males: “Thus, we have the perverse experience that even as their benefits from racial, ethnic, and gender labor queues increase, white men in high-quality jobs experience more integrated workplaces. The perception of competition and reverse discrimination is likely to be highest in labor markets where the benefits from social closure and discrimination are greatest” (p. 205).

<sup>20</sup>It also equates “Hispanic” with “immigrant,” which ignores both natural increase and Hispanics’ centuries-long history in the United States.



We have shown that entering establishments are less diverse than exiting ones; this does not mean that they are more white. Rather, the share of majority-black, majority-Hispanic, and majority-other workplaces has grown in each annual cohort (results are available on request). The growing share of majority-minority establishments, combined with minorities' growing share of the national workforce, has more than offset the increase of minorities in majority-white workplaces.

## Discussion

In this paper, we have analyzed trends in racial segregation between establishments in the United States over the last forty years. We have focused on establishment surveys gathered by the Equal Employment Opportunity Commission. Prior studies, using these data, have documented non-trivial decreases in occupational segregation since the Civil Rights movement (Tomaskovic-Devey et al., 2006; Stainback and Tomaskovic-Devey, 2012). We show that this finding is correct but incomplete. Even as their race has become less predictive of what a person does for a living, it has become more predictive of where they work. This trend is driven more by turnover among establishments than by changes to workforces within them, and it holds even after we separate out differences in the composition of workforces across labor markets, occupations, and industries.

This study makes several contributions. We start with a policy recommendation. In 1983, the EEOC's reporting requirements were changed. Where once establishments had to begin filing EEO-1 surveys once they grew to 50 workers, 25 if they had federal contract work, now they only had to begin at 100 workers, 50 for contractors. We think the threshold should be lowered again. Reports like these have become easier to compile over time, so the regulatory burden has shrunk. More significantly, our figures 7 and 9 show that establishments are comparatively less diverse when they come under regulatory oversight, and that they improve diversity at diminishing rates afterward. We cannot and do not want to declare that the impact of oversight on workforce composition seen here is causal. But presume for a moment that it is. Presume that we could start "bending the curve" when establishments were smaller. Given the survival curves depicted in figure 8, such a change might have a far bigger aggregate impact on hiring and diversity than specific diversity programs adopted in more established workplaces (e.g., Miller, 2017). The cost of experimenting on this would be relatively small, and the benefit potentially large.

This recommendation involves a shift in focus, from the routines used in a population of workplaces to the changes in that population. We do not want to imply that organizational routines, or the study of them, are unimportant. We want to expand the tools at researchers' and practitioners' disposal, not replace them. We think that a parallel with past research on employment discrimination is useful here. Early work on stratification in formal organizations focused on the routines of pay and promotion (Kanter, 1977; Sørensen, 1977; Baron and Bielby, 1980). Much of this work took the pool of workers as given, or beyond the scope of the theory. Researchers quickly noted though that the labor pool places an upper bound on how effective such policies can be. The "best" promotion routines cannot improve diversity if the candidate pool is not diverse. Understanding and improving the integration of an organization requires understanding the dynamics of turnover among its personnel (Fernandez and Weinberg, 1997; Petersen and Saporta, 2004; Sørensen, 2004; Rubineau and Fernandez, 2013; Fernandez-Mateo and Fernandez, 2016). We want to widen the aperture again. Turnover of firms and workplaces places an upper bound on how effective firm policies, including policies that shape firms' labor pools, can be. Understanding and improving the integration of a population of organizations requires understanding the dynamics of turnover among the constituent organizations.

To expand our focus in this way is to immediately increase the avenues for research on em-

ployment segregation. Take one example: the theory of taste-based discrimination first proposed by Becker (1957), in which ascriptive discrimination is a cost that biased employers choose to pay, insofar as they do not necessarily choose the best workers for each job and thus operate less efficiently. From this theory, Becker predicted that more competitive markets would be more likely to drive out discriminating employers. Whether this idea holds is still very much an open question in organizational research (Fernandez, 2015; Pager, 2016). Leveraging representative, longitudinal data across the economy lets us bring more information to bear on the question than previously. At a bare minimum, it would be useful to contrast the relative diversity of entering and exiting establishments based on how competitive their industries are.

This is a straightforward and promising avenue for future research, but we also think that it is exciting for the potential connections it opens with *other* research topics. For example, social scientists have recently become concerned with the growth of oligopoly in the American economy (Decker et al., 2013; Barkai, 2016; Autor et al., 2017). Most such concern centers on the welfare effects of reduced competition (Gordon, 2016; Card et al., 2016) or on the political clout of large firms (Davis, 2013; Furman, 2016), but in this context there is another issue. If there has been a secular reduction in competitive pressure in some industries, can we identify any corresponding change in the pace of integration among workplaces in those industries? If we cannot, then focusing on population-level processes at least gives us grounds to question Becker’s original prediction. But if we can, then stratification researchers would have an interest in competition and anti-trust policy. For policies that shape the population of workplaces can shape the composition of that population’s workforce. The strength of competition in an industry should also present a theoretical bound on how effective specific firms’ diversity policies might be, both in absolute terms and relative to the changes from organizational turnover.

Such connections should be equally exciting for economists. Recent work has shown that much of the increase in pay inequality in the American economy has developed between firms within industries (Barth et al., 2014; Song et al., 2015). Sociologists have tied some of this increase to the differential use of outsourcing by such firms (Cobb and Lin, 2017). We show here that, over much of the same period, racial segregation between establishments within industries has been rising. An exciting and important next step would be to link composition, ownership, and pay data to ask, how much of that surge in between-firm inequality can we attribute to racial stratification, and what role does minority entrepreneurship play in strengthening or offsetting this trend (Fairlie and Robb, 2007)?

Of course, many such questions are of interest to both economists and sociologists. Take the question of whether more diverse firms outlast less diverse ones. As discussed, there is an open question whether this effect varies with the strength of competition; but it is also an open question whether the effect exists *at all*. Very little generalizable work exists on how diversity affects the vital rates or performance of establishments, which is critical for any “business case” for diversity (Black et al., 1996; Bunderson and Sutcliffe, 2002; Kochan et al., 2003; Herring, 2009; Stojmenovska et al., 2017). If data like these, which record both the progress of diversity within organizations and the turnover of organizations, can be linked to richer information on firm performance, then such work becomes possible.

All of these possibilities are directions of productive future research raised by the approach we take here, but there remains the surprising social fact that is our core finding: American workplaces are more segregated by race today than they were a generation ago. We have ruled out several high-level explanations and ruled in mechanisms related to establishment turnover, but by themselves the EEOC’s data do not have enough information to really explain this change. We thus want to discuss some of the limitations of this study and how future work could build on our findings.

First, we have taken care to discuss establishments entering and exiting the EEOC’s data, rather

than being born or dying, and have referred to cohorts in the data; all of this is because we do not have information within the EEOC’s records on firm or establishment age. Useful as the current cohort information is, it is a noisy proxy for age, which is more usefully theorized as affecting organizational structure and performance (Stinchcombe et al., 1968; Hannan and Freeman, 1984). Linking these records to others that have information on age and other characteristics would be very useful. Even if such information is only available for a subset of these establishments, such linking would at least suggest whether additional variables could account for some of the trends seen here. For example, during periods of rapid employment growth, establishments may have less time and ability to recruit and evaluate a diverse applicant pool. This would be consistent with diversity being lower in younger entering workplaces than in older ones.

Second, this same point about richer data holds at the worker level. The EEOC’s records contain invaluable compositional information, but their smallest unit of observation is the race/sex/occupation tuple. Ideally future research would add information about the workers themselves, such as education or language proficiency (Gradín et al., 2015). Such variables could not make the trends seen here go away, but they could help to explain why we observe them in the first place.

Third, we have discussed shifting boundaries of the firm as a likely source of some of the increases we see here, but have also noted that we need more detailed data on jobs in order to explore this idea directly. Such data are gathered elsewhere by the federal government, but at present there is no straightforward crosswalk between the EEOC’s data and other sources. Constructing such crosswalks would not only expedite answering this question but would also simplify all future research efforts that tie EEO-1 data to other sources. Once a crosswalk to more detailed job information is built, addressing the issue of outsourcing becomes relatively straightforward. Outsourcing can be thought of as increasing segregation of *jobs* between establishments and firms, and thus can be included in a hierarchical decomposition like the ones we have performed here. Furthermore the relative contributions of individual job types to a within-job, between-establishment component can be assessed, to see whether increases in segregation are in fact associated with the types of jobs that prior work has found most likely to be outsourced.

Fourth, we have taken the composition of establishments’ labor markets for granted. We do not think that doing so biases our findings in any way; we mention it because it should be possible to use such data to explore the relationship between residential and employment segregation. This study implements a more robust and detailed measure of segregation, the Theil statistic, than work on employment segregation has used. Lichter et al. (2015) have recently taken a similar approach to residential segregation. Their work reveals that, while the racial divide between U.S. cities and suburbs has closed somewhat in recent decades, segregation between the suburbs (and between neighborhoods in cities) has risen. Studies like theirs, combined with this one, suggest that focus on the salient metrics of earlier decades have led social scientists to ignore significant divisions along new lines in American society. There is also a more concrete overlap between that work and ours: how much of the increase in between-establishment segregation might be correlated with and possibly accounted for by between-community macro-segregation like Lichter *et al.* describe? Because the EEOC’s data contain good address information that lends itself to geocoding (Ferguson et al., 2017), it would be possible to situate the establishments observed here in their immediate communities and explore these relationships.

Residential-segregation research also suggests specific processes that might help explain the patterns we find here. Between-establishment segregation declined for about fifteen years, through the mid-1980s, before increasing again. Bader and Warkentien (2016) have documented a similar trajectory in residential segregation, where people of one race steadily replaced people of another. Segregation trends calculated during such periods of “ethnic succession” tend to be U-shaped; the similar high levels at the end points mask the fact that the dominant group has changed. It is an

open question how much of the integration of once heavily-white workplaces in the 1970s was eventually offset by continued white exit from those firms and succession by racial minorities. Methods similar to those developed for studying patterns of integration and re-segregation in neighborhoods could be used to classify industries and workplaces into such types.

We close with some methodological concerns that this work raises. We were quite surprised to find that between-establishment segregation increased and then plateaued over the last generation. In fact, there is virtually no representative work on establishment segregation; yet the sense that it has declined pervades much stratification research. We think that this has two causes. First, occupational segregation, which is what most prior work studies, *has* declined. Second, the better educated you are and the higher status your job, the greater is the the probability that you work in an establishment that *has* diversified. Yet the nature of segregation makes it dangerous to generalize from one’s own social world to the wider one. The major contribution of this paper is an empirical one, but theories rely more than we like to think on empirical assumptions, often all the more so because those assumptions are implicit. If segregation between establishments has declined; if new establishments are more diverse than old ones, and thus turnover and entrepreneurship are not areas of concern for stratification; if most of the action on diversity happens within establishments—if these things are true, then we can ignore population-level processes and concentrate on effects within workplaces over time. If they are not, then by focusing solely within firms we ignore sources of problems and opportunities for solutions.

In sum, these limitations, prospects for future research, and concerns remind us how little we still actually know about the dynamics of employment segregation. Our present situation may be worse than mere ignorance. As in the case of perceived increases in diversity by white workers, many of the social facts that we think we know may have to be unlearned. This may sound depressing. We instead prefer to remember that forthright recognition of ignorance is the foundation of all good science.

## References

- Aaronson, D., Hartley, D., and Mazumder, B. (2017). The effects of the 1930s HOLC ‘redlining’ maps. Technical report, Federal Reserve Bank of Chicago. Working Paper 2017-12.
- Allport, G. W. (1954). *The Nature of Prejudice*. Perseus Books, Cambridge, MA.
- Alonso-Villar, O., Del Río, C., and Gradín, C. (2012). The extent of occupational segregation in the United States: Differences by race, ethnicity, and gender. *Industrial Relations*, 51(2):179–212.
- Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Van Reenen, J. (2017). Concentrating on the fall of the labor share. *American Economic Review Papers and Proceedings*. Forthcoming.
- Bader, M. D. and Warkentien, S. (2016). The fragmented evolution of racial integration since the Civil Rights movement. *Sociological Science*, 3:135–166.
- Barkai, S. (2016). Declining labor and capital shares. University of Chicago Booth School New Working Paper Series No. 2.
- Baron, J. N. and Bielby, W. T. (1980). Bringing the firm back in: Stratification, segmentation, and the organization of work. *American Sociological Review*, 45:737–765.
- Barth, E., Bryson, A., Davis, J. C., and Freeman, R. (2014). It’s where you work: Increases in earnings dispersion across establishments and individuals in the U.S. Technical Report 20447, NBER.
- Becker, G. S. (1957). *The Economics of Discrimination*. University of Chicago Press, Chicago.
- Bender, S., Lane, J., Shaw, K. L., Andersson, F., and Von Wachter, T., editors (2008). *The Analysis of Firms and Employees*. University of Chicago Press, Chicago.

- Bielby, W. T. (2000). Minimizing workplace gender and racial bias. *Contemporary Sociology*, 29:120–129.
- Bielby, W. T. and Baron, J. N. (1984). A woman's place is with other women: Sex segregation within organizations. In Reskin, B., editor, *Sex Segregation in the Workplace: Trends, Explanations, Remedies*, pages 27–55. National Academy Press, Washington, DC.
- Bielby, W. T. and Baron, J. N. (1986). Men and women at work: Sex segregation and statistical discrimination. *American Journal of Sociology*, 91:759–799.
- Black, G., Mason, K., and Cole, G. (1996). Consumer preferences and employment discrimination. *International Advances in Economic Research*, 2:137–145.
- Bunderson, S. J. and Sutcliffe, K. M. (2002). Comparing alternative conceptualizations of functional diversity in management teams: Process and performance effects. *Academy of Management Journal*, 45:875–893.
- Cain, G. (1986). The economic analysis of labor market discrimination: A survey. In Ashenfelter, O. and Layard, R., editors, *Handbook of Labor Economics*, volume 1, pages 693–785. Elsevier, Amsterdam.
- Cappelli, P. (2001). Assessing the decline of internal labor markets. In Berg, I. and Kalleberg, A. L., editors, *Sourcebook of Labor Markets: Evolving Structures and Processes*, pages 207–245. Kluwer Academic/Plenum Publishers, New York.
- Card, D., Heining, A. R. C. J., and Kline, P. (2016). Firms and labor market inequality: Evidence and some theory. California–Berkeley Department of Economics Working Paper.
- Carrington, W. J. and Troske, K. R. (1998). Interfirm racial segregation and the black/white wage gap. *Journal of Labor Economics*, 16(2):231–260.
- Castilla, E. J. (2008). Gender, race, and meritocracy in organizational careers. *American Journal of Sociology*, 113(6):1479–1526.
- Catanzarite, L. (2000). Brown-collar jobs: Occupational segregation and earnings of recent-immigrant latinos. *Sociological Perspectives*, 43(1):45–75.
- Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014). Where is the land of opportunity: The geography of intergenerational mobility in the United States. *Quarterly Journal of Economics*, 129(4):1553–1623.
- Cobb, J. A. and Lin, K.-H. (2017). Growing apart: The changing firm-size wage premium and its inequality consequences. *Organization Science*, 28(3):429–446.
- Cobb, J. A. and Stevens, F. G. (2016). These unequal states: Corporate organization and income inequality in the United States. *Administrative Science Quarterly*, 62(2):304–340.
- Cohen, P. N., Huffman, M. L., and Knauer, S. (2009). Stalled progress? Gender segregation and wage inequality among managers, 1980–2000. *Work and Occupations*, 36(4):318–342.
- Conceição, P. and Ferreira, P. (2000). The young person's guide to the Theil index: Suggesting intuitive interpretation and exploring analytical applications. Technical Report 14, University of Texas Inequality Project. Working Paper.
- Conceição, P. and Galbraith, J. K. (1998). Constructing long and dense time-series of inequality using the Theil index. Technical Report 1, University of Texas Inequality Project. Working Paper.
- Conceição, P., Galbraith, J. K., and Bradford, P. (2001). The Theil index in sequences of nested and hierarchic grouping structures. *Eastern Economic Journal*, 27(4):491–514.
- Cowell, F. A. (1985). Multilevel decomposition of Theil's index of inequality. *Review of Income and Wealth*, 31(2):201–205.
- Davis, G. F. (2013). After the corporation. *Politics and Society*, 41:283–308.
- Decker, R., Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2013). The secular decline in business dynamism in the U.S. University of Maryland Working paper.

- Dezsó, C. L., Ross, D. G., and Uribe, J. (2016). Is there an implicit quota on women in top management? A large-sample statistical analysis. *Strategic Management Journal*, 37(1):98–115.
- Duncan, O. D. and Duncan, B. (1955). A methodological analysis of segregation indexes. *American Sociological Review*, 20(2):210–217.
- Dwyer, R. E. (2013). The care economy? gender, economic restructuring, and job polarization in the U.S. labor market. *American Sociological Review*, 78(3):390–416.
- Fairlie, R. W. and Robb, A. M. (2007). Why are black-owned businesses less successful than white-owned businesses? The role of families, inheritances, and business human capital. *Journal of Labor Economics*, 25(2):289–323.
- Ferguson, J.-P. (2015). The control of managerial discretion: Evidence from unionization’s impact on employment segregation. *American Journal of Sociology*, 121(3):675–721.
- Ferguson, J.-P. (2016). Racial diversity and union organizing in the United States, 1999–2008. *Industrial and Labor Relations Review*, 69(1):53–83.
- Ferguson, J.-P., Snellman, K., and Cummins, M. (2017). Firm relocation and spatial mismatch: Evidence from natural disasters. Stanford GSB Working Paper.
- Fernandez, R. M. (2006). Race, spatial mismatch, and job accessibility: Evidence from a plant relocation. Working Paper.
- Fernandez, R. M. (2015). Does competition drive out discrimination? MIT-Sloan Working Paper.
- Fernandez, R. M. and Fernandez-Mateo, I. (2006). Networks, race and hiring. *American Sociological Review*, 71(1):42–71.
- Fernandez, R. M. and Friedrich, C. (2011). Gender sorting at the application interface. *Industrial Relations*, 50(4):591–609.
- Fernandez, R. M. and Weinberg, N. (1997). Sifting and sorting: Personal contacts and hiring in a retail bank. *American Sociological Review*, 62(6):883–902.
- Fernandez-Mateo, I. and Fernandez, R. M. (2016). Bending the pipeline? Executive search and gender inequality in hiring for top managerial jobs. *Management Science*.
- Furman, J. (2016). Beyond antitrust: The role of competition policy in promoting inclusive growth. Technical report, Council of Economic Advisers. Searle Center Conference on Antitrust Economics and Competition.
- Gelman, A. (2015). Age adjustment mortality update. *Statistical Modeling, Causal Inference, and Social Science (blog)*. [andrewgelman.com/2015/11/06/age-adjustment-mortality-update](http://andrewgelman.com/2015/11/06/age-adjustment-mortality-update).
- Gini, C. (1921). Measurement of inequality of incomes. *The Economic Journal*, 31(121):124–126.
- Goldin, C. and Katz, L. F. (2010). *The Race between Education and Technology*. Harvard University Press, Cambridge, MA.
- Gordon, R. J. (2016). *The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War*. Princeton University Press, Princeton.
- Gradín, C., Del Río, C., and Alonso-Villar, O. (2015). Occupational segregation by race and ethnicity in the United States: Differences across states. *Regional Studies*, 49(10):1621–1638.
- Grigoryeva, A. and Ruef, M. (2015). The historical demography of racial segregation. *American Sociological Review*, 80(4):814–842.
- Grusky, D. B. and Charles, M. (1998). The past, present, and future of sex segregation methodology. *Demography*, 35:497–504.
- Hannan, M. T. and Freeman, J. (1984). Structural inertia and organizational change. *American Sociological Review*, 49:149–164.

- Hellerstein, J., Neumark, D., and McInerney, M. (2008). Changes in workplace segregation in the United States between 1990 and 2000: Evidence from matched employer-employee data. In Bender, S., Lane, J., Shaw, K. L., Andersson, F., and Von Wachter, T., editors, *The Analysis of Firms and Employees*, chapter 5. University of Chicago Press, Chicago.
- Hellerstein, J. K. and Neumark, D. (2008). Workplace segregation in the United States: Race, ethnicity, and skill. *Review of Economics and Statistics*, 90(3):459–477.
- Herring, C. (2009). Does diversity pay? Race, gender, and the business case for diversity. *American Sociological Review*, 74(2):208–224.
- Hirsch, B. T. and Macpherson, D. A. (2004). Wages, sorting on skill, and the racial composition of jobs. *Journal of Labor Economics*, 22(1):189–201.
- Irwin, N. (2017). To understand rising inequality, consider the janitors at two top companies, then and now. *New York Times*.
- Kain, J. F. (1992). The spatial mismatch hypothesis: Three decades later. *Housing Policy Debate*, 2:371–460.
- Kalev, A., Dobbin, F., and Kelly, E. (2006). Best practice or best guesses? Diversity management and the remediation of inequality. *American Sociological Review*, 71:589–617.
- Kanter, R. (1977). *Men and Women of the Corporation*. Basic Books, New York.
- Kaufman, R. L. (2010). *Race, Gender, and the Labor Market: Inequalities at Work*. Lynne Rienner Publishers, Boulder, CO.
- Kim, S. K., Arthurs, J. D., Sahaym, A., and Cullen, J. B. (2013). Search behavior of the diversified firm: The impact of fit on innovation. *Strategic Management Journal*, 34:999–1009.
- King, M. C. (1992). Occupational segregation by race and sex, 1940–1988. *Monthly Labor Review*, pages 30–37.
- Kochan, T. A., Bezrukova, K., Ely, R., Jackson, S., Joshi, A., Jehn, K., Leonard, J., Levine, D., and Thomas, D. (2003). The effects of diversity on business performance: Report of the diversity research network. *Human Resource Management*, 42:3–21.
- Kogut, B., Walker, G., and Anand, J. (2002). Agency and institutions: National divergences in diversification behavior. *Organization Science*, 13(2):162–178.
- Leonard, J. S. (1985). The effect of unions on the employment of blacks, hispanics, and women. *Industrial and Labor Relations Review*, 39(1):115–132.
- Levy, P. B. (1994). *The New Left and Labor in the 1960s*. University of Illinois Press, Champaign, IL.
- Lichter, D. T., Parisi, D., and Taquino, M. C. (2015). Toward a new macro-segregation? Decomposing segregation within and between metropolitan cities and suburbs. *American Sociological Review*, 80(4):843–873.
- Lieberson, S. (1980). *A Piece of the Pie: Blacks and White Immigrants Since 1880*. University of California Press, Berkeley, CA.
- Loewen, J. W. (2005). *Sundown Towns: A Hidden Dimension of American Racism*. The New Press, New York.
- Massey, D. S. and Denton, N. A. (1988). The dimensions of residential segregation. *Social Forces*, 67(2):281–315.
- Miller, C. (2017). The persistent effect of temporary affirmative action. *American Economic Journal: Applied Economics*, 9(3):152–190.
- Mouw, T. and Kalleberg, A. L. (2010). Occupations and wage inequality in the United States. *American Sociological Review*, 75(3):402–431.
- Nelson, B. (2001). *Divided We Stand: American Workers and the Struggle for Black Equality*. Princeton University Press, Princeton.

- Olzak, S. (1989). Labor unrest, immigration, and ethnic conflict in urban America, 1880–1914. *American Journal of Sociology*, 94:1303–1333.
- Pager, D. (2016). Are firms that discriminate more likely to go out of business? *Sociological Science*, 3:849–859.
- Petersen, T. and Saporta, I. (2004). The opportunity structure for discrimination. *American Journal of Sociology*, 109(4):852–901.
- Reardon, S. F. and Firebaugh, G. (2002). Measures of multigroup segregation. *Sociological Methodology*, 32:33–67.
- Reskin, B. F. (2000). The proximate causes of employment discrimination. *Contemporary Sociology*, 29:319–328.
- Reskin, B. F., McBrier, D. B., and Kmec, J. A. (1999). The determinants and consequences of workplace sex and race composition. *Annual Review of Sociology*, 25:335–361.
- Robinson, C., Taylor, T., Tomaskovic-Devey, D., Zimmer, C., and Irvine Jr., M. W. (2005). Studying race/ethnic and sex segregation at the establishment level: Methodological issues and substantive opportunities using EEO-1 reports. *Work and Occupations*, 32:5–38.
- Roediger, D. R. (1991). *The Wages of Whiteness: Race and the Making of the American Working Class*. Verso, New York.
- Rosenfeld, J. and Kleykamp, M. (2012). Organized labor and racial wage inequality in the United States. *American Journal of Sociology*, 117(5):1460–1502.
- Rubineau, B. and Fernandez, R. M. (2013). Missing links: Referrer behavior and job segregation. *Management Science*, 59(11):2470–2489.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423.
- Skaggs, S. (2009). Legal-political pressures and African American access to managerial jobs. *American Sociological Review*, 74:225–244.
- Song, J., Price, D. J., Guvenen, F., Bloom, N., and von Wachter, T. (2015). Firming up inequality. Technical Report 21199, NBER.
- Sørensen, A. B. (1977). The structure of inequality and the process of attainment. *American Sociological Review*, 42(6):965–978.
- Sørensen, J. B. (2004). The organizational demography of racial employment segregation. *American Journal of Sociology*, 110(3):626–671.
- Stainback, K. and Tomaskovic-Devey, D. (2012). *Documenting Desegregation: Racial and Gender Segregation in Private-Sector Employment Since the Civil Rights Act*. Russel Sage, New York.
- Stinchcombe, A. L., McDill, M. S., and Walker, D. R. (1968). Demography of organizations. *American Journal of Sociology*, 74(3):221–229.
- Stojmenovska, D., Bol, T., and Leopold, T. (2017). Does diversity pay? A replication of Herring (2009). *American Sociological Review*, 82(4):857–867.
- Stolberg, S. G. (2014). In Georgia, politics moves past just black and white. *The New York Times*.
- Stromquist, S. and Bergman, M., editors (1997). *Unionizing the Jungles: Labor and Community in the Twentieth-Century Meatpacking Industry*. University of Iowa Press, Iowa City, IA.
- Sugrue, T. J. (1996). *The Origins of the Urban Crisis: Race and Inequality in Postwar Detroit*. Princeton University Press, Princeton, NJ.
- Theil, H. (1972). *Statistical Decomposition Analysis*. North-Holland Publishing Company, Amsterdam.
- Tomaskovic-Devey, D. and Skaggs, S. (1999). Degendered jobs? Organizational processes and sex segregated employment. *Research in Social Stratification and Mobility*, 17:139–172.



Tomaskovic-Devey, D., Zimmer, C., Stainback, K., Robinson, C., Taylor, T., and McTague, T. (2006). Documenting desegregation: Segregation in American workplaces by race, ethnicity, and sex, 1966–2003. *American Sociological Review*, 71:565–588.

Wilson, W. J. (1996). *When Work Disappears: The World of the New Urban Poor*. Knopf, New York.

Zuckerman, E. W. (2000). Focusing the corporate product: Securities analysts and de-diversification. *Administrative Science Quarterly*, 45:591–619.