Universidade Vigo

Departamento de Economía Aplicada

http://webs.uvigo.es/x06

Working Paper: 18/02. February 2018

Social Welfare Losses Due to Occupational Segregation by Gender and Race/Ethnicity in the U.S.: Are There Differences across Regions?

Coral del Río

Olga Alonso-Villar

Coordinator: José María Chamorro Rivas
chamorro@uvigo.es

Social Welfare Losses Due to Occupational Segregation by Gender and Race/Ethnicity in the U.S.: Are There Differences across Regions?*

Coral del Río and Olga Alonso-Villar#

Universidade de Vigo

(ECOSOT-ECOBAS and *EQUALITAS*)

Abstract

Taking into account the well-being losses or gains that each gender-race/ethnicity group has associated with its occupational sorting, this paper explores the social welfare loss that each U.S. large region experiences due to the different circumstances faced by these groups in each regional labor market. To analyze the period 1980–2012 in those terms, we use novel measures that aggregate the well-being losses or gains of the groups consistently with the literature on deprivation. To take into account that disparities among regions may arise from differences in characteristics, this papers uses a propensity score procedure that allows controlling for gender and racial/ethnic composition, immigration profile, educational level, and industrial structure.

JEL Classification: D63; R23; J15; J71

Keywords: Occupational segregation; social welfare; gender; race; regions

---

* We gratefully acknowledge financial support from the *Xunta de Galicia* (GRC 2015/014), the *Ministerio de Economía, Industria y Competitividad*, the *Agencia Estatal de Investigación*, and the *Fondo Europeo de Desarrollo Regional* (grants ECO2016-76506-C4-2-R and ECO2017-82241-R).

Correspondence address: Universidade de Vigo; Facultade de CC. Económicas; Departamento de Economía Aplicada; Campus Lagoas-Marcosende s/n; 36310 Vigo; Spain. Tel.: +34 986812507; fax: +34 986812401; e-mail: ovillar@uvigo.es.

1. Introduction

Women and men occupy different positions in labor markets all over the world. In the United States, women tend to be concentrated in jobs characterized by lower wages, less authority, and fewer promotions (Reskin and Bielby, 2005). In 2012, the median weekly earnings of women working full-time was about 81% that of their male counterparts—and only 73% for workers with a bachelor's degree (U.S. Bureau of Labor Statistics, 2013).[1]

It is important to bear in mind that, in explaining this gender pay gap, occupational segregation (i.e., differences in the occupational sorting of women and men) plays an important role (Petersen and Morgan, 1995; Cotter et al., 2003; Del Río and Alonso-Villar, 2015; Blau and Kahn, 2016). Moreover, many scholars claim that occupations are crucial in generating social stratification. Thus, Mouw and Kalleberg (2010) conclude that a large share of the rise in U.S. wage inequality during the period 1992–2008 resulted from such polarization among occupations.

There is an extensive literature on occupational segregation by gender in the United States. Many of these works document a reduction in segregation at the national level in the second half of the 20th century and stagnation at the beginning of the 21st (Beller, 1985; Bianchi and Rytina, 1986; Levanon et al., 2009; Blau et al., 2013). This evolution is usually explained in terms of entry into the workforce of new cohorts of women with higher educational achievements than their predecessors (Blau et al. 2013) and also as a result of the political pressure for gender equality that became a force in the 1970s yet had essentially halted just two decades later (Tomaskovic-Devey et al., 2006). In the current century, occupational segregation by gender remains significant. In 2010, four out of five women working full-time were employed in "feminized" occupations (i.e., those in which at least 75% of the workers were women); and the proportion of men working in "masculinized" occupations was less (50%) but still considerable (Hegewisch et al., 2011).

The literature on occupational segregation has recently turned its attention to segregation by race and ethnicity. Research has shown that segregation between blacks and non-blacks also decreased (at the national level) in the second half of the past

[1] This average percentage of 81% masks the particular situation of the various racial groups, where the corresponding percentage ranges from 71% for Asians to 90% for African Americans. We shall use the terms "black" and "African American" interchangeably.

century whereas segregation between Hispanics and non-Hispanics increased (Queneau, 2009). Many scholars concur that civil rights legislation drove the progress of some of these minorities, as occurred during the 1960s and 1970s for African American women and men and Hispanic women (Conrad, 2005; Tomaskovic-Devey and Stainback, 2007; Kurtulus, 2012). Yet there is no evidence that Hispanic men benefited from affirmative action (Kurtulus, 2012), and the increasing segregation level of this group since the 1980s seems to be the result of its immigration profile (Del Río and Alonso-Villar, 2015).

It is clear that segregation by race/ethnicity does not affect women and men equally. To the contrary, differences in segregation levels among female groups are lower than among male groups (Spriggs and Williams, 1996; Reskin et al., 2004; Alonso-Villar et al., 2012). Furthermore, neither does segregation by gender affect all racial/ethnic groups in the same way: it is higher for Hispanics and lower for Asians than it is for other groups (Hegewisch et al., 2010; Mintz and Krymkowski, 2011). Both gender and race/ethnicity are important traits that help explain the occupational sorting of a group. However, intersectionality (i.e., the fact that the intersection of two or more social identities creates new categories with their own identities) has received little inquiry in the literature on segregation, which has focused mainly on segregation by either gender or race. This paper deals with occupational segregation in an intersectionality framework in which gender is combined with race/ethnicity.

So far, scholars have mostly measured how unevenly distributed demographic groups are across occupations, but little attention has been paid to its consequences in terms of economic welfare—even though the high concentration of a particular group in a few occupations may prove to be favorable or unfavorable for that group depending on the desirability of those occupations. Recently, new tools have been developed to quantify the well-being loss or gain that a group experiences from its occupational sorting (Alonso-Villar and Del Río, 2017a) or the welfare loss suffered by a society, overall, stemming from the different circumstances (vis-à-vis segregation) faced by its constituent groups (Del Río and Alonso-Villar, 2017). This paper will make use of these measures to explore occupational segregation by gender and race/ethnicity in the U.S.

Most studies on U.S. occupational segregation have been undertaken at the national level (typically considering only two groups in the economy, e.g., women vs. men, blacks vs. non-blacks, etc.) but there have been few inquiries into spatial disparities at subnational scales.[2] Yet it is well known that, in different geographical areas, these groups may be exposed to different cultural or social stereotypes and may also face labor markets featuring different industrial structures, demographic composition, and education levels—factors that may facilitate or hinder the integration of some groups into the labor market. So, when comparing different territories (countries, regions, cities, etc.), it is advisable to disentangle segregation disparities due to spatial differences in population and/or industrial composition from disparities that, after one controls for the composition effect, are evidently the consequence of demographic groups experiencing various degrees of labor market integration.

The aim of this paper is to explore social welfare losses arising from occupational segregation by gender and race/ethnicity at the subnational level. Toward that end, we distinguish among four census regions: Northeast, Midwest, South, and West. We assess whether there are meaningful differences among regions on that score and see how segregation-related losses have evolved in each region over the last three decades. Hence we use the 1980, 1990, and 2000 decennial censuses as well as the 2008–12 5-year sample of the American Community Survey (ACS) provided by the Integrated Public Use Microdata Series (IPUMS; Ruggles et al., 2010).

To quantify a region's social welfare loss, we follow two steps. First, we measure the objective well-being loss or gain that each gender–race/ethnicity group in that region experiences as a result of its distribution across occupations. This task is accomplished using tools proposed by Del Río and Alonso-Villar (2015) and Alonso-Villar and Del Río (2017a), where the "quality" of an occupation in a given region is proxied by its relative wage (i.e., the average wage of that occupation in that region divided by the average wage of the region). Second, we aggregate the well-being losses (gains) of the groups via the approach developed in Del Río and Alonso-Villar (2017). This method is similar to the one followed in the literature on deprivation, since a group's well-being loss can be viewed as shortfalls with respect to the case of no segregation.

Our study also explores the causes of observed interregional disparities in social welfare loss. For that purpose, we use the propensity score procedure proposed by DiNardo et al. (1996) in their study of wage discrimination and adapted by Gradín et al. (2015) to explore spatial disparities in occupational segregation. Following Gradín (2013), the

[2] Exceptions include Abrahamson and Sigelman (1987), Lorence (1992), Alonso-Villar et al. (2012, 2013), Gradín et al. (2015), and Alonso-Villar and Del Río (2017b, 2017c).

contribution of each explanatory factor is obtained using the Shapley decomposition (Sastre and Trannoy, 2002; Shorrocks, 2013).

There are four ways in which this paper departs from most studies on segregation. First, we address segregation in a multigroup context by examining 12 different gender–race/ethnicity groups. Second, we measure the well-being loss or gain experienced by each of these groups that results from its occupational sorting and we do so at a regional (not the national) level. Third, for each U.S. census region, we quantify the social welfare loss caused by the occupational sorting of the groups by aggregating their well-being losses in a manner that is consistent with the normative properties widely assumed in the literature on deprivation. Finally, we account for differences in characteristics that may explain disparities in social welfare losses among regions.

The paper proceeds as follows. Section 2 presents the data and methodology, and in Section 3 we describe the social welfare losses of each region and their evolution. Section 4 presents and applies a propensity score procedure to check whether the observed interregional disparities in social welfare loss can be explained by spatial differences in characteristics such as gender and racial composition, educational achievements, immigration profile, and industrial structure. We conclude in Section 5 with a summary and discussion of our findings.

2. Data and Methodology

2.1 Data

The dataset comes from IPUMS-USA (Ruggles et al., 2010). These data are drawn from the U.S. decennial censuses for the period 1980–2000 and the 2008–12 5-year sample of the American Community Survey—which replaced the census long form after 2000 and offers data on occupation. This dataset harmonizes information in that uniform codes are assigned to variables, which makes long-term comparisons possible. The 5-year sample covers 6.9 million workers and includes the two years before and after 2010. The number of workers in the decennial censuses ranges from 5 million in 1980 to 6.4 million in 2000. As mentioned previously, we distinguish the Northeast, Midwest, South, and West census regions.

With respect to the occupational breakdown, we use the consistent long-term classification provided by IPUMS, which is based on the 1990 Census Bureau classification and accounts for 389 job titles. It is necessary to use a detailed classification of occupations because otherwise differences among demographic groups within broad categories of occupations would not be captured and so the measurement of segregation, and its economic consequences, would be underestimated. The wage of each occupation is proxied by the average hourly wage, which is estimated based on reported wages and number of hours worked—after trimming the tails of the hourly wage distribution to prevent data contamination from outliers (for this we eliminate all workers whose wages are either zero, below the 1st percentile or above the 99th percentile of positive values in that occupation).

We consider the 12 mutually exclusive groups of workers that result from crossing women and men with 6 racial/ethnic groups: the four major single-race groups not of Hispanic origin (which we label as whites, African Americans, Asians, and Native Americans); Hispanics irrespective of race (all labeled as Hispanics); and "other races" (non-Hispanics that self-report some other race or more than one race).[3]

2.2 Methodology

Our focus in this paper is measuring the social welfare loss experienced by each U.S. census region resulting from the well-being loss (or gain) that each gender–race/ethnicity group living there has due to its occupational sorting. We start by explaining how to quantify a group's well-being loss or gain; we then describe how best to aggregate the losses and gains of all the groups examined to calculate the social welfare losses of the respective regions.

Measuring the Well-Being Loss or Gain of a Group Arising from its Occupational Sorting

Recall that occupational sorting may advantage or disadvantage a group depending on the relative wages of occupations in which its members are most concentrated. To quantify a group's well-being loss or gain that is associated with its occupational distribution, we use a family of indices (as in Alonso-Villar and Del Río, 2017a) parameterized by $\varepsilon > 0$:

[3] The residual "other race" category is not consistent across years. In particular, multiple-race responses have been allowed only since year 2000.

$$\Psi_{\varepsilon}^{g}=\begin{cases}\sum_{j}\left(\frac{c_{j}^{g}}{C^{g}}-\frac{t_{j}}{T}\right)\frac{\left(\frac{w_{j}}{\overline{w}}\right)^{1-\varepsilon}-1}{1-\varepsilon} & \varepsilon\neq 1,\\ \sum_{j}\left(\frac{c_{j}^{g}}{C^{g}}-\frac{t_{j}}{T}\right)\ln\frac{w_{j}}{\overline{w}} & \varepsilon=1.\end{cases} \qquad (1)$$

Here c_j^g denotes the number of workers of group g in occupation j, t_j is the number of workers in that occupation, $C^g=\sum_j c_j^g$ is the size of the group, $T=\sum_j t_j$ is the total number of workers in the economy, w_j is the (average) wage of occupation j, $\overline{w}=\sum_j \frac{t_j w_j}{T}$ is the average wage of the economy, $\frac{c_j^g}{C^g}$ is the share of group g in occupation j, and $\frac{t_j}{T}$ is the employment share accounted for by that occupation. The "quality" of any occupation j is given by the ratio $\frac{w_j}{\overline{w}}$, which reflects the relative wage of that occupation as compared to other occupations. All these variables refer to the region under study.

The index Ψ_{ε}^{g} quantifies the per capita (objective) well-being loss or gain of group g that is due to its uneven distribution across occupations. This index is equal to zero if either (a) the group is not segregated (i.e., if $\frac{c_j^g}{C^g}=\frac{t_j}{T}$ in every occupation j) or (b) all occupations have the same wage (since in that case the group neither gains any advantage nor suffers any disadvantage from being unevenly distributed across occupations). The index increases (resp. decreases) when some individuals of the group move from one occupation to another that has a higher (resp. lower) wage.

Note that this index depends on parameter ε, which we use to signify aversion toward the inequality that results when some members of the group work in highly paid occupations while others work in low-paid ones. As ε approaches zero (i.e., when there is inequality neutrality), Ψ_{ε}^{g} becomes the Γ^{g} index proposed by Del Río and Alonso-Villar (2015). For simplicity, we consider

$$\Gamma^{g} \equiv \sum_{j}\left(\frac{c_{j}^{g}}{C^{g}}-\frac{t_{j}}{T}\right)\frac{w_{j}}{\overline{w}} \quad (2)$$

as a particular case of the family Ψ_{ε}^{g} when $\varepsilon = 0$—even though inequality aversion, rather than inequality neutrality, is the usual requirement when measuring objective well-being in the economic literature. Our empirical analysis calculates Ψ_{ε}^{g} not only for $\varepsilon = 0$ but also for $\varepsilon = 1$ and $\varepsilon = 2$; these inequality aversion values are standard in the literature on income distribution. The use of these three indices allows for checking the robustness of our findings to differences in aversion to inequality.

The approach just described allows us to transcend the mere measurement of unevenness, on which most segregation analyses focus, to address the economic consequences of that unevenness in terms of (objective) well-being, which is where the main problem lies.

Measuring Welfare Losses of the Whole Society

The measures just defined allow us to focus on the consequences, in terms of well-being, that each demographic group faces as a result of to its occupational sorting. However, those tools are insufficient for determining the welfare loss of an entire region due to segregation of the *n* mutually exclusive groups it comprises. The reason is that some groups may derive gains—while other groups endure losses—stemming from their uneven distribution across occupations.

One way of dealing with this issue would be to calculate the average well-being losses or gains of the groups involved. However, this approach presumes that advantaged groups' gains offset disadvantaged groups' losses of the same magnitude—an assumption that would be called into question by those people who are inequality averse. Hence a more suitable approach to quantifying a region's social welfare loss resulting from the occupational sorting of its demographic groups is to follow, as proposed by Del Río and Alonso-Villar (2017), a framework similar to the one employed in the literature on deprivation and poverty (Jenkins and Lambert, 1997; Shorrocks, 1998). Here the losses due to segregation are viewed as "gaps" with respect to the case of no segregation.

An important difference between the literature on deprivation and the approach followed by Del Río and Alonso-Villar (2017) is that, in the latter, group membership plays a key role, giving individuals an advantage or a disadvantage depending on their respective gender and race/ethnicity—a notion in line with the literature on social stratification (Darity, 2005; Darity et al., 2015). In what follows, we present the curves and indices proposed by Del Río and Alonso-Villar (2017), which combine elements of the literatures on deprivation and social stratification.

To plot the *social welfare loss curve associated with segregation*, these authors first define a vector $d \equiv (d_1,...,d_g,...,d_n)$ that results from giving each group g a value of zero if the group's occupational sorting yields a well-being gain or, otherwise, the absolute value of the group's well-being loss; here n is the number of mutually exclusive groups into which society is partitioned. Namely, $d_g = \left| \min\left\{\Psi_\varepsilon^g, 0\right\} \right|$.[4] The vector d is assumed to be ordered decreasingly, which means that the groups are ranked from high to low levels of loss. The vector representing the demographic size of the groups is denoted by $C \equiv (C^1,...,C^n)$ and $p^k = \frac{C^1+\cdots+C^k}{T}$ ($0 \leq p^k \leq 1$) is the demographic share of the first k groups, where $k = 1,...,n$. Then the *social welfare loss associated with segregation* (WLAS) at point p^k is defined as the sum of the well-being losses of the first k groups, where each group i is weighted by its population share ($\frac{C^i}{T}$). Namely:

$$W^\varepsilon(p^k) = \sum_{i=1}^{k} \frac{C^i}{T} d_i \,, \tag{3}$$

where ε is the same inequality aversion parameter used before to calculate Ψ_ε^g. At intermediate points ($p \in (0,1)$ with $p \neq p^k$), $W^\varepsilon(p)$ is determined by linear interpolation. This curve provides useful information about the social welfare loss of a region due to the occupational sorting of its constituent gender–race/ethnicity groups (see Figure 1).

[4] To simplify notation, parameter ε is not included in vector d.

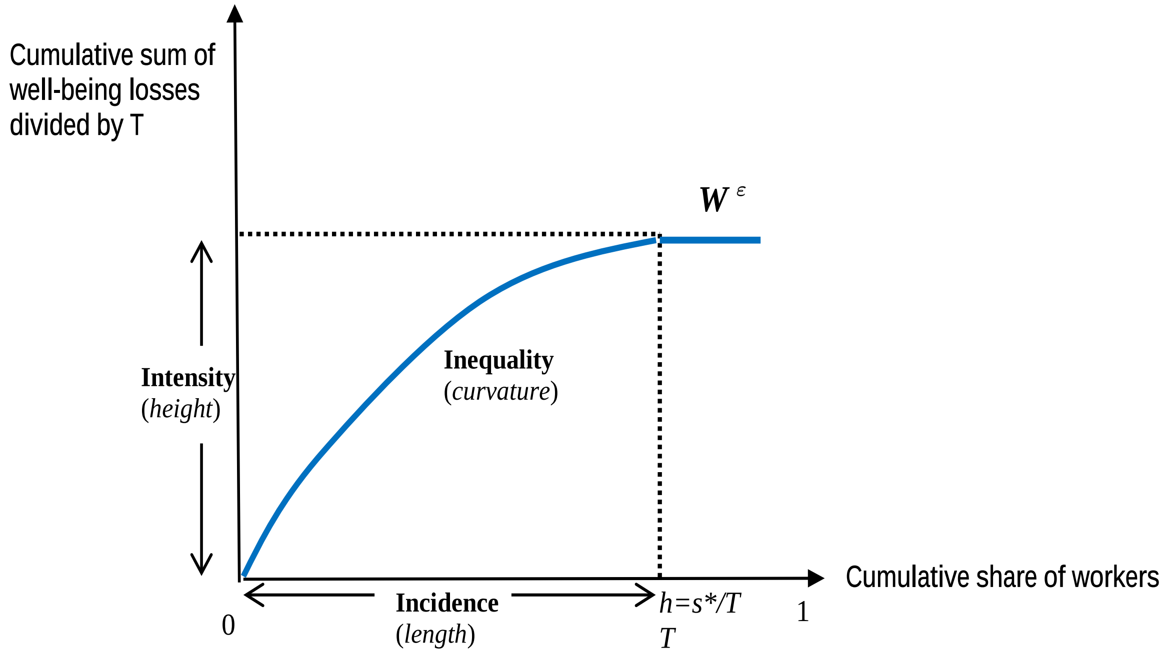

Figure 1. The WLAS curve, W^{ε}
Source: Del Río and Alonso-Villar (2017)

The abscissa value at which the curve becomes horizontal, denoted by $h = s^{*}/T$, represents the incidence of the phenomenon—namely, the population share that the groups with well-being losses account for. The maximum height of the curve conveys the problem's intensity (i.e., the per capita cumulative losses of the groups). Finally, the curvature of the WLAS curve between the origin and point *h* illustrates the inequality in well-being loss among workers in the disadvantaged groups.[5]

These curves are a powerful tool because, when one curve dominates another (i.e., when the former is never above the latter and also is below it at some point) then we can conclude that the social welfare loss in the first situation is lower than that in the second according to a wide range of indices that satisfy basic properties commonly accepted in the literature on poverty and deprivation (Del Río and Alonso-Villar, 2017).

Apart from these curves, to measure the social welfare loss, these authors propose a family of indices that result from adapting the well-known FGT poverty indices developed by Foster et al. (1984) to this context:

$$\mathrm{FGT}_{\alpha}(\tilde{d}) = \frac{1}{T}\sum_{s=1}^{s^{*}}(\tilde{d}_{s})^{\alpha}, \qquad (4)$$

[5] The WLAS curves are based on Jenkins and Lambert's (1997) TIP curves, where TIP stands for "the Three I's of Poverty" (incidence, intensity, and inequality).

where $\alpha \geq 0$ is an inequality aversion parameter associated with the well-being loss inequality among workers belonging to disadvantaged groups (as opposed to inequality aversion, ε, among members of a group), $\tilde{d}_s$ is the well-being loss of worker s (set equal to the per-capita well-being loss of the group to which s belongs), and s^* is the number of individuals for whom $\tilde{d}_s > 0$.

Del Río and Alonso-Villar (2017) show that, when $\alpha > 1$, these indices are consistent with the dominance criterion defined by the WLAS curves. It follows that, when a curve dominates another, we can ensure that with any of these indices the social welfare losses would be lower in the economy represented by the former curve. When no domination exists between the two curves (i.e., if the curves cross) the outcome can change depending on which index is used. Note that index FGT_0 (which represents the proportion of individuals belonging to disadvantaged groups) and index FGT_1 (which measures the average well-being losses of the disadvantaged groups) are not consistent with the WLAS dominance criterion. Nevertheless, our empirical analysis employs both the FGT_0 and FGT_1 indices because they allow measuring the incidence and intensity of the phenomenon separately. Our analysis relies also on the FGT_2 index, which combines the three dimensions of the phenomenon—its incidence, intensity, and inequality among deprived groups—at the same time.

3. Social Welfare Losses by U.S. Regions

We begin the analysis by seeing whether there exist significant differences in the regional social welfare losses associated with the occupational sorting of the gender–race/ethnicity groups that work in each of them. After examining the data at the end of our period of analysis (ACS 2008–12, 5-year sample), we shall analyze the trends observed since 1980 (based on the decennial censuses).

WLAS Curves for Each Region, 2008–12

Figure 2 reveals that the WLAS curve of the Midwest dominates the others (i.e., it is below than or equal to those of the other regions). This means that the Midwest has the country's lowest social welfare losses for a wide range of indices (in particular, all FGT_α indices for which $\alpha > 1$). At the same time, the WLAS curve of the Northeast indicates social welfare losses that are only slightly greater than those in the Midwest, at least in terms of the intensity of the well-being losses (per capita well-being losses of disadvantaged groups) and of the incidence of this phenomenon (percentage of workers who are members of disadvantaged groups). Yet the WLAS curve of the Northeast exhibits a much greater curvature than that of the Midwest, which suggests that the difference between these regions are mainly the result of larger discrepancies in well-being losses among deprived groups in the Northeast than in the Midwest.

Figure 2 also shows that the West's WLAS is clearly dominated by that for the other regions, which implies that social welfare losses are the greatest in this U.S. region according to many indices. Observe that, in the West, the population share belonging to groups with losses is substantially lower than in the other regions (32% vs. more than 50%). The reason of this is that the West is the only region where white women had well-being gains at the end of the study period (see Figure A4 in the Appendix). The West's being dominated occurs because there the per capita well-being losses of disadvantaged groups—that is, the phenomenon's intensity—far exceeded that in the other regions. Finally, we remark that a ranking between the South and the Northeast is not possible because the curves intersect, although the intensity is clearly higher in the South.[6]

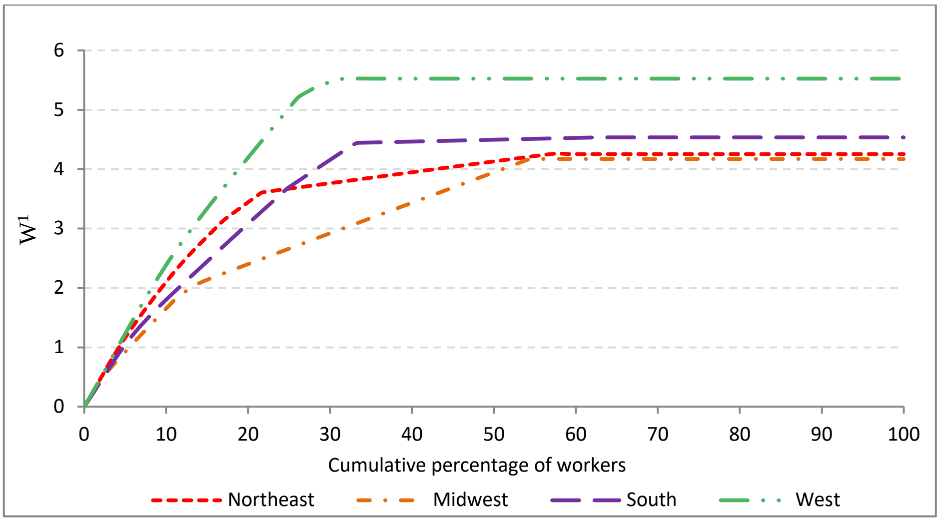

Figure 2. The WLAS curve by region, 2008–12 ($\varepsilon = 1$)

[6] The chart looks substantially the same when the inequality aversion parameter, ε, is set to 0 or 2: the WLAS curve of the West is still dominated by those for the other regions, the South and Northeast curves intersect, and the curve of the Midwest dominates those for the West and South. When $\varepsilon = 2$, the curves of the Midwest and Northeast intersect—with the main change being that the phenomenon's intensity is slightly higher in the Midwest.

Figure 3, which plots the WLAS for all regions in 1980, reveals a considerably different scenario.[7] Here, all curves cross, so we are unable to determine which regions are better-off or worse-off. In fact, when the cumulative percentage of disadvantaged workers reaches about 20%, the four curves are remarkably similar. The implication is that, if we ignore all groups except those with the largest per capita well-being losses, then the four regions would be virtually indistinguishable. Only when the cumulative percentage of such workers exceeds 20% does the curve of the Midwest start to deviate from those for the other regions, which is indicative of a more severe problem in that region. More specifically, this means that some disadvantaged groups did have greater losses in this region and so, when they are lumped with workers in the second quintile, the cumulative losses are higher than those in the other regions. Thus in 1980, the phenomenon was most intense in the Midwest region. Even so, its incidence was lower in that region: the percentage of individuals in disadvantaged groups was slightly lower in the Midwest than in the three other regions (i.e., the WLAS curve becomes horizontal at a more leftward part of the graph). The chart also shows that, in 1980, the curves of the other three regions differed little from each other.

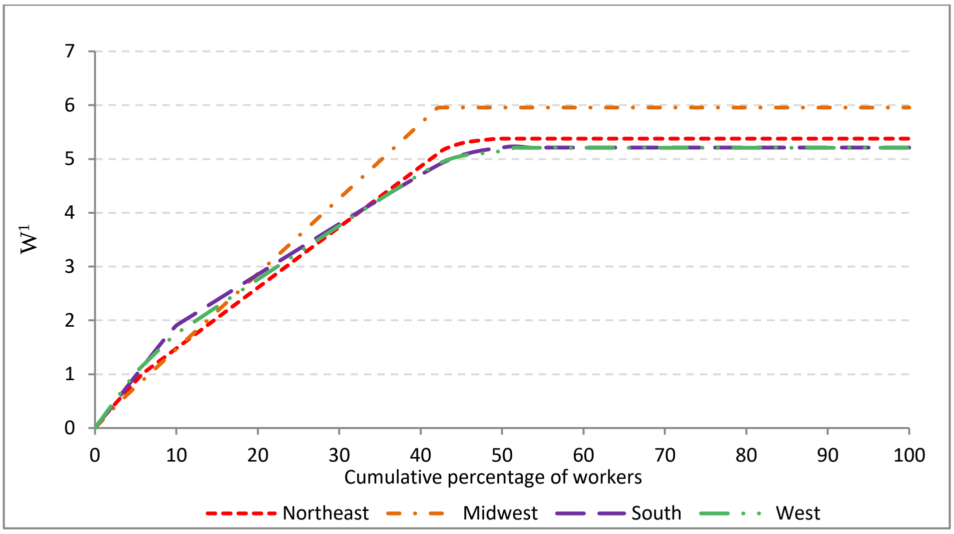

Figure 3. The WLAS curve by region, 1980 ($\varepsilon = 1$)

FGT Indices for Each Region over the Period 1980–2012

Figure 4 illustrates the evolution of the FGT_2 index for each region when $\varepsilon = 1$.[8] As mentioned earlier, the FGT_2 index measures a region's social welfare loss due to the

[7] These graphs are likewise similar when $\varepsilon = 0$ or $\varepsilon = 2$.

[8] The graphs for $\varepsilon = 0$ and 2 are almost identical to Figure 4.

occupational sorting of its 12 gender–race/ethnicity groups when one accounts not only for the incidence (FGT_0) and intensity (FGT_1) of the phenomenon but also the disparities that exist among the losses of the individuals belonging to the disadvantaged groups. The evolution of the FGT_0 (resp. FGT_1) index is plotted in the left (resp. right) panel of Figure 5.[9] The values of the three FGT indices under the different settings of the aversion parameter ε are given in Table A2 of the Appendix.

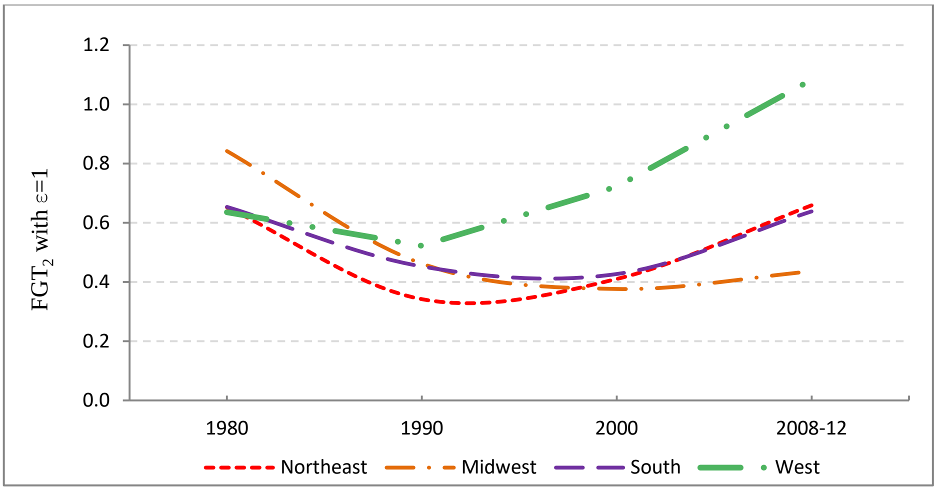

Figure 4. FGT_2 index by region ($\varepsilon = 1$)

The FGT_2 index exhibits a U-shape trend in every region, as also happens at the national level (Del Río and Alonso-Villar, 2017). Thus, all regions saw a decreasing index at the start of the study period, but in each case the index eventually bottomed out and then began rising. Despite those similarities, there were significant interregional differences in the index value and also in its evolution. First, the Midwest began with an FGT_2 index above that of the other three regions, which shared a similar starting point. Second, the Midwest's U-shape evolution was smoother than elsewhere, which gave that region the lowest FGT_2 index in 2008–12; this observation accords with its WLAS curve dominating the others, as mentioned previously. Third, the differences among the regions were much greater in the last decade than in previous ones. In 2008–12, the West had an FGT_2 index more than double that of the Midwest, while the Northeast and South shared an intermediate value.

[9] The graphs for $\varepsilon = 0$ and 2 are analogous.

Hence, the Midwest, which is less racially diverse than the other regions (see Table A1 in the Appendix), improved its relative position in terms of the FGT_2 index; it went from being the region with the highest social welfare losses in 1980 to the one with the smallest such losses in 2008–12.[10] This transformation resulted, at least in part, from the remarkable reduction in the per capita well-being losses associated with the occupational sorting of the groups (see FGT_1 in Figure 5).[11] Note that white women in the Midwest accounted for a larger (and increasing) share of workers than in the other regions, so how they fared has an important effect on the region's losses. In 1980, the greatest well-being losses for white women did, in fact, occur in the Midwest (see Figures A1–A4 in the Appendix). Although this comparison has remained stable over time, the actual amount of these losses has decreased considerably since 1980. On the other hand, the relatively small group of Asian women experienced notable occupational advances in the period; thus they were experiencing well-being gains (rather than losses) by the 1990s and, as a consequence, no longer constituted a disadvantaged group.

However, the experience of male groups in the Midwest does not conform to reductions in the FGT_2 index. African Americans and Hispanics worsened over the period, as they shifted from having small well-being gains in 1980 to experiencing losses soon thereafter; this trend increased both the incidence and the intensity of the problem. Because the other main racial groups (whites and Asians) were never disadvantaged groups in the period of analysis, they had no effect on the FGT_2 index.[12]

[10] Although not shown here, the Midwest´s WLAS curve for 1980 is dominated by its curve for 2000, which implies that the improvement experienced in this region during the first two decades is robust to changes in the particular index used. In other words, we obtain the same result not only when using the FGT_2 index but also when using any FGT_α index for which $\alpha > 1$. Yet from 2000 onward, we can draw no conclusive results because the WLAS curves cross: the outcome depends on which FGT index is used.

[11] However, the incidence of the problem (as captured by FGT_0) increased slightly throughout the period (Figure 5 and Table A2). The reason was the rise in the share of two disadvantaged groups, Hispanic women and Hispanic men (Table A1).

[12] Note that African American and Hispanic men accounted for only 7% of all workers in 2008–12, which may explain why their worsening situation did not prevent the index from decreasing.

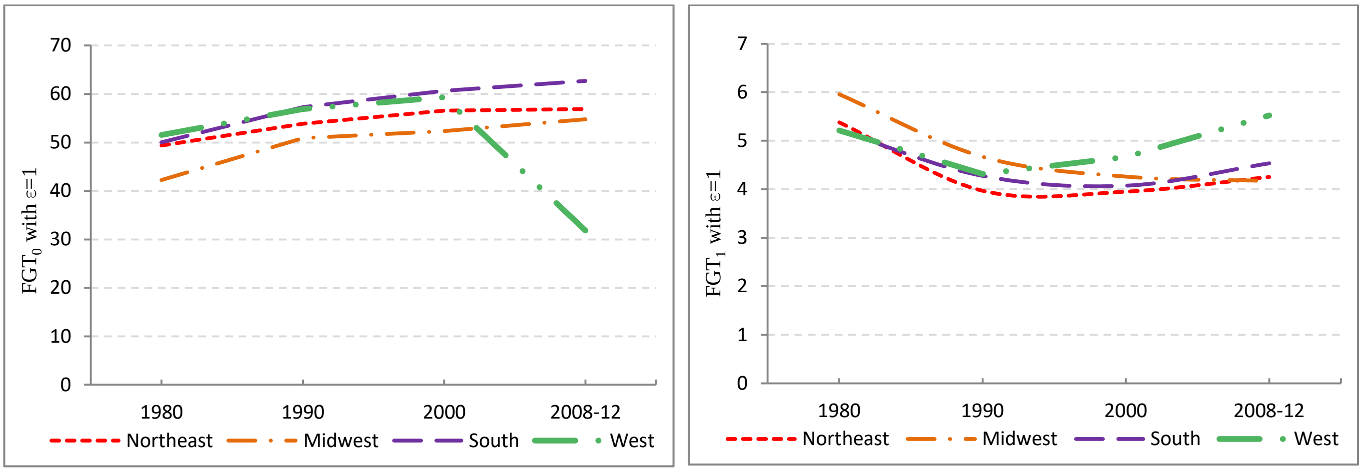

Figure 5. Indices FGT_0 and FGT_1 by region ($\varepsilon = 1$)

In the West region, the social welfare losses in terms of the FGT_2 index not only became greater after 1990 (unlike the other regions, where that process started later) but were also persistently much higher than those of the other regions (Figure 4).[13] Note, however, that the problem in this region over the last decade was not a rise in the incidence of the phenomenon (see FGT_0 in Figure 5); in fact, the percentage of workers belonging to groups experiencing well-being losses due to their occupational sorting actually declined (for $\varepsilon = 1$, those proportions were 59% in 2000 and 32% in 2008–12). The reason for this reduced percentage of disadvantaged workers may be that white (and Asian) women began to experience small well-being gains in the 2000s (Figure A4). The disparities among workers belonging to groups with well-being losses in 2008–12 cannot help us to explain why the FGT_2 index is higher in the West than elsewhere.[14] Rather, it was the continuous increase in the initially high share of Hispanic women and men in the West—two disadvantaged groups that experienced greater losses in this period (Figure A4)—that seems to explain the problem's increasing intensity in this region from 1990 onward (see FGT_1 in Figure 5). This

[13] Although not shown in the paper, the West's WLAS curve for 1990 dominates its curve for 2000, which in turn dominates the curve for 2008–12. These results confirm the robustness of our finding. The West's WLAS curve for 1980 also dominates that for 2008–12. So the problem in the West is more severe now than in 1980 according not only to the FGT_2 indices mentioned before but also to a wide range of indices (those consistent with the dominance criterion of the WLAS curves).

[14] As seen in Figure 2, the WLAS curve for this region is almost a straight line in the increasing part of the curve, which implies that the per capita losses in well-being of the disadvantaged groups are quite similar. In fact, the coefficient of variation for these losses is 0.36 in the West, as compared with 0.61 in the Midwest, 0.97 in the South, and 1.04 in the Northeast.

increase appears to have more than offset the positive effect of the evolution of the white and Asian women.

As for the Northeast and South regions, the FGT_2 index evolved similarly in both regions—in that the values were similar at both the beginning and end of the period—but the U-shape is more pronounced in the Northeast because of a sharper fall during the 1980s.[15] This reduction in losses seems to arise from a stronger decrease in the well-being losses of African American women in the Northeast during the first decade (see Figures A1 and A3). This improvement did not last long though: the losses suffered by these women had already increased slightly in 2000, and the process continued (with increasing intensity) until the end of the period. In the South, however, African American women had in 1980 greater well-being losses than they did in the Northeast (and also in the other regions) but these losses became less at a fairly steady rate so that, by the end of the period, these women caught up with their counterparts in the Northeast.

4. Controlling for Regional Characteristics

The analysis so far has revealed substantial disparities among regions as regards losses in social welfare due to the occupational sorting of their respective gender–race/ethnicity groups. However, these differences could arise not only because some demographic groups may find it more difficult to secure “good” jobs in some regions than in others but also because of regional differences in such factors as gender and racial/ethnic composition, immigration profile (including both years of U.S. residence and English proficiency), and education levels, all of which could affect the availability of occupations. A region’s industrial composition could also affect the occupational sorting of our demographic groups by altering the number of job openings in occupations traditionally associated with a group members’ employment.

The main question we pose in this section is whether the regional disparities in social welfare losses would remain if there were no differences in the characteristics just mentioned. To address this question, we take a “reference region” and then build, for each of three other regions, a counterfactual economy such that the share of each

[15] In the Northeast, the WLAS curve for 1990 dominates that for 1980, which corroborates the region’s improvement during the first decade for a wide range of indices. However, since 1990 this trend has reversed: the earlier decade’s WLAS curve dominates that of the later decade’s. In the South, the curve for 1980 is dominated by the one for 2000; thus the improvement in this region lasted for two decades.

subgroup defined by the combination of those characteristics is the same in all regions—but with the occupational distribution of each subgroup unchanged from what we observe in the actual data. We refer to the social welfare loss calculated using this counterfactual distribution as the conditional social welfare loss, and it represents the social welfare loss that each region would experience if there were no regional differences in characteristics. When a region's conditional social welfare loss is strongly similar to its unconditional loss, we can surmise that the difference between that region and the reference region does not result from differences in characteristics but rather from differences in the extent to which some gender–race/ethnicity groups are integrated into the labor market. When instead there is a significant difference between the conditional and unconditional welfare losses, it is almost certain that regional characteristics account (at least in part) for such regional disparities. In this latter case, we could also identify the main explanatory factors. Next we present the methodology used to calculate the conditional social welfare loss, after which we report our findings.

4.1 Propensity Score Procedure

In the empirical analysis, we "homogenize" the four regions according to six key characteristics that differ among regions and may help to explain the observed regional disparities in social welfare losses: (i) gender (2 groups); (ii) racial/ethnic composition (5 groups: non-Hispanic whites, African Americans, and Asians, Hispanics of any race, and others);[16] (iii) years of U.S. residence (3 categories: born in the U.S., resided there up to 10 years, and resided there for more than 10 years); (iv) English proficiency (4 categories: speaking only English, speaking English very well, well, and not well or not at all); (v) educational achievements (4 levels: less than high school, high school diploma, some college, and bachelor's degree); and (vi) industrial structure (11 sectors).[17] These are the characteristics or attributes to which we refer hereafter.

[16] Because of their small group size, Native Americans were subsumed within the group of individuals from "other" races.

[17] The sectors are: "agriculture, forestry, fisheries, and mining"; "construction"; "manufacturing-1" (which includes some durable goods: metal industries; machinery and computing equipment; electrical machinery, equipment, and supplies; transportation equipment; professional and photographic equipment; and watches); "manufacturing-2" (which includes nondurable goods and the remaining durable goods); "transportation, communications, other public utilities and wholesale trade"; "retail trade"; "finance, insurance, and real estate"; "business and repair services"; "personal services, and entertainment and recreation services"; "professional and related services"; and "public administration and active duty military".

The propensity score procedure, initially proposed in the context of wage discrimination by Di Nardo et al. (1996) and adapted by Gradín et al. (2015) to explore spatial disparities in occupational segregation, consists of building a counterfactual distribution for each region so that each “cell” or subgroup resulting from the “crossing” of the main attributes mentioned above (e.g., Asian immigrants who have lived up to 10 years in the U.S., speak English very well, have a university degree, and work in the professional services sector) has the same weight in all regions whereas the occupational sorting of that subgroup is kept unaltered (i.e., it is the one we observe in the data). This procedure requires that we first take a reference region with respect to which the remaining regions will be homogenized. Suppose, for example, that the region of reference is the South. We must then reweight the original observations from the other regions by the probability (as predicted by a logit model) that each worker—who has specific attributes in terms of gender, race/ethnicity, immigration profile, and education, and works in a certain sector—resides in the South rather than that worker’s own region. To streamline the presentation, we shall explain how to build the counterfactual distribution for a single region: the Midwest.

Let $z \equiv (z_1,...,z_k)$ denote the vector of the *k* covariates describing the attributes of each subgroup, and let *R* be a dummy variable indicating regional membership; thus *R*= *S* for workers living in the South and *R*=*M* for those living in the Midwest. The weighting scheme, Ψ_z, by which we give the Midwest the same characteristics as the South can be estimated from the data as follows, where the vertical bar is shorthand for “conditional on”:

$$\Psi_z = \frac{\frac{\Pr(R=S|z)}{\Pr(R=S)}}{\frac{\Pr(R=M|z)}{\Pr(R=M)}} = \frac{\Pr(R=M)}{\Pr(R=S)} \frac{\Pr(R=S|z)}{\Pr(R=M|z)} .$$

The first term can be approximated by the ratio of the Midwest’s population to the South’s population samples. The second term can be obtained by estimating the probability of an individual with attributes *z* residing in the South (rather than the Midwest). For that estimation, we use a logit model over the pooled sample of observations from both regions:

$$\Pr(R=S|z) = \frac{\exp(z\hat{\beta})}{1+\exp(z\hat{\beta})},$$

where $\hat{\beta}$ is the associated vector of estimated coefficients.

We employ this procedure to construct a counterfactual economy in the Midwest. Then we can calculate the WLAS curve and the FGT_α indices for this economy and compare them with those based on our data.

The difference between a conditional FGT_α index for the Midwest derived from our counterfactual distribution and the one obtained using the real distribution gives us a measure of the difference in social welfare loss between the Midwest and the South that is explained by our vector *z* of covariates. Following Gradín (2013), this explained part can be further disaggregated into the respective contributions of each factor (which can be either a single covariate or a set of covariates) via the Shapley decomposition—a technique commonly used in the literature on income distribution (Sastre and Trannoy, 2002; Shorrocks, 2013).[18]

Our conditional analysis yields the welfare loss that the Midwest would have had if it did not differ from the South with regard to gender and racial/ethnic composition, years of residence, educational achievements, and industrial structure. The same procedure is then followed for the West and Northeast regions. Any differences (among the four regions) that remain after the complete conditional analysis give us a picture of the true comparative difficulty encountered, from one region to another, by our gender–race/ethnicity groups when seeking to become integrated into the labor market—that is, because we have controlled for basic characteristics that would distort such conclusions if only the raw data were used.

4.2 Are There Regional Differences in Conditional Welfare Losses?

Figure 6, which displays each U.S. region's conditional WLAS curves in 2008–12 (with the South as the reference region), reveals that the picture changes substantially as

[18] To obtain the contribution of race/ethnicity, for example, we use the logit coefficients as follows. First, we calculate the prediction of $\Pr(R = S|z)$ by assuming that all coefficients except for those of race/ethnicity dummies are zero; then we compare the social welfare loss in the Midwest resulting from this new counterfactual distribution to the social welfare loss with the real distribution. Next, we calculate the prediction of the mentioned probability while assuming zero coefficients for all covariates except for race/ethnicity and one other covariate (e.g., years of U.S. residence). The resulting counterfactual is compared to the counterfactual where only the variable years of U.S. residence is taken into account. The analysis is repeated but with educational achievements (rather than years of U.S. residence) as the other covariate accounted for, and so on. This inform us about the marginal contribution of race/ethnicity when this is the second factor we control for. We continue by following the same procedure while considering all possible sequences where race/ethnicity is the third (rather than the second) factor to change and so on. By averaging over all possible marginal contributions of race/ethnicity, we compute the contribution that this covariate makes to explaining the difference between (i) the Midwest's loss of social welfare under the counterfactual distribution and (ii) its loss under the real distribution.

compared with the unconditional analysis (Figure 2).[19] First, although the West's curve is still dominated by the others, the intensity of that region's social welfare losses no longer differs much from that in those other regions.[20] Second, the Midwest's WLAS curve no longer dominates the others, and its maximum height (which embodies the intensity of losses) is no longer the lowest.[21] Third, and despite these changes, the conditional analysis shows notable differences among regions. More specifically, the phenomenon clearly reaches its lowest intensity in the Northeast. Also, the Northeast's WLAS curve dominates those for the South and West. Thus we conclude that, for a wide range of indices, the social welfare losses are lower in the Northeast than in the South and West once we control for characteristics.[22]

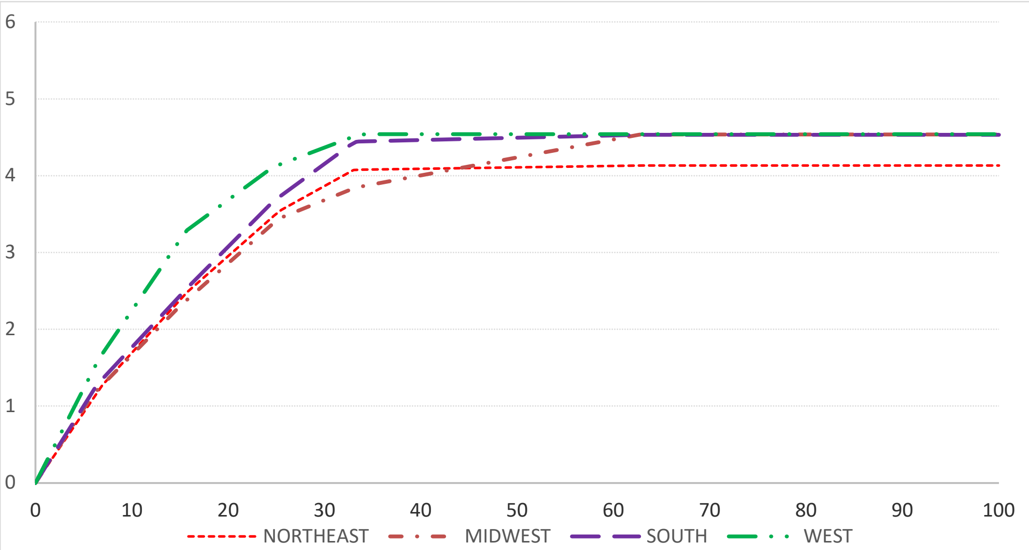

Figure 6. The conditional WLAS curves, 2008–12 (reference region: South; $\varepsilon = 1$)

It follows then that controlling for characteristics reduces some of the interregional disparities in social welfare losses. In particular, according to the coefficient of variation, regional disparities in values of the FGT_2 index—an index that is consistent with the WLAS dominance criterion—decline by 50% when the South is the reference

[19] The coefficients of the logit regressions are shown in Table A3.
[20] When the region of reference is other than the South, we also find that the curves become less distinctive because—when we control for characteristics—the problem's intensity in the West differs less from its intensity in the other regions. Moreover, when either the Midwest or Northeast is used as the reference region, the West's curve intersects some of the other curves, which implies that the West can have a lower social welfare loss than other regions according to some indices.
[21] When reference regions other than the South are used, the Midwest also loses its superior "ranking" because the phenomenon's intensity becomes higher there than in other regions. Moreover, in some cases, the Midwest actually has the highest intensity (although that in the South and West are not much lower).
[22] The results for the Northeast are robust to changing the reference region.

region and $\varepsilon = 1$; with other reference regions, the reduction is even greater.[23] This finding suggests that our covariates explain at least half of the variability in regional welfare losses, even though some regional disparities are not explicable by the characteristics we describe.

The question that now arises is: Which of the characteristics we consider are the most explanatory? To answer that question, we use the FGT_2 index and decompose the change between each region's conditional and unconditional social welfare loss (i.e., the explained part of the difference between a region's unconditional loss and the South's unconditional loss) into the contribution of each factor: gender composition, racial/ethnic composition, immigration profile (which combines the variables of years of U.S. residence and English proficiency), educational achievements, and industrial structure. Figure 7 reports the contribution of these factors—determined via Shapley decomposition—for each region (where the South is again the reference region).

We start by explaining how to interpret this chart. First of all, the South is our reference region and so there is no difference there between the conditional and unconditional social welfare losses. Second, with reference to the figure's horizontal axis, the positive factors (resp. negative) are those that would cause the FGT_2 index to increase (resp. decrease). So, for example, if workers in the West region were of the same educational level as those in the South, then the index would be higher than when calculated using the real (i.e., not counterfactual) distribution. Yet, if the West were characterized by the same racial/ethnic composition, immigration profile, and industrial structure as the South, then the index would be lower than is actually the case.

The figure clearly shows that, in the West, the net effect of all our factors taken together is both negative and large. Therefore, if this region had the same attributes as the South, then, according to the FGT_2 index, its social welfare loss would be 26% lower than what we actually observe. This result is consistent with the West's WLAS curve being much closer to the other regions' curves in Figure 6 than it is in Figure 2; thus there is a significant reduction in intensity when we control for characteristics. Figure 7's outcomes in the Northeast and Midwest—with the former region having a negative net effect and the latter a positive one—are also consistent with Figures 2 and 6 in the sense

[23] Regional differences in FGT_2 values are reduced by nearly 50%, 65%, and 70% when (respectively) the West, Northeast, and Midwest serve as the reference region.

that, when we control for characteristics, intensity (resp. inequality among disadvantaged groups) is slightly (resp. markedly) reduced in the Northeast whereas, in the Midwest, both intensity and incidence are considerably increased.

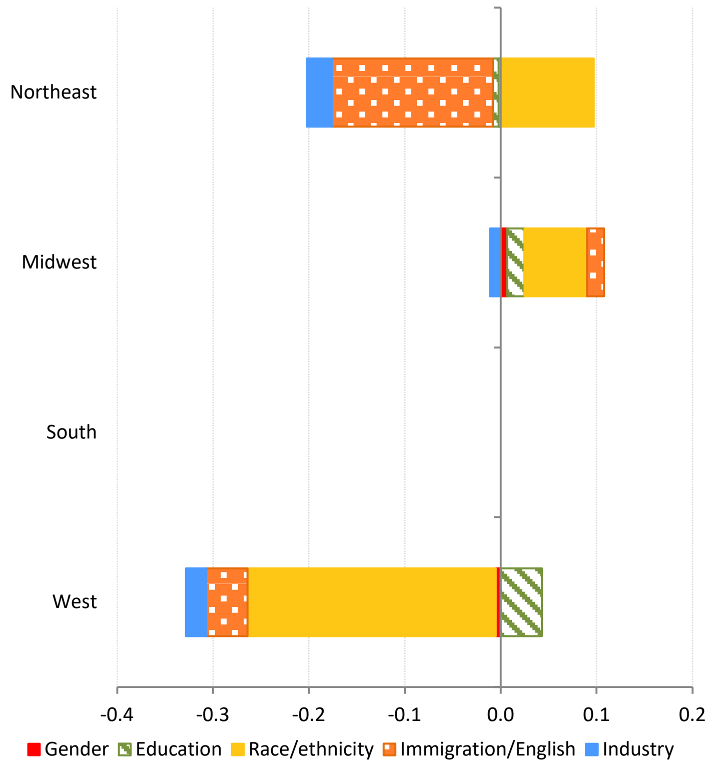

Figure 7. Conditional minus unconditional social welfare loss in each region (FGT_2 index; $\varepsilon = 1$; reference region: South).

We infer from Figure 7 that racial/ethnic composition and also the immigration profile are key drivers of regional disparities in social welfare losses; education achievements and industrial structure play lesser roles.[24] Recall, however, that those spatial differences reflect more than differences in the attributes of our gender–race/ethnicity groups and the regions' industrial structures. The extent to which the groups are integrated into the labor market is not the same across regions.

5. Final Comments

The extant literature on occupational segregation has focused mainly on measuring the aggregate or overall segregation that arises from the occupational sorting of mutually exclusive groups into which a society can be partitioned—typically two groups, as with

[24] Although not shown in the paper, these main outcomes are unaffected by our choice of the reference region.

women and men, or whites and non-whites. However, hardly any scholars have analyzed the implications of segregation for social welfare. If all occupations were of the same quality (as measured, e.g., by wages), then differences in the occupational sorting of groups would not be a source of severe economic problems. Yet different occupations do offer different wages, and they differ also in terms of promotional opportunities, physical risk, required hours per week, and so on. Occupations can be viewed as a mechanism that generates economic inequalities among demographic groups (Mouw and Kalleberg, 2010). Occupational segregation by gender and/or race/ethnicity helps perpetuate those inequalities and the social hierarchies they engender (Browne and Misra, 2003).

In this paper we have addressed the social welfare losses—due to occupational segregation by gender and race/ethnicity—experienced over the last three decades (from 1980 to 2012) in the four census regions of the U.S.: the Midwest, Northeast, South, and West. The analysis has considered 12 mutually exclusive demographic groups and nearly 400 occupational titles. Our findings indicate that the phenomenon is not homogenous across the country and also that regional disparities have increased over time.

In focusing on the per capita well-being gains or losses of the largest racial minorities associated with their occupational sorting, we found that Asian women and men seem to be better-off in the Midwest, which is the region with the least racial diversity. Since 1980, these groups exhibited larger well-being gains (or smaller losses) associated with their occupational sorting than did their white counterparts, although Asian men saw gains throughout the period whereas Asian women did not begin to exhibit gains until the 1990s. Furthermore, the differential in favor of both male and female Asians (as compared with other groups) was greater in the Midwest than in the other regions and also increased over the period. The circumstances of African Americans in the Midwest are quite different. Men of this minority group experienced a dramatic worsening: despite starting with gains (though small ones) in 1980, they suffered large well-being losses associated with their occupational sorting at the end of the period (2008–12). In contrast, the trend for African American women in this region is described by a bell-shaped curve (i.e., the losses first decreased and then increased), although this group's losses were always greater than those of their male peers. As for Hispanic men and women, their initial situation in the Midwest was similar to that of African Americans, although Hispanic well-being losses associated with their occupational sorting increased more steeply over the period.

In the Northeast, the losses incurred by both African American and Hispanic workers of each gender follow a pattern similar to that observed in the Midwest, but the relative advantage of Asians (with respect to whites) was smaller in the Northeast and also a more recent phenomenon there. In the South, where African Americans account for a larger share of workers than in other regions, men of this minority group experienced a less severe worsening; even so, in 2008–12 their losses differed little from those of African American men in the Midwest and Northeast—mainly because it was in the South where this group was worse-off in 1980. The trend in losses for African American women in the South is also distinctive: their well-being losses associated with their occupational sorting, which in 1980 were the greatest in the country among all gender–race/ethnicity groups, steadily declined over the study period. Hence African American women in the South experienced in 2008–12 losses that were not much worse than those of their Northeastern and Midwestern peers. With regard to Hispanic women and men, in the South these groups were initially better-off than African American women and men, especially the men, but these groups' relative position had reversed by the end of the period.

In the West, the situation of African American men in 1980 was quite similar to that in the South. However, in the former region the well-being losses of this group remained relatively stable over time while in the latter region the losses increased. In addition, the well-being losses of African American women in the West, which started by differing little from those by this group in the Northeast and Midwest, declined significantly in the study period's first decade and so despite its later leveling out, in 2008–12 African American women in the West had smaller well-being losses associated with their occupational sorting than in the other three regions.

It is in the West where the well-being losses of Hispanic women and men, who account for a large share of the workers in that region, deviated more dramatically from their African American counterparts—a gap that widened over the period. This is also the only region in which Asians, whose population share was not negligible and increased over the period, exhibited smaller well-being gains (or larger losses) than did whites. However, these differences shrank with each succeeding decade, allowing the Asian groups to catch up with the white groups in 2008–12 (i.e., for women and men both).

Our analysis has revealed that, in all regions, the position of white women improved substantially during the period. However, it is only in the West that this group ultimately made any well-being gains associated with their occupational sorting, and they were small ones. The gains of white men remained relatively stable over time and were similar across the four regions (though slightly more notable in the West since 1990).

If the goal is to compare the social welfare losses across regions then, in addition to deriving the per capita well-being gain or loss of each group, we must also account for the size of each group. Thus a region will be worse-off if disadvantaged groups—those experiencing well-being losses associated with their occupational sorting—are large. When we view trends in the per capita gains or losses of the 12 demographic groups that we track while considering their relative population sizes, it is clear that the interregional disparities with respect to social welfare losses changed dramatically since 1980. In 2008–12, the West (resp. Midwest) was the region with the largest (resp. smallest) losses in social welfare according to a wide range of indices.

However, these results do not imply that the labor integration of women and racial minorities was necessarily more difficult in the West and/or easier in the Midwest. Regional differences may reflect spatial disparities in the attributes of the groups and also disparities in their industrial structures, which may foster or hinder the concentration of the groups in the occupations at which they are traditionally employed. After controlling for regional characteristics such as gender and racial/ethnic composition, educational achievements, workers' years of U.S. residence, English proficiency, and industrial structure (which are captured by 27 indicator variables), we found that at least half of the interregional differences in social welfare loss disappear, although some spatial disparities persist. Hence there may well exist factors, other than those addressed here, that affect the integration of women and racial/ethnic minorities into the labor market; examples include citizens' attitudes toward gender and race, government policies, social capital, and so forth. The role played by these other factors in the regional variation of social welfare losses arising from occupational segregation by gender and race/ethnicity is beyond the scope of this paper, but it offers fruitful avenues for further research on this topic.

References

Abrahamson M. and Sigelman, L. (1987): “Occupational sex segregation in metropolitan areas,” *American Sociological Review* 52, 588-597.

Alonso-Villar, O. and Del Río, C. (2017a): “Occupational Segregation and Well-being,” *Review of Income and Wealth* 63(2), 269–287.

Alonso-Villar, O. and Del Río, C. (2017b): “The Occupational Segregation of African American Women in the United States: A Look at its Evolution from 1940 to 2010,” *Feminist Economics* 23(1), 108-134.

Alonso-Villar, O. and Del Río, C. (2017c): “Mapping the Occupational Segregation of White Women in the U.S.: Differences across Metropolitan Areas,” *Papers in Regional Science* 96(3), 603-625.

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.

Alonso-Villar, O., Gradín, C. and Del Río, C. (2013): "Occupational segregation of Hispanics in U.S. metropolitan areas", *Applied Economics* 45(30), 4298-4307.

Beller, A. H. (1985): “Changes in the Sex Composition of U.S. Occupations, 1960-1981,” *Journal of Human Resources* 20, 235-250.

Bianchi, S. M. and Rytina. N. (1986) “The Decline of Occupational Sex Segregation during the 1970s: Census and CPS Comparisons,” *Demography* 23, 79-86.

Blau, F., Brummund, P. and Liu, A. (2013): “Trends in Occupational Segregation by Gender 1970-2009: Adjusting for the Impact of Changes in the Occupational Coding System,” *Demography* 50, 471-492.

Blau, F. D. and Kahn, L. M. (2016): “The Gender Wage Gap: Extent, Trends, and Explanations,” NBER Working Paper 21913.

Browne, I. and Misra, J. (2003): “The Intersection of Gender and Race in the Labor Market”, *Annual Review of Sociology* 29, 487-513.

Conrad, C. (2005): “Changes in the labor market status of black women, 1960-2000”, in C. Conrad, J. Whitehead, P. Mason, and J. Steward (Eds.), *African Americans in the U.S. economy*, Lanham (MD): Rowman & Littlefield Publishers, 157-162.

Cotter, D., Hermsen, J., and Vanneman, R. (2003): “The effects of occupational gender segregation across race,” *The Sociological Quarterly* 44, 17-36.

Darity Jr., W. A. (2005): “Stratification Economics: The Role of Intergroup Inequality,” *Journal of Economics and Finance* 29 (2), 144-153.

Darity Jr., W.A., Hamilton, D., and Stewart, J.B. (2015): “A Tour de Force in Understanding Intergroup Inequality: An Introduction to Stratification Economics,” *Review of Black Political Economy* 42, 1-6.

Del Río C. and Alonso-Villar, O. (2015): “The Evolution of Occupational Segregation in the U.S., 1940-2010: The Gains and Losses of Gender–Race/Ethnic Groups,” *Demography* 52(3), 967-988.

Del Río, C. and Alonso-Villar, O. (2017): "Segregation and Social Welfare: A Methodological Proposal with an Application to the U.S.", *Social Indicators Research* (forthcoming). DOI: 10.1007/s11205-017-1598-0.

DiNardo, J., Fortin, N., and Lemieux, T. (1996): “Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach,” *Econometrica* 64, 1001-1044.

Foster, J.E., Greer, J., Thorbecke, E. (1984): “A Class of Decomposable Poverty Measures,” *Econometrica* 52(3), 761–766.

Gradín, C. (2013): “Conditional Occupational Segregation of Minorities in the US,” *Journal of Economic Inequality* 11(4), 473-493.

Gradín, C., del Río, C., and Alonso-Villar. O. (2015): “Occupational Segregation by Race and Ethnicity in the US: Differences across States,” *Regional Studies* 49 (10), 1621-1638.

Hegewisch, A., Liepmann, H., Hayes, J. and Hartmann, H. (2010): “Separate and Not Equal? Gender Segregation in the Labor Market and the Gender Wage Gap,” Institute for Women’s Policy Research Briefing Paper, IWPR C377, Washington, D.C.

Hegewisch, A., Willians, C. and Henderson, A. (2011): “The Gender Wage Gap by Occupation,” Institute for Women’s Policy Research Fact Sheet, IWPR C350a, Washington, D.C.

Jenkins, S.P., Lambert, P.J. (1997): “Three ‘I’s of Poverty Curves, with an Analysis of UK Poverty Trends,” *Oxford Economic Papers* 49, 317–327.

Kurtulus, F. (2012): “Affirmative Action and the Occupational Advancement of Minorities and Women during 1973-2003,” *Industrial Relations* 51, 213-246.

Levanon, A., England, P., and Allison, P. (2009): “Occupational Feminization and Pay: Assessing Causal Dynamics Using 1950-2000 U.S. Census Data,” *Social Forces* 88, 865-981.

Lorence, J. (1992): “Service Sector Growth and Metropolitan Occupational Sex Segregation,” *Work and Occupations* 19, 128-156.

Mintz, B. and Krymkowski, D. (2011): “The Intersection of Race/Ethnicity and Gender in Occupational Segregation,” *International Journal of Sociology* 40(4), 31-58.

Mouw, T. and Kalleberg, A. (2010): “Occupations and the Structure of Wage Inequality in the United States, 1980s to 2000s,” *American Sociological Review* 75(3), 402-431.

Petersen, T. and Morgan, L. (1995): “Separate and unequal: Occupation-establishment sex segregation and the gender wage gap,” *American Journal of Sociology* 101, 329-365.

Queneau, H. (2009): "Trends in occupational segregation by race and ethnicity in the USA: evidence from detailed data," *Applied Economics Letters* 16, 1347-1350.

Reskin, B. and Bielby, D. (2005): "A Sociological Perspective on Gender and Career Outcomes," *Journal of Economic Perspectives* 19(1), 71-86.

Reskin, B., Hargens, L. and Hirsh, E. (2004): "Picturing Segregation: The Structure of Occupational Segregation by Sex, Race, Ethnicity, and Hispanicity," mimeo, University of Washington.

Ruggles, S., Alexander, T., Genadek, K., Goeken, R., Schroeder, M. and Sobek. M. (2010): "Integrated Public Use Microdata Series: Version 5.0" [Machine-readable database]. Minneapolis, MN: Minnesota Population Center.

Sastre, M. & Trannoy, A. (2002): "Shapley Inequality Decomposition by Factor Components: Some Methodological Issues," *Journal of Economics* 9, 51-89.

Shorrocks, A. (1998): "Deprivation Profiles and Deprivation Indices," in Jenkins, S.P., Kapteyn, A., and van Praag, B.M.S. (eds.), *The Distribution of House Welfare and Household Production*, Cambridge University Press, New York, 250-267.

Shorrocks, A. (2013): "Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value," *Journal of Economic Inequality* 11, 99-126.

Spriggs, William E. and Rhonda M. Williams (1996): "A Logit Decomposition Analysis of Occupational Segregation: Results for the 1970s and 1980s," *Review of Economics and Statistics* 78, 348-355.

Tomaskovic-Devey, D. and Stainback, K. (2007): "Discrimination and Desegregation: Equal Opportunity Progress in U.S. Private Sector Workplaces since the Civil Rights Act," *Annals of the American Academy* 609, 49-84.

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.

U.S. Bureau of Labor Statistics (2013): "Highlights of Women's Earnings in 2012," BLS Reports 1045 (October).

Appendix

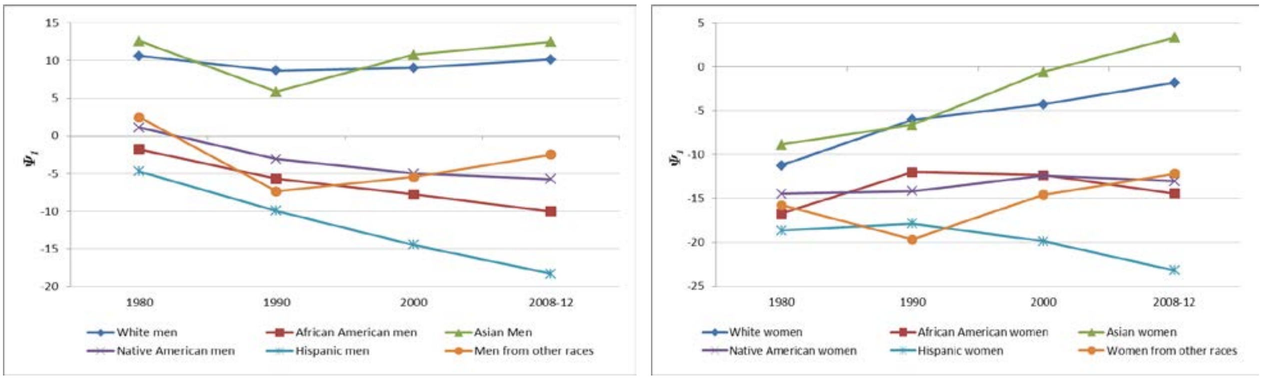

Figure A1. Well-being losses (gains) of the main gender–race/ethnicity groups, Northeast

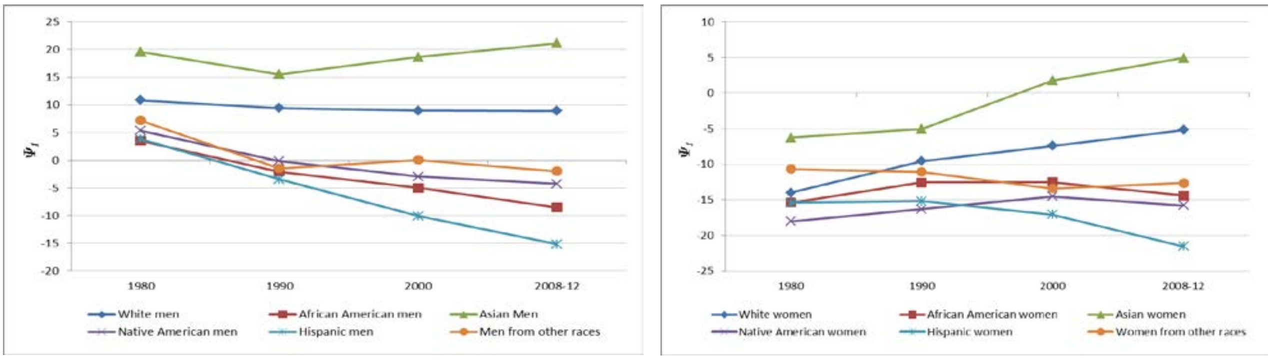

Figure A2. Well-being losses (gains) of the main gender–race/ethnicity groups, Midwest

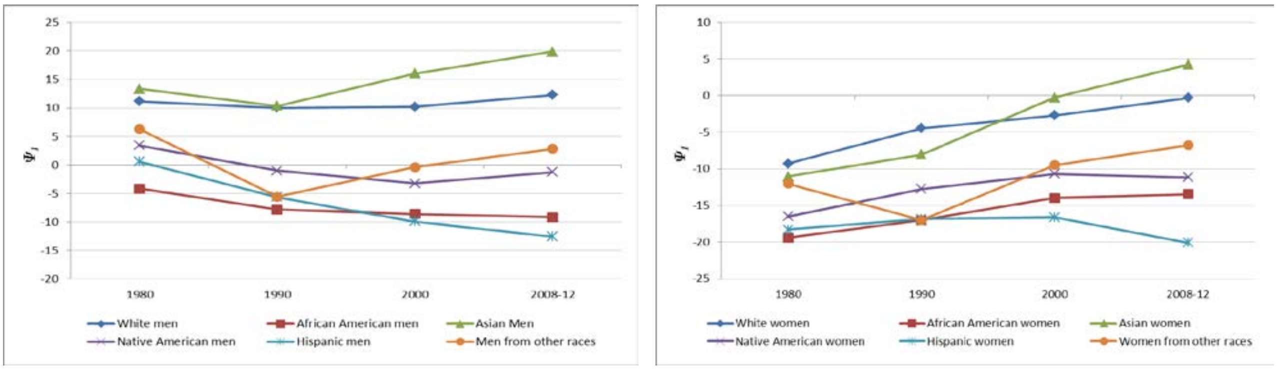

Figure A3. Well-being losses (gains) of the main gender–race/ethnicity groups, South

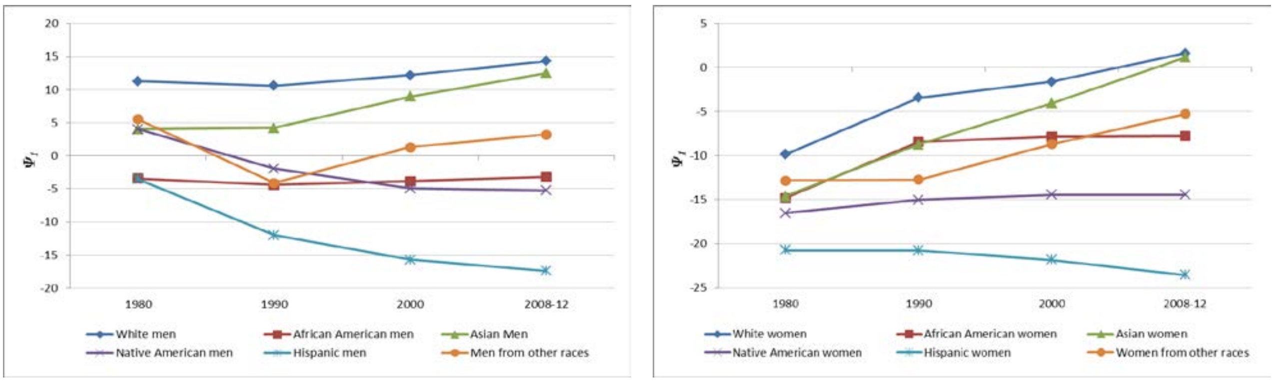

Figure A4. Well-being losses (gains) of the main gender–race/ethnicity groups, West

Table A1. Demographic weights of gender–race/ethnicity groups

	1980					1990					2000					2008-12				
Gender-race/ethnicity groups	**U.S.**	**North east**	**Mid west**	**South**	**West**	**U.S.**	**North east**	**Mid west**	**South**	**West**	**U.S.**	**North east**	**Mid west**	**South**	**West**	**U.S.**	**North east**	**Mid west**	**South**	**West**
White men	48.3	49.8	52.5	46.0	45.0	43.5	44.7	48.5	42.0	39.2	39.8	41.2	45.3	38.2	34.9	35.5	37.3	42.5	33.4	30.3
African American men	4.9	3.9	3.5	8.1	2.5	4.7	4.1	3.3	7.6	2.3	4.6	3.9	3.4	7.4	2.0	4.8	4.2	3.4	7.7	2.0
Asian Men	0.9	0.7	0.4	0.3	2.7	1.5	1.4	0.7	0.7	3.9	2.0	2.2	1.0	1.1	4.3	2.7	3.1	1.4	1.7	5.3
Native American men	0.3	0.1	0.2	0.3	0.6	0.3	0.1	0.2	0.3	0.6	0.3	0.1	0.2	0.0	0.6	0.3	0.1	0.2	0.2	0.5
Hispanic men	3.4	2.4	1.1	3.3	7.6	4.6	3.4	1.5	4.2	10	6.0	4.2	2.4	6.0	11.5	8.5	6.2	3.6	8.9	14.6
Men from other races	0.1	0.1	0.0	0.1	0.1	0.0	0.1	0.0	0.0	0.1	0.9	0.9	0.6	0.7	1.4	0.8	0.7	0.6	0.7	1.2
White women	34.3	36.6	37.6	31.7	31.7	35.4	37.7	40.3	33.3	31.1	34.0	36.5	39.9	31.7	29.2	31.7	34.6	39.2	29.1	26.1
African American women	4.7	4.0	3.5	7.6	2.1	5.1	4.7	3.8	8.1	2.1	5.3	4.8	4.1	8.6	2.0	5.8	5.2	4.4	9.4	2.0
Asian women	0.8	0.5	0.3	0.3	2.4	1.3	1.1	0.5	0.6	3.5	1.8	1.8	0.8	1.0	4.0	2.5	2.7	1.2	1.5	5.2
Native American women	0.2	0.1	0.1	0.2	0.5	0.3	0.1	0.2	0.3	0.6	0.3	0.1	0.2	0.3	0.6	0.3	0.1	0.2	0.2	0.5
Hispanic women	2.2	1.7	0.7	2.1	4.8	3.1	2.5	1.0	2.9	6.5	4.2	3.5	1.6	4.1	8.1	6.4	5.1	2.5	6.3	11.0
Women from other races	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.0	0.1	0.7	0.8	0.5	0.6	1.2	0.8	0.7	0.6	0.7	1.2

Table A2. Social welfare losses indices (x 100) by regions

	FGT_0			**FGT_1**			**FGT_2**		
Northeast	$\Psi_{\varepsilon=0}$	$\Psi_{\varepsilon=1}$	$\Psi_{\varepsilon=2}$	$\Psi_{\varepsilon=0}$	$\Psi_{\varepsilon=1}$	$\Psi_{\varepsilon=2}$	$\Psi_{\varepsilon=0}$	$\Psi_{\varepsilon=1}$	$\Psi_{\varepsilon=2}$
1980	49.44	49.36	49.36	5.32	5.38	5.84	0.62	0.65	0.79
1990	53.84	53.84	53.84	3.96	3.97	4.41	0.34	0.34	0.43
2000	54.81	56.59	56.59	4.25	3.95	4.33	0.45	0.41	0.50
2008-12	56.91	56.91	56.91	4.52	4.25	4.73	0.66	0.66	0.88
	FGT_0			**FGT_1**			**FGT_2**		
Midwest	$\Psi_{\varepsilon=0}$	$\Psi_{\varepsilon=1}$	$\Psi_{\varepsilon=2}$	$\Psi_{\varepsilon=0}$	$\Psi_{\varepsilon=1}$	$\Psi_{\varepsilon=2}$	$\Psi_{\varepsilon=0}$	$\Psi_{\varepsilon=1}$	$\Psi_{\varepsilon=2}$
1980	42.27	42.27	42.27	5.64	5.96	6.75	0.76	0.84	1.08
1990	50.89	50.89	50.67	4.51	4.67	5.39	0.43	0.46	0.63
2000	52.32	52.32	53.15	4.39	4.26	4.78	0.39	0.38	0.48
2008-12	54.76	54.76	54.76	4.25	4.17	4.86	0.44	0.44	0.59
	FGT_0			**FGT_1**			**FGT_2**		
South	$\Psi_{\varepsilon=0}$	$\Psi_{\varepsilon=1}$	$\Psi_{\varepsilon=2}$	$\Psi_{\varepsilon=0}$	$\Psi_{\varepsilon=1}$	$\Psi_{\varepsilon=2}$	$\Psi_{\varepsilon=0}$	$\Psi_{\varepsilon=1}$	$\Psi_{\varepsilon=2}$
1980	53.34	50.04	50.04	5.14	5.21	5.70	0.61	0.65	0.83
1990	57.32	57.32	56.99	4.36	4.28	4.77	0.44	0.45	0.59
2000	59.70	60.66	60.66	4.49	4.07	4.37	0.48	0.43	0.52
2008-12	62.69	62.69	62.43	4.91	4.53	5.01	0.66	0.64	0.85
	FGT_0			**FGT_1**			**FGT_2**		
West	$\Psi_{\varepsilon=0}$	$\Psi_{\varepsilon=1}$	$\Psi_{\varepsilon=2}$	$\Psi_{\varepsilon=0}$	$\Psi_{\varepsilon=1}$	$\Psi_{\varepsilon=2}$	$\Psi_{\varepsilon=0}$	$\Psi_{\varepsilon=1}$	$\Psi_{\varepsilon=2}$
1980	51.54	51.54	51.54	5.20	5.21	5.62	0.61	0.64	0.77
1990	56.88	56.88	56.88	4.43	4.32	4.71	0.51	0.52	0.66
2000	59.31	59.31	59.31	5.04	4.67	5.00	0.75	0.72	0.91
2008-12	31.84	31.84	37.01	5.62	5.52	6.29	1.09	1.08	1.45

Table A3. Logit regressions for the probability of working in the South (pool samples of the South and other region): estimated coefficients (standard errors below).

	Northeast	Midwest	West
Gender:			
Male	--	--	--
Female	0.017 (0.003)	-0.045 (0.003)	0.043 (0.003)
Education:			
Less than High School	--	--	--
High School	-0.249 (0.006)	-0.084 (0.006)	0.007 (0.005)
Some College	-0.104 (0.006)	-0.134 (0.006)	-0.316 (0.005)
Bachelor's Degree	-0.314 (0.006)	-0.025 (0.006)	-0.266 (0.005)
Race/ethnicity:			
White	--	--	--
Black	0.788 (0.005)	1.041 (0.005)	1.339 (0.006)
Asian	0.257 (0.008)	0.256 (0.010)	-1.229 (0.007)
Hispanic (any race)	0.900 (0.007)	1.044 (0.007)	-0.641 (0.005)
Other	0.427 (0.011)	0.367 (0.011)	-0.723 (0.008)
Years of residence:			
Born in the US	--	--	--
Immigrant <=10 years	-0.426 (0.008)	0.204 (0.010)	0.349 (0.008)
Immigrant > 10 years	-0.557 (0.006)	0.233 (0.007)	-0.101 (0.005)
English:			
Only English	--	--	--
Very well	-0.307 (0.006)	0.083 (0.007)	-0.037 (0.005)
Well	-0.371 (0.009)	-0.117 (0.011)	-0.111 (0.008)
Not well or not at all	-0.374 (0.010)	-0.083 (0.012)	-0.224 (0.009)
Industry:			
Agriculture, forestry, fisheries, and mining	--	--	--
Construction	-0.438 (0.012)	0.127 (0.009)	0.250 (0.009)

Manufacturing-1	-0.626 (0.012)	-0.650 (0.009)	0.076 (0.009)
Manufacturing-2	-0.675 (0.012)	-0.392 (0.009)	0.303 (0.009)
Transportation, communications, other public utilities and wholesale trade	-0.645 (0.011)	-0.050 (0.009)	0.104 (0.008)
Retail trade	-0.628 (0.011)	-0.078 (0.008)	0.139 (0.008)
Finance, insurance, and real estate	-0.844 (0.011)	-0.092 (0.009)	0.125 (0.009)
Business and repair services	-0.589 (0.012)	0.002 (0.010)	0.058 (0.009)
Personal services, and entertainment and recreation services	-0.665 (0.012)	0.002 (0.010)	-0.159 (0.009)
Professional and related services	-0.808 (0.011)	-0.107 (0.008)	0.115 (0.008)
Public administration and active duty military	-0.333 (0.012)	0.399 (0.010)	0.147 (0.009)
Intercept	1.446 (0.011)	0.404 (0.009)	0.630 (0.008)
Number of observations	3,796,796	4,071,446	4,070,021
Pseudo-R2	0.030	0.043	0.063
Wald chi2(23)	81,059.1	111,381.2	168,385.8