EQUALITAS
Economics of Inequality and Poverty Analysis

EQUALITAS Working Paper No. 60

The Welfare Effects of Occupational Segregation by Gender and Race: Differences Across U.S. Regions

Coral del Río

Olga Alonso-Villar

February, 2019

The Welfare Effects of Occupational Segregation by Gender and Race: Differences Across U.S. Regions*

Coral del Río and Olga Alonso-Villar#

Universidade de Vigo (ECOSOT-ECOBAS and EQUALITAS)

Abstract

Using tools rooted in welfare economics, this paper explores the social welfare loss that arises from occupational segregation by gender and race in the U.S. at the subnational level. Our findings indicate that the phenomenon is not homogenous across the country (and also that spatial variation has increased over time, 1980–2012). After controlling for characteristics, some regional disparities in welfare losses persist. The (conditional) losses are lower in the Northeast than in the South and West according to a wide range of indicators, including those that take into account the relative size of disadvantaged groups (incidence), the magnitude of their losses (intensity), and the inequality among those groups. The intensity of the phenomenon is also lower in the Northeast than in the Midwest. On the contrary, the West has the highest (conditional) losses, although the intensity of the phenomenon barely differs from that in the South or Midwest.

JEL Classification: D63; J15; J71; R23

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

---

* We gratefully acknowledge financial support from 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) and *Xunta de Galicia*.

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

1. Introduction

Women and men occupy different positions in worldwide labor markets, and the United States is not an exception. Women tend to be concentrated in jobs characterized by lower wages, less authority, and fewer opportunities for promotion (Reskin and Bielby, 2005). In fact, differences in the occupational sorting of women and men play an important role in explaining the gender pay gap (Petersen and Morgan, 1995; Cotter et al., 2003; Del Río and Alonso-Villar, 2015). Moreover, occupational (and industrial) segregation has become more important than human capital in explaining the wage gap (Blau and Kahn, 2017).

Extensive literature exists on occupational segregation by gender in the United States. Many of these works document a reduction in segregation in the second half of the 20th century and stagnation at the beginning of the 21st (Beller, 1985; Bianchi and Rytina, 1986; 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 as a result of the political pressure for gender equality that became a force in the 1970s yet 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 (one out of two) but still considerable (Hegewisch et al., 2011).

The literature on occupational segregation has more recently turned its attention to segregation by race and ethnicity. Research has shown that segregation between blacks and non-blacks also decreased in the second half of the past century whereas segregation between Hispanics and non-Hispanics increased (Queneau, 2009).[1] 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).

Gender and race/ethnicity are important traits that help explain why individuals work in the occupations they do. On one hand, some jobs are socially considered more suitable for women and others for men (Reskin and Bielby, 2005). On the other hand, the white population has access to better occupations than African Americans (Branch, 2007; Del Río and Alonso-Villar, 2015). However, segregation by race/ethnicity does not affect women and

[1] We use the terms "black" and "African American" interchangeably.

men equally. To the contrary, there are fewer differences in segregation levels among female groups 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). In other words, the position of an individual in the labor market is not independent of that of the group to which she/he belongs; gender and race contribute to generate social hierarchies, as feminist theorists have largely discussed (Collins, 2000; Glenn, 1999; Browne and Misra, 2003). However, intersectionality (i.e., the fact that the intersection of two or more social identities creates new categories with their own identities; Darity et al., 2015) has received little inquiry in the literature on segregation, which has focused mainly on segregation by either gender or race.

The aim of this paper is to explore, in an intersectional framework, the economic consequences of occupational segregation by gender and race/ethnicity in the U.S. at the subnational level. These consequences are measured by the (objective) welfare loss that a society experiences due to the high concentration of some gender–race/ethnicity groups in low-paid occupations, which means these groups earn below average wages. We distinguish among the four census regions: Northeast, Midwest, South, and West, which have a long tradition in comparative statistical analyses because they group states based on historical, demographic, and economic characteristics.

Using the tools proposed by Alonso-Villar and Del Río (2017) and Del Río and Alonso-Villar (2018), which are rooted in the literature on welfare economics and deprivation/poverty, we assess whether there are meaningful regional differences in segregation-related losses, and show how these 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 five-year sample of the American Community Survey (ACS). We also explore the causes of observed interregional disparities in social welfare losses. For that purpose, we use the propensity score procedure proposed by DiNardo et al. (1996) as adapted by Gradín et al. (2015) to build a counterfactual economy in which no regional differences exist in terms of gender–race composition, education levels, immigration profile, or industrial structure. Following Gradín (2013), the contribution of each explanatory factor is obtained using the Shapley decomposition, which is commonly used in income distribution literature and has the advantage of being independent of the sequence in which the factors are introduced in the analysis.

There are five 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 distinguishing among nearly 400 occupational categories. Second, we deal with occupational segregation at a regional (not national) level. In different geographical areas, 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. Third, we measure the welfare loss or gain experienced by each of these groups that results from its occupational sorting, thus transcending the mere measurement of unevenness, on which most segregation analyses focus, to address the economic consequences of that unevenness in terms of (objective) welfare, which is where the main problem lies. All of the above will allow us to answer the following questions: Are some gender–race/ethnicity groups systematically more concentrated in low-paid jobs using a fine occupational classification? How have the occupational achievements of the various groups evolved over time in each region?

Fourth, for each U.S. census region, we quantify the social welfare loss that arises from segregation by aggregating the groups' welfare losses in a manner consistent with extant literature on deprivation and poverty. Is segregation by gender and race a more severe problem in the West than it is in Northeast? In particular, is the incidence of the problem higher in the West? In other words, is the relative size of groups who tend to concentrate in low-paid jobs higher in the West than in the Northeast? Regarding the intensity of the phenomenon (i.e., the magnitude of the welfare losses of the groups due to their overrepresentation in low-paid jobs), is it higher in the West than in the South? Additionally, are the groups overrepresented in low-paid jobs more unequal in terms of occupational achievements in the Midwest than they are in the Northeast?

Fifth, we account for differences in characteristics that may explain those regional disparities. Would segregation by gender and race be a more important issue in some regions than in others if there were no regional differences in education, gender and racial composition, immigration profile, and industrial structure? Do gender and race/ethnicity make it more difficult for some groups to integrate in some regional labor markets than in others?

2. Data and Methodology

2.1 Data

We use the U.S. decennial censuses (covering 1980, 1990, and 2000) and the 2008–12 five-year sample of the American Community Survey—which replaced the census long form after 2000 and offers data on occupation. The data were provided by the Integrated Public Use Microdata Series (IPUMS-USA) at the Minnesota Population Center (Ruggles et al., 2010).[2] 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-USA, which is based on the 1990 Census Bureau classification and accounts for 389 job titles. We 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 combining gender 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

To quantify a region's social welfare loss, we follow two steps. First, we measure the (objective) welfare loss or gain that each gender–race/ethnicity group in that region

[2] This dataset is publicly available at https://usa.ipums.org/usa/.

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

experiences as a result of its uneven distribution across occupations. This task is accomplished using the index proposed by Alonso-Villar and Del Río (2017), which is based on notions from the literature on welfare economics. So long as the underrepresentation (resp. overrepresentation) of the group occurs in occupations with wages above the average, the group will have a welfare loss (resp. gain) with respect to the case of no segregation. Second, we aggregate the welfare losses of the groups (in each region) via the approach developed in Del Río and Alonso-Villar (2018). This method is similar to the one followed in the literature on deprivation and poverty, since a group's welfare loss can be viewed as a shortfall with respect to the case of no segregation.

Measuring the Welfare Loss or Gain of a Group Arising from its Occupational Sorting

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 welfare loss or gain that is associated with its occupational sorting, expressed in per capita terms, we use the index Ψ_1^g proposed by Alonso-Villar and Del Río (2017):

$$\Psi_1^g = \sum_j \left(\frac{c_j^g}{C^g} - \frac{t_j}{T} \right) \ln \frac{w_j}{\overline{w}}, \tag{1}$$

where 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, and $\overline{w} = \sum_j \frac{t_j w_j}{T}$ is the average wage of the economy. The ratio $\frac{w_j}{\overline{w}}$ reflects the relative wage of occupation j as compared to the economy's average wage. All these variables refer to the region under study.

The index Ψ_1^g results from using a well-known welfare function to compare the welfare associated with the occupational sorting of the group and the welfare the group would have if it were not segregated (i.e., if $\frac{c_j^g}{C^g} = \frac{t_j}{T}$ in every occupation j). In the former case, the welfare function is applied over an artificial "income" distribution where each member of the group is given an "income" equal to the average wage of the occupation in which she/he works divided by the average wage of the economy. In the latter case, the group is assumed to have a representation in each occupation equal to its weight in the economy and the "income" distribution is defined accordingly. As Alonso-Villar and Del Río (2017) explain, the welfare level of a group associated with its occupational sorting depends not only on the group's earnings but also on the within-group inequality that arises from the fact that some group's members may work in low-paid occupations and others in highly paid ones.

The index is equal to zero if either (a) the group is not segregated 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 is positive (resp. negative) when the group is overrepresented (resp. underrepresented) in highly paid occupations and underrepresented (overrepresented) in low-paid ones.

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) welfare, which is where the main problem lies.

Measuring Welfare Losses of the Whole Society

The above tool is insufficient for determining the welfare loss of an entire region due to segregation. The reason is that some groups may derive gains—while other groups endure losses—stemming from their occupational sorting. One way of dealing with this issue would be to calculate the average welfare 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. A more suitable way of quantifying a region's social welfare loss resulting from the occupational sorting of its demographic groups is to use, as proposed by Del Río and Alonso-Villar (2018), a framework similar to the one employed in the literature on deprivation and poverty. The losses due to segregation are viewed as "gaps" with respect to the case of no segregation and those losses are aggregated following some criteria widely assumed in the literature on poverty. This is the methodology we follow. Note, however, that this approach departs from the literature on deprivation in that 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 et al., 2015).

To obtain the welfare loss of a region due to occupational segregation by gender and race/ethnicity, first, we calculate the welfare loss or gain of each gender–race/ethnicity group

g using the index Ψ_1^g defined earlier. Then we rank the groups with welfare losses (i.e., those with $\Psi_1^g < 0$) from high to low levels of loss whereas the groups with no losses (i.e., those with $\Psi_1^g \geq 0$) come next in the ranking, in no particular order. If we denote by $C \equiv (C^1, ..., C^n)$ the vector representing the demographic size of the n gender–race/ethnicity groups and by $p^k = \frac{C^1 + \cdots + C^k}{T}$ the demographic share of the first k groups ($k = 1, ..., n$), the *social welfare loss curve associated with segregation* (WLAS) at point p^k is defined as the weighted sum of the welfare losses of the first k groups. Namely:

$$W(p^k) = \sum_{g=1}^{k} \frac{C^g}{T} d^g , \qquad (2)$$

where d^g is equal to the absolute value of Ψ_1^g if the group has a welfare loss and zero otherwise (at intermediate points p, $W(p)$ is determined by linear interpolation). This curve provides useful information about the social welfare loss of a region (see Figure 1).

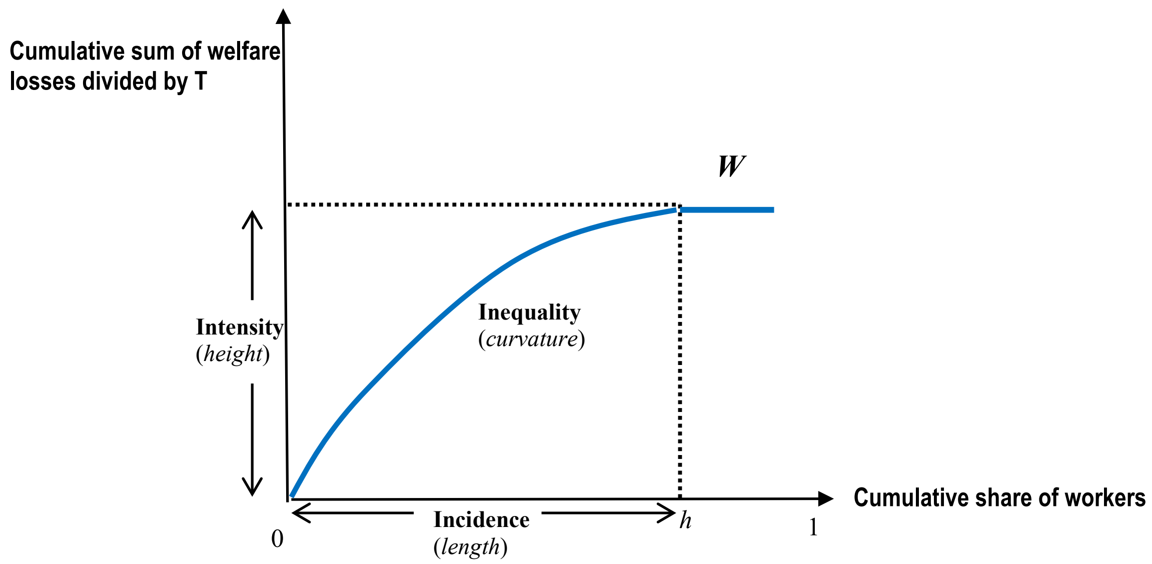

Figure 1. The WLAS curve, W
Source: Del Río and Alonso-Villar (2018)

The abscissa value at which the curve becomes horizontal, denoted by h, represents the incidence of the phenomenon—namely, the population share that the groups with welfare losses account for. The maximum height of the curve conveys the problem's intensity (i.e., the cumulative losses of the groups divided by T). Finally, the curvature of the WLAS curve between the origin and point h illustrates the inequality that exits among disadvantaged groups (i.e., those with welfare losses).[4]

These curves are a powerful tool because, when one curve dominates another (i.e., when the former is never above the latter and 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.

Apart from these curves, to measure the social welfare loss, Del Río and Alonso-Villar (2018) propose a family of indices that result from adapting the well-known FGT poverty indices developed by Foster et al. (1984) to this context:

$$\mathrm{FGT}_{\alpha} = \frac{1}{T}\sum_{s=1}^{s^*}(d_s)^{\alpha}, \tag{3}$$

where $\alpha \geq 0$ is an inequality aversion parameter associated with the welfare loss inequality among groups with welfare losses, d_s is the welfare loss of worker s (set equal to the per-capita welfare loss of the group to which s belongs), and s^* is the number of individuals for whom $d_s > 0$.[5]

Del Río and Alonso-Villar (2018) 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, i.e., h) and index FGT_1 (which measures the welfare losses of the disadvantaged groups divided by T, i.e., the height of W) 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. 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.

[4] 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).

[5] Note that the incidence of the phenomenon, h, and s^* are related given that $h = \frac{s^*}{T}$.

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 will 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 and of the incidence of this phenomenon (i.e., taking into account, respectively, the welfare losses of the groups divided by T and the percentage of workers who belong to groups with welfare losses). 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 welfare losses among deprived groups in the Northeast than in the Midwest.

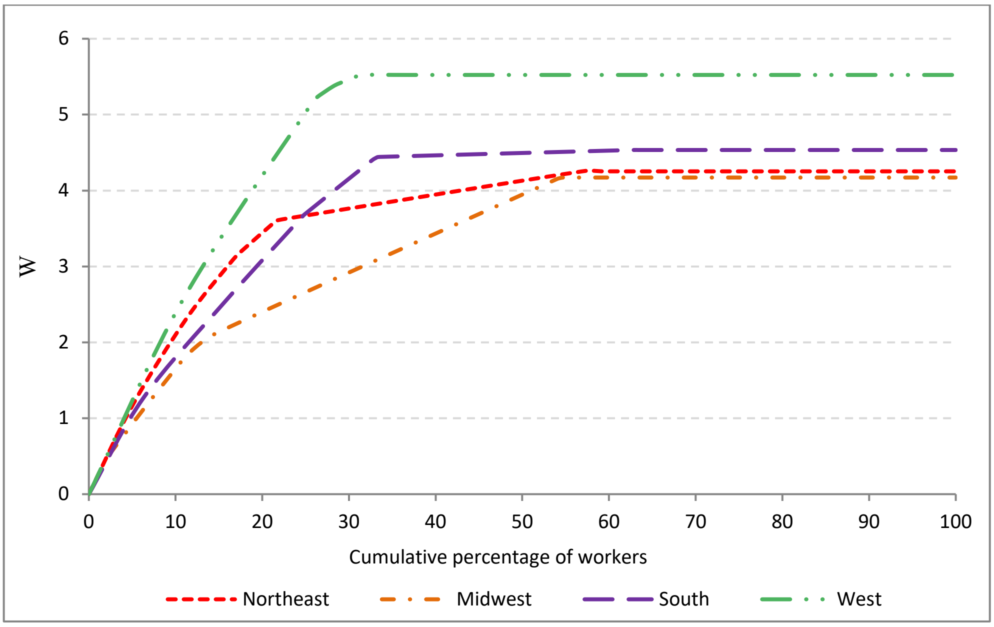

Figure 2. The WLAS curve by region, 2008–12

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 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 gains associated with their occupational sorting (see Figure A4 in the Appendix). The West's being dominated occurs because there 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.

WLAS Curves for Each Region, 1980

Figure 3, which plots the WLAS for all regions in 1980, reveals a considerably different scenario. Here, all curves cross, so we are unable to determine which regions are better-off or worse-off.

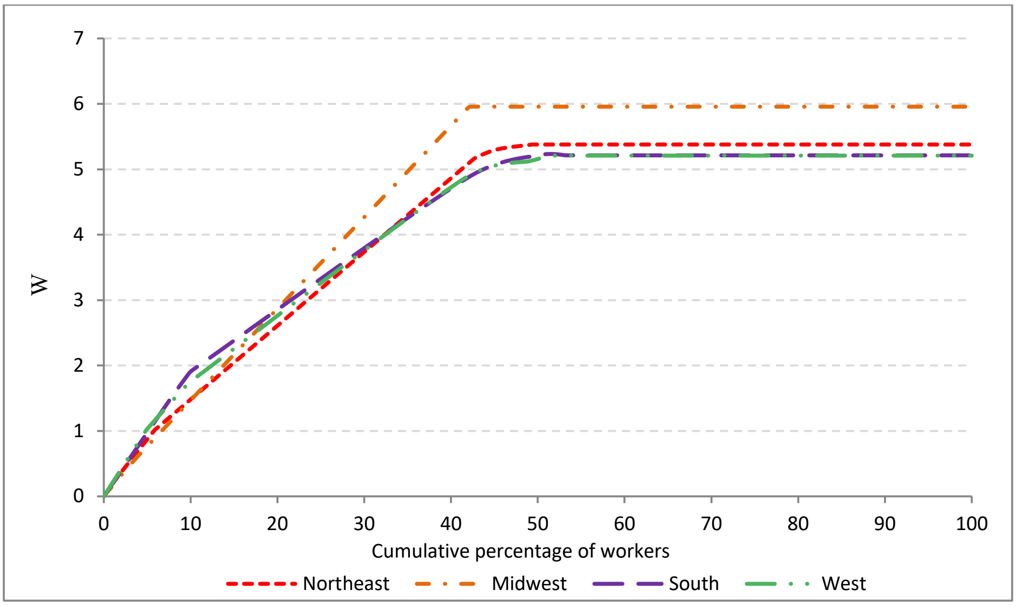

Figure 3. The WLAS curve by region, 1980

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 welfare 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 problem 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.

FGT Indices for Each Region over the Period 1980–2012

Figure 4 illustrates the evolution of the FGT_2 index for each region. As mentioned earlier, the FGT_2 index measures a region's social welfare loss due to the 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 and FGT_1 indices is plotted in Figures 5 and 6, respectively.

The FGT_2 index exhibits a U-shape trend in every region, as also happens at the national level (Del Río and Alonso-Villar, 2018). 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. [6] Third, the differences among the regions were much greater in the last decade than in

[6] 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.

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.

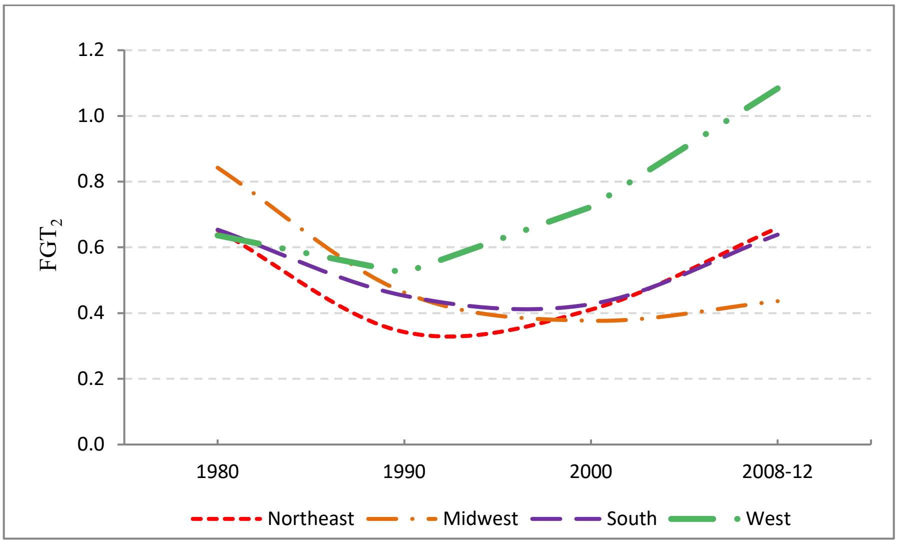

Figure 4. Index FGT_2 (x100) by region

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, at least in part, due to the remarkable reduction in the intensity of the phenomenon (Figure 6).[7] 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 welfare losses for white women did, in fact, occur in the Midwest (see Figures A1–A4 in the Appendix). Although this pattern 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 as well. They were experiencing welfare gains (rather than losses) by the 1990s.

[7] 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).

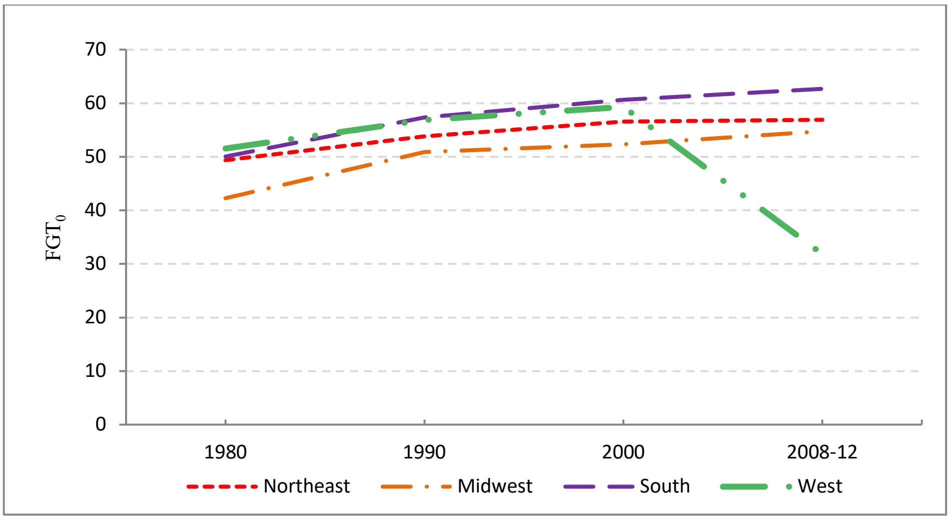

Figure 5. Index FGT_0 (x100) by region

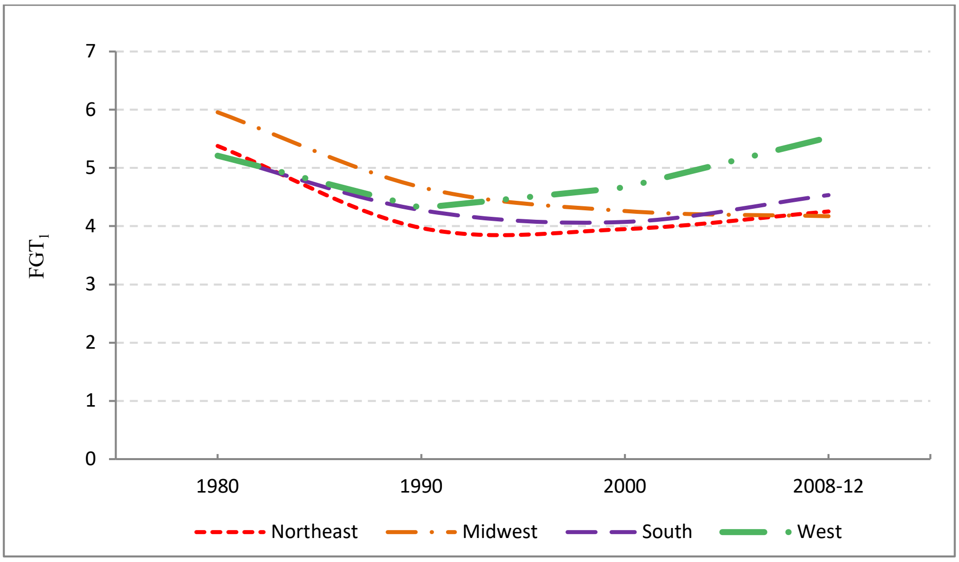

Figure 6. Index FGT_1 (x100) by region

However, the experience of male groups in the Midwest does not conform to reductions in the FGT_2 index (Figure A2). African Americans and Hispanics worsened over the period, as they shifted from having small welfare gains in 1980 to experiencing losses soon thereafter. 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. On the other hand, the other main racial groups (whites and Asians) had always welfare gains in the period of analysis and, therefore, they have no effect on the FGT_2 index.

In the West, the social welfare losses in terms of the FGT_2 index not only became greater after 1990 (unlike the other regions, where the increase started later) but were also persistently much higher than those of the other regions (Figure 4).[8] Note, however, that the problem in this region over the last decade was not a rise in the incidence of the phenomenon (Figure 5); in fact, the percentage of workers belonging to groups with welfare losses actually declined (those proportions were 59% in 2000 and 32% in 2008–12). The reason for this may be that white (and Asian) women began to experience small welfare gains in the 2000s (Figure A4). The disparities among groups with welfare losses in 2008–12 cannot help us either to explain why the FGT_2 index is higher in the West than elsewhere.[9] 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 (Figure 6). This 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.[10] This reduction in losses seems to arise from a stronger decrease in the welfare 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 segregation-related 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 welfare 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.

[8] 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).

[9] 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 welfare losses 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.

[10] 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.

Apart from helping to explain the evolution of welfare losses in each region, Figures A1–A4 also reveal that Asian and white men have the highest occupational attainments (i.e., the highest welfare levels according to the Ψ_1 index) whatever the region or the year. On the contrary, Hispanics (particularly women) men tend to have the lowest occupational attainments. The rise in their demographic shares along the study period has gone hand in hand with worsening job prospects for these two groups in all regions. We also find that all regions witness a certain convergence between African American women and men driven by the improvement of the former and the worsening of the latter. Despite this, the occupational attainments of African American women are lower than those of African American men and those of either Asian or white women. The gap between African American women and white/Asian women has increased over time as a consequence of the largest occupational advancements of the latter groups. In 2008–12, Asian women were the only female group with occupational attainments above the average in all regions ($\Psi_1 > 0$) although these attainments were much lower than those of either white or Asian men.

4. Controlling for Regional Characteristics

The analysis so far has revealed substantial disparities among regions as regards losses in social welfare due to occupational segregation by gender and race/ethnicity. 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–race composition, immigration profile, 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 the other three regions, a counterfactual economy such that the share of each 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 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 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 loss, 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

We "homogenize" the four regions according to six key characteristics that 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);[11] (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).[12] These are the characteristics or attributes to which we refer hereafter.

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 levels, consists of building a counterfactual distribution for each region so that each "cell" or subgroup resulting from combining the main attributes mentioned above (e.g., Asian immigrant men 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.

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

[12] 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".

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 will explain how to build the counterfactual distribution for a single region: the Midwest.

Let $z \equiv (z_1,\dots,z_k)$ 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":

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

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 actual 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).[13]

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, English proficiency, 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.

4.2 Are There Regional Differences in Conditional Welfare Losses?

Figure 7, which displays each region's conditional WLAS curves in 2008–12 (with the South as the reference region), reveals that the picture changes substantially as compared with the unconditional analysis (Figure 2).[14] First, although the West's curve is still dominated by the others, the intensity of that region's welfare loss no longer differs much from that in those other regions.[15] Second, the Midwest's WLAS curve no longer dominates all others, and its maximum height (which embodies the intensity of the phenomenon) is no longer the lowest.[16] Third, although controlling for characteristics reduces by half the interregional disparities in

[13] 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 actual 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 actual distribution.

[14] The coefficients of the logit regressions are shown in Table A3.

[15] When the region of reference is other than the South, we also find that the curves become less distinctive. 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.

[16] 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 those in the South and West are not much lower).

welfare losses,[17] the conditional analysis shows that notable differences among regions persist. 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.[18] In particular, the values of the FGT_2 index are 0.55 for the Northeast, 0.64 for the South, and 0.8 for the West (0.53 for the Midwest).

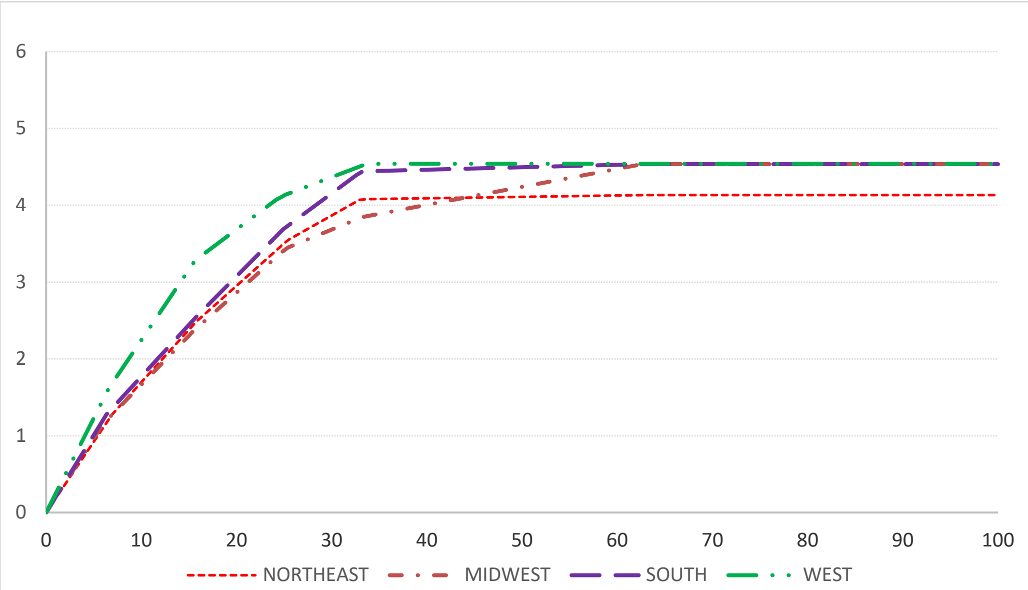

Figure 7. The conditional WLAS curves, 2008–12 (reference region: South)

The question that now arises is: Which of the characteristics we consider in the conditional analysis are the most explanatory of the actual regional disparities? To answer that question, we use the FGT_2 index and decompose the change between each region's conditional and unconditional welfare 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 8 reports the contribution of these factors—determined via Shapley decomposition—for each region.

[17] According to the coefficient of variation, regional disparities in terms of the FGT_2 index decline by 50% when the South is the reference region; with other reference regions, the reduction is equal or greater (50%, 65%, and 70% when, respectively, the West, Northeast, and Midwest serve as the reference region). This finding suggests that our covariates explain at least half of the variability in regional welfare losses.

[18] The results for the Northeast are robust to changing 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 welfare loss. 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 actual (i.e., not the counterfactual) distribution. Yet, if the West were characterized by the same gender and 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 welfare loss would be 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 7 than it is in Figure 2. Figure 8'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 (the welfare loss decreases in the Northeast and increases in the Midwest).

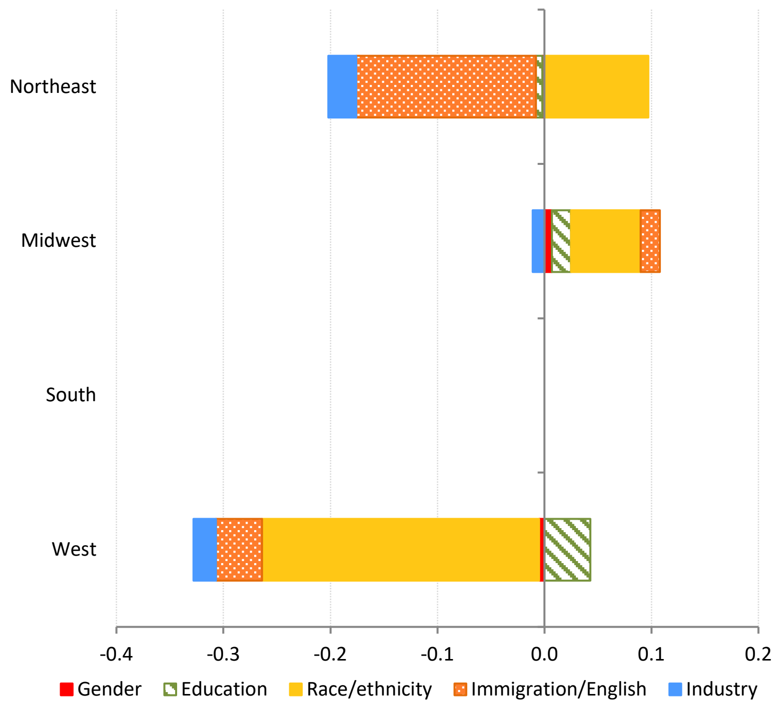

Figure 8. Conditional minus unconditional FGT_2 index (x100) (reference region: South)

We infer from Figure 8 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.[19] The analysis suggests that a large part of the high unconditional welfare losses in the West comes from its racial/ethnic composition. On the contrary, the lower losses in the Midwest arise mainly from its lower racial diversity. Notwithstanding, we have to bear in mind that regional disparities still persist after these characteristics have been taken into account. The Midwest and Northeast have lower welfare losses associated with occupational segregation by gender and race/ethnicity than the South and West.

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 the mutually exclusive groups into which a society can be partitioned. 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 also differ in terms of promotional opportunities, physical risk, required hours per week, and so on. Occupational segregation 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).

Distinguishing among 12 gender–race/ethnicity groups and nearly 400 occupational titles, this paper has estimated the (per capita) welfare gain or loss of each group associated with its occupational sorting, which reflects its occupational achievements. This has been done for the last three decades (from 1980 to 2012) and each of the four census regions (Midwest, Northeast, South, and West). We have found that Asian and white men have higher occupational achievements than any other male or female group, a pattern consistent across time and regions. The lowest position in the ranking tends to be for Hispanics, especially women. African American women systematically rank lower than either white or Asian women or African American men.

This paper has also quantified the social welfare losses of each region, accounting not only for the welfare loss of each group but also its size. Our findings indicate that the phenomenon is

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

not homogenous across the country and that regional disparities have increased over time. In 1980, the Midwest exhibited the greatest losses, surpassing those of the South, Northeast, and West, which shared a common value. Three decades later, the Midwest had the lowest losses, whereas the West's losses exceeded by far those of the other three regions (a pattern that started in 1990).

The reduction in the Midwest's losses appears to be explained by the occupational advancements of white women—who account for an important (and increasing) share of workers and were in a worse situation there in 1980 than in the other regions—and Asian women. In the West, the occupational achievements of Hispanic women and men, who account for a large share of the workers, deviated more dramatically (and increasingly) from their African American counterparts (who are less concentrated in low-paid occupations in this region than in the others). The increasing proportion of Hispanic population in the West seems to explain why social welfare losses increased there over time (despite the occupational advancements of African American women).

After controlling for regional characteristics—racial composition and, to a lesser extent, immigration profile being the most important factors—we found that at least half of the interregional differences in social welfare losses disappear, although some spatial disparities persist. The (conditional) losses associated with occupational segregation by gender and race/ethnicity are lower in the Northeast than in the South and West according to a wide range of indicators, including those that take into account the relative size of disadvantaged groups (incidence), the magnitude of their losses (intensity), and the inequality among those groups. The intensity of the phenomenon is also lower in the Northeast than in the Midwest. On the contrary, the West has the highest (conditional) losses, although the intensity of the phenomenon barely differs from that in the South or Midwest. The analysis suggests that the integration of women and racial/ethnic minorities into the labor market differs across regions beyond spatial disparities in groups' attributes and industrial structures; hence there may well exist other factors associated with the characteristics of the regions—such as citizens' attitudes toward gender and race, government policies, and social capital—that help to explain these differences. The role played by these other factors is beyond the scope of this paper, but our findings offer fruitful avenues for further research on this topic.

References

Alonso-Villar, Olga, and Coral Del Río. 2017. “Occupational Segregation and Well-being.” Review of Income and Wealth 63(2): 269-287.

Alonso-Villar, Olga, Coral Del Río, and Carlos Gradín. 2012. “The Extent of Occupational Segregation in the United States: Differences by Race, Ethnicity, and Gender.” Industrial Relations 51(2): 179-212.

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

Bianchi, Suzanne, and Nancy Rytina. 1986. “The Decline of Occupational Sex Segregation during the 1970s: Census and CPS Comparisons.” Demography 23: 79-86.

Blau, Francine, Peter Brummund, and Albert Liu. 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, Francine, and Lawrence Kahn. 2017. “The Gender Wage Gap: Extent, Trends, and Explanations.” Journal of Economic Literature 55(3): 789-865.

Branch, Enobong H. 2007. “The Creation of Restricted Opportunity due to the Intersection of Race & Sex: Black Women in the Bottom Class.” Race, Gender & Class 14(3-4): 247-264.

Browne, Irene, and Joya Misra. 2003. “The Intersection of Gender and Race in the Labor Market.” Annual Review of Sociology 29: 487-513.

Collins, Patricia H. 2000. *Black Feminist Thought: Knowledge, Consciousness and the Politics of Empowerment*. New York: Routledge (second edition).

Conrad, Cecilia. 2005. “Changes in the Labor Market Status of Black Women, 1960-2000. ” In African Americans in the U.S. Economy, edited by Cecilia Conrad, John Whitehead, Patrick Mason, and James Steward, pp. 157-162. Lanham, MD: Rowman & Littlefield Publishers.

Cotter, David, Joan Hermsen, and Reeve Vanneman. 2003. “The Effects of Occupational Gender Segregation across Race.” The Sociological Quarterly 44: 17-36.

Darity Jr., William, Darrick Hamilton, and James Stewart. 2015. “A Tour de Force in Understanding Intergroup Inequality: An Introduction to Stratification Economics.” Review of Black Political Economy 42: 1-6.

Del Río, Coral, and Olga Alonso-Villar. 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, Coral, and Olga Alonso-Villar. 2018. “Segregation and Social Welfare: A Methodological Proposal with an Application to the U.S.” Social Indicators Research 137(1): 257-280.

DiNardo, John, Nicole Fortin, and Thomas Lemieux. 1996. “Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach.” Econometrica 64: 1001-1044.

Foster, James, Joel Greer, and Erik Thorbecke. 1984. “A Class of Decomposable Poverty Measures.” Econometrica 52(3): 761–766.

Glenn, Evelyn Nakano. 1999. “The Social Construction and Institutionalization of Gender and Race: An Integrative Framework.” In Revisioning Gender, edited by Myra Marx Ferree, Judith Lorber, and Beth B. Hess, pp. 3-43. Thousand Oaks, CA: Sage publications.

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

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

Hegewisch, Ariane, Hannah Liepmann, Jeff Hayes, and Heidi Hartmann. 2010. “Separate and Not Equal? Gender Segregation in the Labor Market and the Gender Wage Gap.” Briefing Paper No. C377. Washington, D.C.: Institute for Women’s Policy Research.

Hegewisch, Ariane, Claudia Willians, and Vanessa Harbin. 2011. “The Gender Wage Gap by Occupation.” Briefing Paper No. C350a. Washington, D.C.: Institute for Women’s Policy Research.

Jenkins, Stephen, and Peter Lambert. 1997. “Three ‘I’s of Poverty Curves, with an Analysis of UK Poverty Trends.” Oxford Economic Papers 49: 317–327.

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

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

Petersen, Trond and Laurie Morgan. 1995. “Separate and unequal: Occupation-establishment sex segregation and the gender wage gap.” American Journal of Sociology 101: 329-365.

Queneau, Hervé. 2009. “Trends in occupational segregation by race and ethnicity in the USA: evidence from detailed data.” Applied Economics Letters 16: 1347-1350.

Reskin, Barbara, and Denise Bielby. 2005. “A Sociological Perspective on Gender and Career Outcomes.” Journal of Economic Perspectives 19(1): 71-86.

Reskin, Barbara, Lowell Hargens, and Elizabeth Hirsh. 2004. “Picturing Segregation: The Structure of Occupational Segregation by Sex, Race, Ethnicity and Hispanicity.” Presented at the American Sociological Association Annual Meeting, San Francisco, CA, August 15, 2004.

Ruggles, Steven, Trent Alexander, Katie Genadek, Ronald Goeken, Matthew Schroeder, and Matthew Sobek. 2010. “Integrated Public Use Microdata Series: Version 5.0.” [Machine-readable database]. Minneapolis, MN: Minnesota Population Center.

Sastre, Mercedes, and Alain Trannoy. 2002. “Shapley Inequality Decomposition by Factor Components: Some Methodological Issues.” Journal of Economics 9: 51-89.

Shorrocks, Anthony. 2013. “Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value.” Journal of Economic Inequality 11: 99-126.

Spriggs, William, and Rhonda 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, Donald, and Kevin Stainback. 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, Donald, Catherine Zimmer, Kevin Stainback, Corre Robinson, Tiffany Taylor, and Tricia McTague. 2006. "Documenting desegregation: Segregation in American Workplaces by Race, Ethnicity, and Sex, 1966-2003." American Sociological Review 71: 565-588.

Appendix

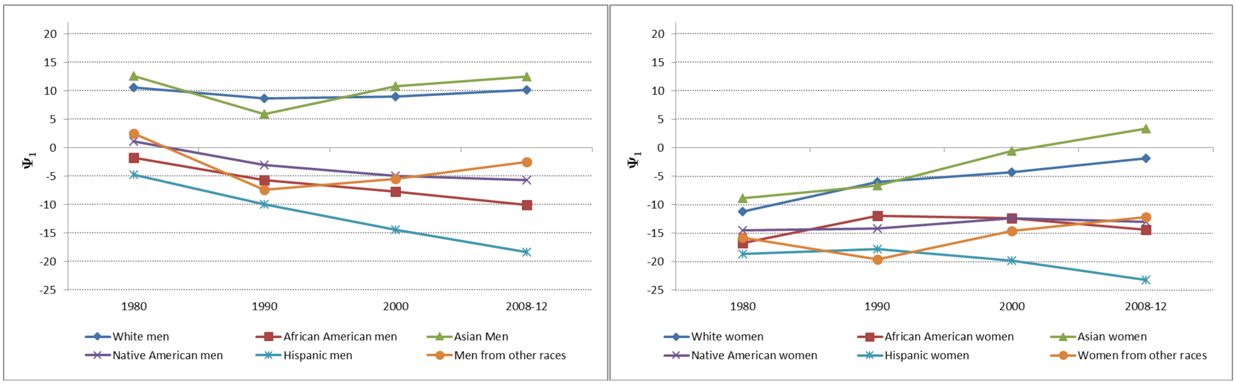

Figure A1. Welfare losses (gains) of the gender–race/ethnicity groups, Northeast

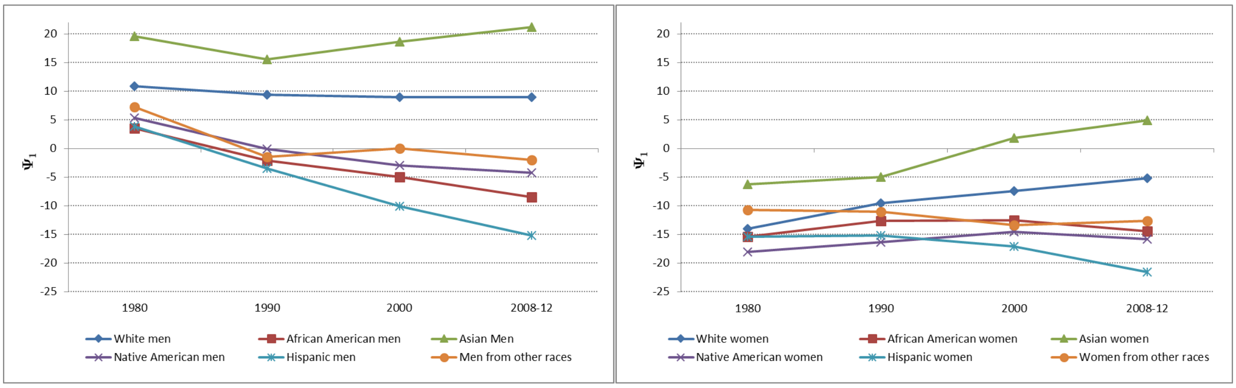

Figure A2. Welfare losses (gains) of the gender–race/ethnicity groups, Midwest

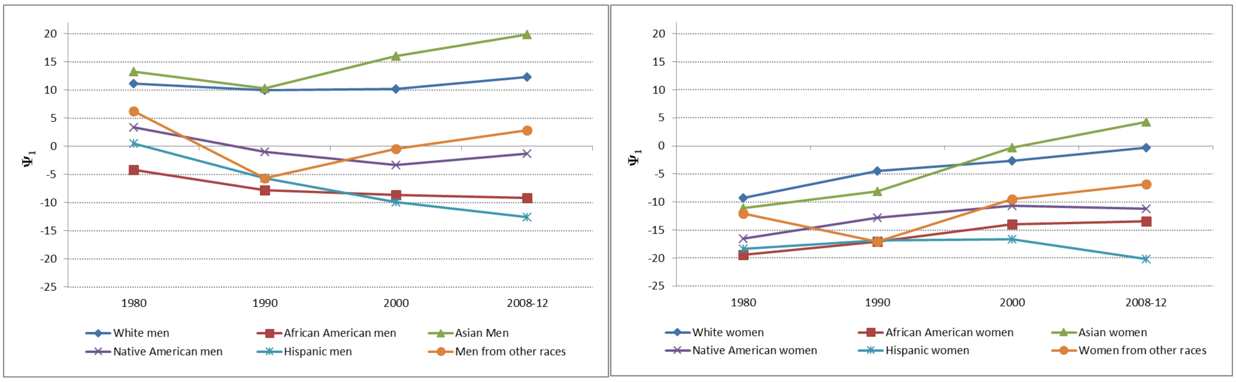

Figure A3. Welfare losses (gains) of the gender–race/ethnicity groups, South

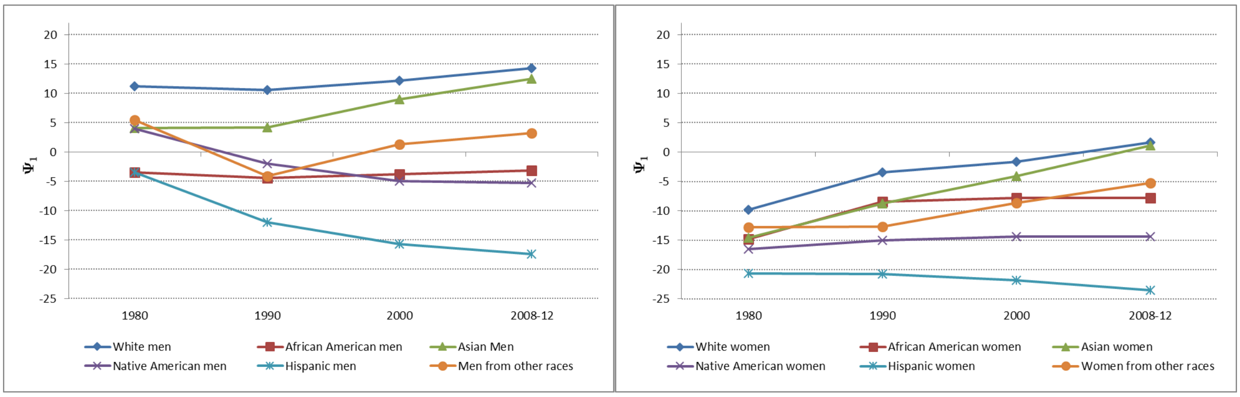

Figure A4. Welfare losses (gains) of the gender–race/ethnicity groups, West

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 region

Northeast	FGT_0	FGT_1	FGT_2
1980	49.36	5.38	0.65
1990	53.84	3.97	0.34
2000	56.59	3.95	0.41
2008-12	56.91	4.25	0.66
Midwest	FGT_0	FGT_1	FGT_2
1980	42.27	5.96	0.84
1990	50.89	4.67	0.46
2000	52.32	4.26	0.38
2008-12	54.76	4.17	0.44
South	FGT_0	FGT_1	FGT_2
1980	50.04	5.21	0.65
1990	57.32	4.28	0.45
2000	60.66	4.07	0.43
2008-12	62.69	4.53	0.64
West	FGT_0	FGT_1	FGT_2
1980	51.54	5.21	0.64
1990	56.88	4.32	0.52
2000	59.31	4.67	0.72
2008-12	31.84	5.52	1.08

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