amse
école d'économie d'aix-marseille
aix-marseille school of economics

Working Papers / Documents de travail

Residential Segregation Matters to Racial Income Gaps Evidence from South Africa

Florent Dubois
Christophe Muller

WP 2022- Nr 05

RESIDENTIAL SEGREGATION MATTERS TO RACIAL INCOME GAPS

Evidence from South Africa*

FLORENT DUBOIS† AND CHRISTOPHE MULLER‡

† University of Reading, Whiteknight streets, Reading, United Kingdom

EconomiX, University Paris Nanterre, 92000 Nanterre, France

‡ Aix-Marseille Univ., CNRS, AMSE, 13001 Marseille, France

March 2022

We contend that residential segregation should be an essential component of the analyses of socio-ethnic income gaps. Focusing on the contemporary White/African gap in South Africa, we complete Mincer wage equations with an Isolation index that reflects the level of segregation in the local area where individuals dwell. We decompose the income gap distribution into detailed composition and structure components. Segregation is found to be the main contributor of the structure effect, ahead of education and experience, and to make a sizable contribution to the

*This work was supported by the French National Research Agency Grant ANR-17-EURE-0020, and by the Excellence Initiative of Aix-Marseille University - A*MIDEX. We would like to thank Jean-Paul Azam, Yann Bramoullé, Olga Cantó-Sánchez, Eliane El Badaoui, Fabien Moizeau, and Michel Lubrano for their invaluable comments. This paper also benefited from insights of conference participants at the 2018 LAGV conference and the 2020 ADRES conference. We would like to express our gratitude to Takwanisa Machemedze, Alex Montgomery, and Lynn Woolfrey at DataFirst for their expert assistance with the data. All remaining errors are ours. Any views expressed in this paper are solely the authors' and do not reflect the position of any institution.

†email address: f.m.g.dubois@reading.ac.uk; postal address: Department of Economics, School of Politics, Economics and International Relations, University of Reading, Whiteknights, Reading RG6 6AA, Berkshire

‡email address: christophe.muller@univ-amu.fr; postal address: Maison de l'économie et de la gestion d'Aix, 424 chemin du viaduc, 13080 Aix-en-Provence

composition effect. Moreover, segregation is found to be harmful at the bottom of the African income distribution, notably in relation to local informal job-search networks, while it is beneficial at the top of the White income distribution. Specific subpopulations are identified that suffer and benefit most from segregation, including for the former, little educated workers in agriculture and mining, often female, confined in their personal networks. Finally, minimum wage policies are found likely to attenuate most segregation's noxious mechanisms, while a variety of policy lessons are drawn from the decomposition analysis by distinguishing not only compositional from structural effects, but also distinct group-specific social positions.

JEL Codes: J15, D31, R23

Keywords: Residential segregation, Post-Apartheid South Africa, Distribution analysis, Generalized decompositions

I. Introduction

Persistent racial differences in socio-economic outcomes is a major concern as it threatens the stability of diverse societies. Earlier attempts to legally justify such differences were met with fierce resistance. In South Africa, the Apartheid regime clashed with the African resistance during most of the twentieth century, while in the United States, the Civil Rights movement battled relentlessly against the Jim Crow legislation. However, explaining the persistence of such differences, most notably income gaps (Bayer and Charles 2018; Leibbrandt et al. 2010b), even after these discriminatory legislations were repealed still constitutes a challenge for social scientists.

To explain these gaps, most of the literature invokes labour market mechanisms, operating either through human capital endowment gaps or discrimination (Altonji and Blank 1999). In addition, Chetty et al. (2020) shed lights on the role played by differences in intergenerational mobility. Yet, among non-market factors, residential segregation has received much less attention than education or family backgrounds. Often, the analysis is limited to the spatial mismatch hypothesis: the tendency for minority members living far from job centers to select poorer job alternatives (Kain 1968). However, because working decisions and selected job opportunities may depend also on people's interactions within their neighborhood, alternative mechanisms based on racial preferences or neighborhood effects may participate in generating income differences.

Nevertheless, there is only limited evidence on the effect of segregation on incomes, almost exclusively based on the United States, and the relative importance of segregation compared to other determinants is rarely assessed. Cutler and Glaeser (1997) find on average a negative effect, while

Edin et al. (2003) estimate a positive effect and Oreopoulos (2003), consistently with the results on the Moving To Opportunity experiment (Katz et al. 2001), do not spot any effect at all. Still in the US, Levy (2022) finds that neighborhoods have sizable effects on wealth that vary significantly with race. These disparate findings may be partly explained by heterogeneity in the mechanisms of segregation. In this regard, Cutler et al. (2008) show that segregation impacts differently immigrants' wages depending on their education group, and Chetty et al. (2016) uncover a positive effect on children's future wages had they moved into lower-poverty neighborhoods before they had reached 13, but a small negative effect for older ones.

In this paper, we provide new evidence on the importance of segregation for racial income differences by using distribution decomposition methods. While not providing causal evidence of mechanisms driving income gaps, this methods provide valuable stylized facts that reveal promising research directions. We use South Africa as a benchmark case for examining these issues. Arguably, this is a most relevant case as the country combines the longest and most pronounced experience of legally enforced segregation with one of the largest racial income gap in the world.

Our analysis is organized into two parts. First, we document the effect of segregation on incomes in South Africa using a simple Mincerian framework. This analysis of the mean effect of segregation has a twofold virtue. Firstly, it establishes a premier estimate for South Africa, as the South African wage gap literature is actually silent on this phenomenon (Sherer 2000; Gradín 2012, 2014; Leibbrandt et al. 2010b). Secondly, it provides a useful benchmark of the aggregated pattern. Then, we scrutinize the heterogeneous association of incomes with residential segregation with RIF-regressions, generalized decompositions (Firpo et al. 2009; Fortin et al. 2011) and sorted effects (Chernozhukov et al. 2018), which allows us to identify for whom segregation matters most.

Consistently with Cutler and Glaeser (1997), we find that segregation has, on average, a negative effect on incomes for Africans. The average effect for Whites is positive but unstable over time. Segregation is the main contributor to the racial income gap, even ahead of education, particularly for the structure effect. The differential effect of segregation emerges even clearer in the distributional analysis. Segregation is associated negatively with income at the bottom of the African distribution, while positively for the top of the White distribution. Segregation appears as the main contributor in the structure effect with the strongest positive contribution below the median.

We complete our analysis of the heterogeneity of segregation effects by detailing the socio-demo-economic characteristics of the African main winners and losers from segregation with a classification analysis in the spirit of Chernozhukov et al. (2018). This points at gender differences, network effects, and unionship as important correlates of these effects of segregation. Finally, we explore our estimations through the prism of the 2018 minimum wage reform in South Africa which is found likely to attenuate the segregation influence on wages for Africans.

The remainder of the paper proceeds as follows: Section II. discusses the potential economic channels through which segregation might affect income. Section III. describes the measure of segregation used and discusses the associated inference problem in mean and distribution decompositions. Section IV. reviews segregation during Apartheid, and the post-Apartheid trends in income inequality, and presents the data. Section V. expounds on the results obtained by decomposing the mean income gap. Section VI. extends the analysis to the entire income distribution, and to the 2018 minimum wage reform. The last section concludes the paper.

II. How Segregation Relates to Incomes

II.A. Individual Preferences for Segregation

In the housing market, segregation, through racial preferences, may transform neighborhoods into clubs and restrict the access to their amenities. In that case, individuals, when deciding where to live, can take into account the racial composition of their targeted neighborhood. Schelling (1971) demonstrates that only a small preference for one's own ethnic group is enough to yield highly segregated local contexts. Realtors also play a role. They can employ discriminatory tactics, such as redlining, because they are themselves racist or because they care for the racial preferences of their current or potential customers (Yinger 1986). Once established somewhere, individuals will vote for their contribution to local public goods. Alesina et al. (1999) show that most individuals in more diverse communities vote for less spending in education when it also benefits to the minority group. This may generate gaps in human capital accumulation across neighborhoods through differences in education quality, which finally could materialize into wage gaps. Consequently, local segregation and income levels could be correlated in that case. This mechanism is amplified under an initial correlation between income and segregation. This would be the case if racial groups are hierarchized by income. Besides, economic gaps between races may be at the origin of racial prejudice (Blumer 1958).

II.B. Neighborhood and Peer Effects

Segregation may act on income through behavior diffusion. Namely, individuals living in segregated areas may be more prone to develop specific work habits when they belong to some local group, and hence to be subjected to group-specific income processes. For example, in the US, Black workers living in ghettos are sometimes believed to be characterized by tardiness, absenteeism, or unreliability, and this may be one reason for their lower incomes. Wilson (1987) claimed that it was inner-city isolation that generated bad work habits. In particular, there is some evidence of a 'ghetto culture' of bad habits that tends to reinforce these habits through social pressure. Even children often feel peer pressure to perform poorly at school. In these conditions, it may be difficult to escape unemployment and poorly paying jobs from within the ghetto. Bénabou (1993) shows that neighborhood and peer effects can explain some individuals' low quality of work. If peers are defined in connection to ethnicity, then the isolation index that we use measures the extent of such social pressure. Besides, social pressure may foster bad practices in one group and good work habits in another, which may further pull apart the incomes of the two groups under segregation.

In addition, ethnic networks may provide differential access to jobs and work promotions (Magruder 2010). In particular, local segregation against one group may limit its access to professional information obtained by other groups (Ioannides and Loury 2004). Ethnically isolated individuals may have lower incomes, ceteris paribus, because their information is more restricted.

II.C. Segmented Labor Markets, Capital Ownership, Trade Unions, and Spatial Mismatch

Segregation may generate earnings differentials by contributing to the segmentation of the labor market (Dickens and Lang 1985; Magnac 1991). Entrepreneurs may pay lower wages to the discriminated group because they are themselves racist and perceive a cost of employing a minority worker (Becker 1957). In addition, if racist entrepreneurs settle disproportionately in the same segregated areas, due to the proximity of an industrial park, for instance, then a correlation between local segregation and the wage gap across groups emerges.

Alternatively, racial discrimination from the employees, potentially sustained by trade unions, may serve as a device for protecting some insider workers' privileges and higher wages in the primary sector. Historically, this was the case in the mining industry in South Africa (Thompson 2001, chap. 4). White miners were collectively organized and had laws passed that gave them a monopoly on well-paid jobs in mines, whereas African miners could have done the same work for a small fraction of their wage. Segregation eases the formation of such collective action by facilitating coordination within politically proactive groups, and the exclusion of discriminated groups. In these conditions, once again, segregation and incomes correlates. Note that in post-Apartheid South Africa, trade unions instead strive at reducing the wage gap, yielding a negative correlation with segregation.

Even when racial discrimination is statistical, rather than taste-based, segregation may still contribute to segmenting the labor market by limiting the information of minority workers for reaching recruiters, which again generates wage gaps correlated with segregation. Large differences in capital (and human capital) ownership across groups, as is the case in South Africa, would strengthen this mechanism.

Segmentation can be implemented through entry barriers and entry costs into the primary sector. Thus, segregation may affect income levels by forcing minority workers to live far away from job opportunities (Kain 1968), thereby raising their search costs and commuting costs. In South Africa, post-Apartheid housing programs have been reinforcing the estrangement of many African workers from job opportunities for at least a decade (Bebbington et al. 2010). Lastly, segregation may make traditional solidarity more salient by isolating some groups from the rest of the economy. For instance, in African communities, traditional redistribution within extended families, neighborhoods and kin groups, may deter workers to search for well-paid jobs (Mhlongo 2019).

III. Methodology

III.A. Measuring Segregation

Segregation is often measured as the propensity of individuals to live with peers, separately from other groups. The most standard approaches assume a given partition of a city[1] as given and use information on the subdivision of the city's population to compute an index. Massey and Denton (1988) propose considering five dimensions of segregation: evenness, exposure, concentration, centralization, and clustering. We focus on evenness and exposure for several reasons. First, they are, by far, the most popular approaches. Second, the other dimensions appear less specific to the notion of segregation, less politically salient, and may require fine-gridded data, which are typically not available.

Evenness refers to the degree of overlap between the spatial distributions of two considered groups. The most common index in the empirical literature on segregation is the Dissimilarity Index, which quantifies the proportion of the minority group that would have to relocate to achieve an equal spatial distribution. Its formula in the case of two groups, say Africans and Whites, for a partition of

[1] Segregation measures can also consider the country level. We will use the 'city' vocabulary in the remainder of this paper, as residential segregation is typically discussed at the city level.

the city into a set I of locations, is:

$$Dissimilarity = \frac{1}{2}\sum_{i\in I}|\frac{White_i}{White_{Population}} - \frac{African_i}{African_{Population}}| \tag{1}$$

where $Group_i$ is the number of Group individuals in location i, $Group_{Population}$ is the total number of Group individuals in the population, and the two groups are Africans and Whites in our case.

In contrast, exposure measures the degree of potential contact between the two groups. One widely used measure of exposure is the Isolation index, which measures the expected probability of interacting with a member of the same group. Its formula, for example for a White individual, is:

$$Isolation = \sum_{i\in I}\frac{White_i}{White_{Population}}\frac{White_i}{Total_i} \tag{2}$$

where $Total_i$ is the total population of location i.

Since we cannot observe the local network structure, our approach is pragmatic and relies on the Isolation index. Beyond its attractive axiomatic properties,[2] this choice is motivated by econometric identification assumptions that are discussed below.

The isolation index is simply a function of population shares (the population share of each group in the considered subunit, multiplied by the constant population share of the same group in the supra-unit). Therefore, using the index is equivalent to using a functional transformation that concentrates the information on population shares. It directly reflects the presumed probability that an individual of a certain group would interact with another individual of the same group in the location. For peer effects and racial preferences arguments, it directly indicates with whom an individual is likely to have most of her/his social interactions. Its appropriateness is less obvious for the segmented labour market argument, while if labour networks are based on the own racial group, the probability described by the index should usefully proxy the capacity of the network to convey information or provide job opportunities. Moreover, all the socio-economic mechanisms considered in Section II assume that Africans and Whites not only differ in their characteristics, but also that these characteristics do not provide the same return when it comes to determining income levels. Therefore, having a segregation index able to identify with whom an individual is interacting and providing some asymmetric flexibility of effects is essential. Dissimilarity/Evenness indices cannot identify these kinds of social effects as they only quantify how much people should move to correct an imbalanced compo-

[2]It is notably asymmetric (Massey and Denton 1988) and respects the Independence and School Division properties (Frankel and Volij 2011).

sition at the supra-unit level. In particular, to these measures, it doesn't matter whether Africans or Whites are moved from one location to another.

III.B. Segregation and Income Decomposition

The aim of this paper is to quantify the contribution of segregation to the income gap distribution. Oaxaca-Blinder type decompositions can assist in quantifying additive contributions of variables. They can also suggest explanations by factors or reciprocal links. As is typical in decomposition approaches (DiNardo et al. 1996; Sherer 2000), selection or endogeneity issues are not addressed and there is no causal interpretation of the decomposition, in general. The role of decomposition methods is to provide a preliminary examination of the data, perhaps before specifying a causal or a structural model that would include the factors found with substantial contributions. This descriptive-predictive approach is endorsed, for instance, in the survey of Fortin et al. (2011, pp. 96-97) on decomposition methods.

In a linear setting, the difference in mean income Y between two groups, A and B, can be decomposed as:

$$\mathbb{E}[Y_A] - \mathbb{E}[Y_B] = (\mathbb{E}[X_A] - \mathbb{E}[X_B])\beta_A + \mathbb{E}[X_B](\beta_A - \beta_B)$$

where the composition effect, $(\mathbb{E}[X_A] - \mathbb{E}[X_B])\beta_A$, stems from the average difference in the characteristics X between the two groups, and the structure effect, $\mathbb{E}[X_B](\beta_A - \beta_B)$, comes from the difference in the coefficients β between the two groups (e.g., Jann 2008). In particular, this simple adding-up property is automatically satisfied in the above standard Oaxaca-Blinder decomposition that relies on linear regressions to describe the means of the compared distributions. This is also the case when examining unconditional quantiles with RIF regressions because the last stage of their estimation is a linear regression. Each of the expectations and parameter vectors that appear in the above decomposition must be estimated from some data, which may involve usual sampling, estimation, specification, and measurement errors. Accordingly, we examine the potential specification error associated with the usual omission of the segregation variable in Mincer equations.

More generally, decomposition methods make possible some quantitative assessment of the relative contributions of each covariate to the gap between the distributions of the two groups. In this paper, we focus on the contribution of the local segregation variable, while controlling for essential explanatory factors of earnings: the education and experience of the individuals. A debatable, albeit rather common, interpretation of the structural component is as a measure of the discrimination in the labor market.[3]

Thus, although the fundamental meaning of the used segregation indicator is clear – it refers to the probability of meeting locally someone from one's own group – its connection with the local population shares of each group varies with the group (e.g., in terms of the proportion of Whites for a White individual, and the proportion of Africans for an African individual). This is expected as the indicator manifests that each of these population shares has different consequences for Africans and Whites. This is not only because the proportion of one's own group is to be reversed when the reference group is changed, but also because this proportion could be associated with different mechanisms. For example, Africans may suffer for living in a disadvantaged ghetto confined to their own group, while Whites may instead benefit from residing a in privileged neighborhood, which may function like a club around their own group. In particular, one may generally consider that predominantly African neighborhood are poor, while predominantly White neighborhood are wealthy.

Therefore, the segregation index, as any other variable, may capture some influence of unobserved correlated characteristics that may not only affect incomes, but also operates differently for the two groups. Besides, one could extend these reflections to the education and experience variables that could be associated with unobserved group-specific mechanisms (e.g., education programs that rarely promotes African role models). This issue, often overlooked, while analyzing wage gaps, is therefore neither new nor specific to the study of segregation impact.

A first perspective on this issue is to consider that the included variables in a decomposition analysis legitimately represent not only their specific (causal or reciprocal) effects, but also their correlations with unobservable factors. For example, finding that the contribution of education is essential to understand wage gaps does not preclude to pursue later the analysis with causal models of education, e.g. as: an input in a productivity process, a signal of ability in the labor market, a correlation with family background in communities, etc. In that sense, observing a sizeable contribution of education may be only a first step towards more specific, and structural, investigations of its causal effects. The same can be said about the contributions of segregation. Some of the potential mechanisms, or causes, of its contribution have been evoked in Section II. Another potential channel is that segregation is related to selectivity. The role of segregation in the selectivity in the labor market can be declined on similar lines that what was already discussed in Section II. Selectivity through determination of the residence location is another distinct possibility, through internal or external

[3]For instance, Sherer (2000, p.319), referring to the structure effect, states that: "*To determine the extent to which earnings differentials represent labour-market discrimination, Blinder-Oaxaca decompositions are constructed using the output from OLS earnings regressions.*".

migrations, and would directly affect segregation measures. These features remain largely hidden behind the results of the decomposition analysis.

However, a way to attenuate an excessive influence of unobserved local group-specific characteristics or mechanisms is to include local fixed effect in the model for each considered subplace, and even to interact these fixed effects with each group dummy. Even though these fixed effects are unobserved, they can still be scrutinized in the formulae of decomposition analyses. Since we do not avail of panel data, the fixed effects cannot be eliminated. However, they can be included and examined in a Oaxaca-Blinder decomposition. Let

$$y_{it} = f_{s(i)} + X_{it}' b_{G(i)} + \epsilon_{it},$$

where $b_{G(i)}$ is the coefficient vector for the group $G(i)$ of individual i, $f_{s(i)}$ is a fixed effect for the subplace $s(i)$ of individual i, and ϵ_{it} is an error term. Namely, $G(i)$ = A (respectively, B) if i is in A (B). Then,

$$\mathbb{E}[Y^A]–\mathbb{E}[Y^B] = (\mathbb{E}[X^A]–\mathbb{E}[X^B]) \cdot b_A + \mathbb{E}[X^B] \cdot (b_A–b_B) + \{\mathbb{E}[f_{s(i)} \cdot \mathbb{I}_{[\text{i in A}]}] - \mathbb{E}[f_{s(i)} \cdot \mathbb{I}_{[\text{i in B}]}]\},$$

where Y^A (Y^B) is the income in group A (B), similar exponent notations for X, and the curly bracket term is the new contribution brought by the fixed effects.

Therefore, the term $\mathbb{E}[f_{s(i)} \cdot \mathbb{I}_{[\text{i in A}]}] := f^A$ (respectively, $\mathbb{E}[f_{s(i)} \cdot \mathbb{I}_{[\text{i in B}]}] := f^B$) summarizes the mean fixed characteristics of the subplaces of group-A (group-B) individuals, weighted by the number of these individuals in each subplace. What is interesting here is that the mean effect of the characteristics of the subplace, for each group, only appears in a constant term, which is therefore not obviously connected to the effects of the X or the variations of the coefficients, for each group. In that case, the interpretation of the relative contributions of the observed factors does not change, providing f^A and f^B are not or little related to $(\mathbb{E}[X^A]–\mathbb{E}[X^B]) \cdot b_A$ or to $\mathbb{E}[X^B] \cdot (b_A–b_B)$. Note that since $f^A - f^B$ can be seen as the difference in coefficients of a variable always equal to 1, there is no associated composition effect in that case.

If, in addition, group-specific fixed-effects, $f^A_{s(i)} \cdot \mathbb{I}_{[\text{i in A}]}$ and $f^B_{s(i)} \cdot \mathbb{I}_{[\text{i in B}]}$, are instead assumed and included in the model, respectively for each group, one obtains for the new contribution term:

$$\{\mathbb{E}[f^A_{s(i)} \cdot \mathbb{I}_{[\text{i in A}]} \cdot \mathbb{I}_{[\text{i in A}]}] + \mathbb{E}[f^B_{s(i)} \cdot \mathbb{I}_{[\text{i in B}]} \cdot \mathbb{I}_{[\text{i in A}]}] - \mathbb{E}[f^A_{s(i)} \cdot \mathbb{I}_{[\text{i in A}]} \cdot \mathbb{I}_{[\text{i in B}]}] - \mathbb{E}[f^B_{s(i)} \cdot \mathbb{I}_{[\text{i in B}]} \cdot \mathbb{I}_{[\text{i in B}]}]\}$$
$$= \mathbb{E}[f^A_{s(i)} \cdot \mathbb{I}_{[\text{i in A}]}] - \mathbb{E}[f^B_{s(i)} \cdot \mathbb{I}_{[\text{i in B}]}] := f^A - f^B,$$

since the two groups are exclusive. Therefore, we have a similar formula as before, and only the intercept is affected by the introduction of fixed effects. Similar reflections apply to the structural effects in distribution decompositions. This shows that examining the contribution of the intercept to the structural effect, often overlooked in distribution analyses, is also interesting in that it reflects the role of group-specific (or not) missing factors.

However, a segregation measure that is only specific to the location (such as the Dissimilarity Index), instead of being also specific to the group (like the Isolation index), would ignore different segregation contexts for distinct groups. For example, in South Africa, for analyzing the link between segregation and income, an all-African township is sociologically and economically different from an all-White suburb.

As we need an asymmetric measure of segregation,[4] we primarily use the Isolation Index. Since, in our application, segregation is measured using the initial information taken from the 2001 Census and is fixed for all individuals, it is consistent with the idea that segregation may act on income over relatively long run. This allows the measure of segregation to be the same in the two studied periods. Of course, segregation measures may imperfectly reflect equilibrium since households are mobile across space, notably with African residential locations being less restricted after the end of Apartheid. However, describing residential segregation as relatively fixed is a common convention as residency is much less flexible than wage setting and job changes.

IV. Context and Data

IV.A. Segregation

There is a long history of racial segregation in South Africa (Thompson 2001). ‘Color bar’ discriminatory legislation, against Africans and other non-White inhabitants, was in force from the early days of the Union of South Africa. This culminated during the Apartheid period, with a nationwide pol-

[4] In the words of Frankel and Volij (2011) (p.6): *“Although [Symmetry] is a standard property which is satisfied by most indices, it may not be suitable for work that focuses on the problems that face a particular ethnic group. For instance, if one is interested in the social isolation of blacks from all other groups, then one may prefer an index that treats blacks differently.”*

icy of separate socio-economic development supported by the Afrikaner minority (Thompson 2001, Chap. 5-6; Giliomee 2003). The 1950 Population Registration Act categorized and recorded racial identities on individual identification documents into 'Blacks', 'Whites', 'Coloureds' and 'Indians'.[5] The 1950 Group Area Act allocated separate settlement regions to distinct races. A permit was needed to cross the internal borders of racial regions, which contributed to stabilizing the population composition of each region. Under the 1953 Reservation of Separate Amenities Act, the different races had access to separate hospitals, universities and other public amenities. The 1953 Bantu Education Act introduced separate schools for different races. Over time, several additional laws restricted citizen's movements within the country. In practice, Africans were often kept apart from cities and towns, unless they could justify their presence there with a work permit. Although racial spatial segregation has declined since the end of the Apartheid in 1994, it remains common de facto and would not fail to strike any casual observer traveling across the country. The coincidence of these facts with the history of discriminatory remunerations along ethnic lines suggests that different work compensation rules, notably with respect to ethnicity, may have been and are still implemented in low- and high-segregation areas.

IV.B. Racial income gap

At the demise of Apartheid in 1994, people's aspirations and expectations turned toward greater economic equality and improvements in their standards of living. However, the next decade was instead characterized by increasing inequality (Leibbrandt et al. 2012), poverty traps (Adato et al. 2006), and anti-poor growth (Özler 2007). South Africa is one the most unequal countries in the world, with especially large racial gaps in living conditions.

Over the 1993-2006 period, aggregate inequality increased (Agüero et al. 2007). Then, Statistics South Africa (2017) notes that, while the Gini index modestly declined from 0.72 to 0.68 over 2006-2015, it has remained stable since 2009. Most of the former increase in aggregate inequality is associated with an increase in within-group inequality (Leibbrandt et al. 2012), especially for Africans (Özler 2007). Despite an initial reduction in within-group inequality after 2006, by 2015, every group had nearly returned to its original income level (Statistics South Africa 2017). In contrast, evidence regarding between-group inequality is rather scarce. Leibbrandt et al. (2010b) find an increase in it, whereas Leibbrandt et al. (2012) report a decreasing contribution of it to aggregate inequality. However, this is relative to an extreme maximal counterfactual, which does not imply that

[5] We employ these categories, except replacing 'Blacks' with 'Africans' as, in post-Apartheid South Africa, 'Blacks' refers to all the non-Whites groups together.

between-group inequality, in absolute terms, has become low. Finally, the emergence of an African middle class is a major evolution in the South African society. However, Bhorat and Khan (2018) claim that the size of the phenomenon might have been overestimated.

The post-Apartheid land reform redistributing some land from Whites to Africans could have reduced the gap. However, despite many political discourses and mediatic claims, very little distribution took place so far. [6] Therefore, land redistribution can only have had a negligible impact, and is not worth including in the analysis.

IV.C. Data

Our first source of data is the community profiles from the 2001 South African Census (Statistics South Africa 2003). They provide the total counts of the South African population aggregated at geographic levels ranging from enumeration areas to provinces. The data are exhaustive but only provide summary statistics about the distributions of some socio-demographic characteristics within each location.

In South-Africa, there are 278 municipalities: 8 metropolitan, 44 districts and 226 local municipalities. These administrative units are focused on local economies and provision of public services and infrastructures. Municipal districts constitute the second layer of the South African geographic frame designed and maintained by the Municipal Demarcation Board and used for the 2001 census. They include the 8 Metropolitan areas or Category A municipalities (Cape Town, Buffalo City, Nelson Mandela Bay, Mangaung, Ethekwini, Ekurhuleni, Johannesburg, and Tshwane) and 44 Category C municipalities (also called district councils or district municipalities). Metropolitan areas are defined as ”conurbations featuring high population density; intense movement of people, goods and services; extensive development; and multiple business and industrial districts. Other features include a complex and diverse economy, a single area where integrated development is desirable and strong interdependent social and economic linkages between its constituent units exist” (Statistics South Africa 2015). Category C municipalities represent a group of (Category B) local municipalities. The most notable difference between metropolitan areas and Category C municipalities lies in their distribution of powers and functions, the formers have ”exclusive municipal executive and legislative authority in its area” (South Africa 1996) while the latter focus on the district-wide missions such as planning, infrastructure development, or transport, and let Category B municipalities manage local aspects like public service provisions.

[6] See Zimmerman (2000); Sonneborn (2010); Kepe and Hall (2016) for more details on land redistribution.

Subplaces constitute the fifth layer of the geographic frame. They were designed initially to answer a need for more disaggregated data in the first post-Apartheid Censuses. Subplaces represent suburbs, sections, sub-villages and some small non-urban areas, which usually correspond to local housing markets. There is no legislative or executive authority embodied into subplaces. They are on average inhabited by 2000 inhabitants (mean=2109.86, 1st quartile=351, median=768, 3rd quartile=2048) For comparison with the familiar US context, subplace would be equivalent to US census tracts, and municipal districts with US metropolitan statistical areas.

Labor markets may stretch across subplaces, including for commuting workers at least occasionally. But it is unlikely that they stretch across Municipal districts and most low-wage informal jobs are very local. It is most likely that all people in a subplace access the same schools and health services, and receive the same public goods.

Electoral wards are all included into the municipal districts. They are at a comparable scale than subplaces, but there is no correspondence between these two layers. In fact, they are entirely designed by two different authorities, Statistics South Africa for the subplaces, and the Municipal Demarcation Board for the electoral wards.

The level of analysis is the Municipal district. However, a subunit level is needed to compute a segregation index as such an index is based on comparing the composition of the subunit relative to the supra-unit composition. Having a too aggregated subunit level may make cities appear more integrated than they really are because ghettos become diluted into the mass of more integrated subunits. Subplaces, by representing local housing markets well, constitute a good approximation to a neighbourhood where individual preferences, peer effects, or spatial mismatch will play a role.

Our second data source is the National Income Dynamics Study (NIDS hereafter), which is an individual panel data survey conducted every two years with a nationally representative sample. There are four waves available that cover the period 2008-2014. However, we will use only the 2008 (SALDRU 2018) and 2014 (SALDRU 2016) waves to avoid the short-term fluctuations due to the 2008-2009 economic crisis[7] that may obscure the contributions of the main regressors in the decomposition. Data on incomes are usually considered relatively reliable (Leibbrandt et al. 2012). But, the sub-sample sizes by racial groups are sometimes relatively small.

The NIDS provide information about individual characteristics and income, while the community profiles serve for calculating the measure of segregation in each municipal district, subdivided by subplaces to obtain a more precise sense of local segregation. The lowest geographic sampling level

[7] According to a report on poverty levels by Statistics South Africa (2017), *"The number of people living below the food line increased to 15,8 million in 2009 from 12,6 million in 2006, before dropping to 10,2 million people in 2011."*

in the NIDS is the municipal district. As we are interested in studying income differences, we restrict our sample to individuals older than 15 who report a positive total personal monthly real income. Amounts are deflated to November 2014 rands with the CPI. Income is measured as the monthly take-home pay from the main job. Other secondary sources of income are excluded to avoid contaminating the analysis with substantial measurement errors, notably from omissions and non-responses. The design of the NIDS explicitly separates self-employed from employees on which we concentrate. As a result, our base sample consists of 2922 Africans and 440 Whites in 2008 and 5291 Africans and 229 Whites in 2014. In the Online Appendix, we provide evidence of robustness with an alternative sample definition addressing issues with seasonal and part-time workers, workers older than the retirement age, and early retirement. Given the small number of White female workers in the sample (99), we prefer to discuss our main results by pooling the two genders. However, we will also briefly turn to separate (and less statistically significant) estimations by gender in subsection VI.C, and detailed results by gender are provided in the Appendix.

We focus on the African-White gap only, as these are the two most prominent groups in South Africa. Whites occupy the highest economic positions, while Africans are the most disadvantaged group and crystalized the fear of the Afrikaner minority during Apartheid. Though also often discriminated against, Coloureds and Asians stand economically between Africans and Whites. Table I reports the mean and standard deviation of the variables used in the analysis, across ethnic groups and survey rounds. As expected, Whites are generally more educated, older, and richer than Africans. They usually have more interactions with the other group, as shown by the statistics on isolation. In the next section, we report the results of the decomposition.

[Table 1 about here.]

V. Mean Analysis

We assume that expected incomes are determined by the individuals' education and experience,[8] possibly quadratically. Then, we augment this specification with a measure of segregation. As stated above, our measure of segregation is fixed in the year 2001 because we cannot measure segregation from the NIDS and have to rely on a measure constructed from the 2001 Census.[9] However, since segregation might affect income levels with a delay, it does not seem unreasonable to adopt this

[8]We measure education and experience as the number of years of completed education and of experience (age minus years of completed education minus 6) while measuring them in months instead does not change the results.

[9]We also considered a measure of segregation coming from the 2011 Census (adjusted to the 2001 administrative boundaries) for 2014. It produced very similar results.

approach. For example, bad habits may develop over several years before becoming ingrained. Our model takes the following form for each individual i:

$$
\begin{aligned}
Income_i = \alpha + \beta_1 \times Education_i + \beta_2 \times Education_i^2 + \beta_3 \times Experience_i \\
+ \beta_4 \times Experience_i^2 + \beta_5 \times Segregation(2001)_i + \epsilon_i
\end{aligned}
\tag{3}
$$

where $\alpha, \beta_1, \beta_2, \beta_3, \beta_4$, and β_5 are parameters to estimate, and ϵ_i is a centered error term. We first run this OLS regression separately for Africans and Whites. The results are displayed in Table II.

[Table 2 about here.]

Several expected effects emerge. We find a positive and significant effect of experience for both groups with decreasing returns, as the coefficient for its square is negative. However, as experience is a function of age, these coefficient estimates might also capture a life-cycle phenomenon, older people being generally wealthier than their younger counterparts. The effect of isolation on mean income is positive for Whites and negative for Africans, although it loses its significance in 2014 for Whites. Finally, the effect of education is dominated by the quadratic term, which yields an overall positive effect for Whites and for Africans with more than 6 years of schooling, which concerns at least 72 percent of Africans aged 15 or older. This U-shape pattern may be explained by the skills mismatch characterizing the South African labor market.

Then, we decompose the mean, using the pooled sample as the reference group,[10] in Table III, to elicit the magnitudes of the roles of the different correlates, notably segregation.[11] The income gap between the two groups corresponds to the difference between the mean predicted incomes of the two groups obtained via OLS.

First, consistent with the high and rising inequality levels observed since the end of Apartheid (Agüero et al. 2007; Leibbrandt et al. 2012), the average real monthly income gap between Whites and Africans is considerable. It corresponds to 6658 rands in 2008 and rises to 6886 rands in 2014, almost twice the national minimum wage in 2019.

[Table 3 about here.]

The magnitudes of the composition and structure effects are comparable. Despite the emergence of an African middle class, Africans continue to lag behind on many socio-economic characteristics.

[10] Using the Whites coefficient as the reference does not change the main results.

[11] To avoid transferring part of the structure effect into the composition effect, we add a group dummy to the pooled model for the decomposition (Jann 2008).

A sizable and significant composition effect is thus an expected finding. However, the finding that this composition effect is roughly equal to the structure effect is less expected. It implies that Africans with similar socioeconomic characteristics as Whites benefit much less on average than Whites from these characteristics and that this matters as much as the differences in socio-economic characteristics. This might be a consequence of racial discrimination in the job or housing markets (Kain 1968). Alternatively, it might reflect different work habits between Africans and Whites or different professions and activity sectors. For instance, if Africans work mainly in rural areas or the industrial sector, having a master's degree might give them access to a lower wage than Whites working in the financial service sector in an urban area. Thus, the racial gap in returns to education might signal a premium for urban areas and/or the financial sector. Over time, the share of the composition effect increases from 46.4 percent in 2008 to 52.7 percent in 2014.

When we examine more closely the detailed decomposition, we first note that all the groups of variables contribute positively to the gap through the composition effect. Education is the main contributor to the composition effect, accounting for 95.5 percent of the effect in 2008 and for 73 percent in 2014. This reinforces our discussion above: Africans lag behind Whites in terms of education and experience.[12]

The results for segregation present a different pattern. Its mean composition effect is close to zero and not significant in 2008, whereas it is positive and significant in 2014. However, the contribution of segregation in the composition effect increased by more than fivefold between 2008 and 2014, and while, in 2008, it represents around 3 percent of the total composition effect, it accounts for more than 15 percent in 2014, being the second-greatest contributor to the composition effect, after education. The 2008-2009 economic crisis, which struck South Africa during the last quarter of 2008, and the following turmoil drove many people into poverty.[13] As the most deprived are usually the most isolated, this might explain the massive increase of the role of segregation between the two periods.

In the structure effect, segregation emerges as the main relative contribution to the total structure effect. Its contribution is of similar order to education in 2008, although education reduces the gap

[12]As the contributions of each factor sum up to the total effect, the total contribution of education is the sum of the contribution of the education variable and that of education squared. This is also true for experience and the distributional analysis in the next section.

[13]"it is clear that the [2008 global financial] crisis was particularly tough on those most deprived in our society" [...] "The last five years, notably between 2011 and 2015, have been a rough economic rollercoaster for South Africa" [...] "This period has seen the financial health of South African households decline under the weight of [this] economic [pressure] and, in turn, pulled more households and individuals down into poverty." (Statistics South Africa 2017, p.14 and 16)

while segregation increases it, and segregation is the only significant contributor in 2014. Education comes second, accounting for 30 percent of the structure effect in 2008, while its effect is not precisely estimated.This may be partly due to the dual school system inherited from Apartheid. For instance, in 2009, grade three pupils in formerly White schools outperformed grade five pupils in formerly African schools on a standardized test designed for grade three students (Spaull 2013). This suggests that segregation may be partly responsible for the contribution of education through this duality, and provide an additional motivation for including segregation as a control. When interpreting these figures, we should bear in mind that the constant term still represents a large share of the structure effect (17.6 percent in 2008 and 37.7 percent in 2014), hinting at substantial group-specific hidden factors.

VI. Decomposing Income Distributions

VI.A. Distribution Analysis

The main interest in a distribution analysis of the racial income gap is as a device for investigating the heterogeneity of the effects of segregation on this gap. As before, we pursue an agnostic perspective on endogeneity and selectivity phenomena. As a matter of fact, the distribution analysis may provide hints about where in the distribution these issues may matter most.

Then, instead of comparing the distribution means of the two groups and decomposing the mean gap, we compare the marginal distribution quantiles of the two groups for the same quantile index (for example, for the median). In that case, the composition effect still solely describes the effect of the differences in the characteristics between the two groups, while permitting the comparison for the same given quantile index in the two distributions.

We depart from common approaches by decomposing the income distribution with RIF regressions (Firpo et al. 2009) instead of the reweighting approach (DiNardo et al. 1996). The reweighting approach suffers from path-dependence in the detailed decomposition, which does not sum to the aggregate decomposition. RIF regressions are much simpler and perform better in practice for detailed decompositions.

VI.B. Decompositions

The estimation results of the detailed Oaxaca-Blinder decompositions applied to the $\mathrm{RIF}(y, q_\tau)$ dependent variable are presented in Figure I. In the top-left panel, we first display the quantiles of the racial gap in log income [14] for the two studied years. For both 2008 and 2014, the gap is always significantly different from zero and keeps the same sign for all quantiles. This first-order stochastic dominance result implies that mean income and income inequality among Whites is unambiguously higher and lower, respectively, than the corresponding aggregate income indicators for Africans, in both years. Finally, 2014 also first-order stochastic dominates 2008, within each racial group, which confirms the unambiguous improvement of each of these two income distributions over the studied period.

[Figure 1 about here.]

However, we observe two distinct patterns. In 2008, the income gap increases from the bottom quantile to the median and decreases thereafter. In 2014, it is relatively stable from the second decile to a little before the sixth decile and then declines as we approach the top of the distribution. More important, at any quantile, the income gap is always smaller, in log points, in 2014 than in 2008, but at a higher real income level than in 2008, which corresponds to an increase in the income gap. For instance, at the first quartile, the income gap of 1.45 log points in 2008 and 1.25 log points in 2014 coincides to gaps of 4026 rands in 2008 and 4390 rands in 2014.

We test the null hypothesis of no differences between quantiles levels in each year with t-tests on each quantile index. Dashed quantiles represent quantiles for which the null hypothesis is not rejected. The decline in log income differences occurs only significantly for the intermediate quantiles ranging from the 32nd to the 68th income quantile. Therefore, this slower increase of the income gap for middle classes might be linked to the emergence of an African middle class.

We report the aggregate decomposition of the racial gap, for each year and each quantile, in the top-right panel of Figure I. This allows us to disentangle the differences in observed characteristics from the influences of market and social mechanisms that are captured by differences in the parameters. The dashed parts of the curves represent quantile composition and structure effects that are significantly different from zero at the 5 percent level. In 2008, the structure effect continuously decreases with quantiles, while the composition effect is increasing and plateaus near the 6th decile. At the upper end of the distribution, the structure effect actually contributes to reducing the income gap. In terms of relative size, the structure effect is slightly larger than the composition effect up to the 35th quantile. In 2014, the pattern is similar, although the two elicited components of the income gap are much closer and their change over quantiles much slower. Thus, the magnitude of the

[14] In the remainder of the paper, we refer to 'income' for simplicity, but it should be understood that we employ the natural logarithm of income.

composition effect overtakes that of the structure effect beginning at the median, with the latter not contributing at all after the 65th quantile. This suggests that the hidden mechanisms that separate the incomes of Whites and of Africans operate primarily among the lower classes of these groups. This particularity will be exploited below in the analysis of the minimum wage reform.

To complete this description, t-tests are performed to compare the structure and composition effects in 2008 with their respective counterparts in 2014. Then, we examine whether the structure effect is significantly different from the composition effect in 2008 and in 2014. Regarding the temporal trend, there are no significant variations for the structure effect, except potentially for a small group of quantiles after the median. This stability over time may indicate relatively permanent socio-economic mechanisms, some of which might be linked to segregation. For the composition effect, a notable decline over time is observed from the 33rd quantile. Thus, the relative importance of each effect has changed over time. In 2008, the two effects are significantly different except around the 35th quantile. However, in 2014, both the structure and composition effects contribute equally below the median. Ultimately, the observed reduction in the income gap observed for the middle quantiles appears to be driven primarily by the reduction in the composition effect.

We delve deeper into the detailed relative contribution of each factor to the composition (Figure I, middle panels) and the structure (bottom panels) effects. In 2008, experience does not play any role in the composition effect. Education is the most important contributor to the composition effect, followed by segregation, with the former representing twice the latter's contribution across almost the entire distribution.[15] Both are increasing throughout the distribution, except after the 6th decile, after which the contribution of education slightly decreases and that of segregation continues to increase. This parallel pattern explains the increasing contribution of the composition effect across quantiles, and when education decreases, segregation compensates for its reduction to form the plateau observed. In 2014, each contribution is ranked similarly as in 2008, but experience now contributes positively to the income gap from the first quintile, although it remains the smallest contributor. Education's contribution is stable across quantiles up to the median, at which point it begins to rise to its maximum around the third quartile, and declines slightly thereafter. In 2014, segregation's contribution slowly decreases until the median before recovering from its minimum around the third quartile. Then, it plateaus until it spikes dramatically in the very top quantiles. Both the rise and decline of education's contribution from the median and the tremendous spike exhibited by segregation at the very top materialize directly in the aggregate composition effect. The relative stability of

[15] A t-test suggests that the contribution of education is significantly equal to twice that of segregation up to the 84th quantile.

the latter in the first half of the distribution comes from the contributions of education and experience compensating for the weakening of the contribution of isolation. As is typical in quantile analyses, substantial variations at extreme quantiles are suspected to be statistical artifacts due to the restricted sample sizes used in the calculations for these quantiles.

Regarding the structure effect, the ranking of the contributors differs drastically from that for the composition effect. Segregation is now the dominant factor at almost all quantiles, before experience, followed by education. The intercept parameter is specific to the structure effect and bears a precise interpretation in this context. Usually, in mean regressions, the intercept is viewed as the average income level individuals obtain once the effects of the other covariates have been removed. In quantile regressions, it is instead the minimum income level at the specified quantile regardless of the effect of other covariates. Thus, in the decomposition, a significant difference between two intercepts may suggest intrinsic discrimination between the two groups. However, one cannot infer anything about the origin of this statistical discrimination, whether it is true racial discrimination inherited from Apartheid or something else related to omitted factors. In both 2008 and 2014, the contribution of the intercept is positive and significant for approximately 20 percent of the population above the 6th decile. However, in both years, this positive contribution is systematically compensated for by a negative contribution of the same magnitude from education. The two terms statistically cancel out throughout the distribution, which is a consequence of the additive normalization of the decomposition and of the insignificance of the contributions of the other factors at these quantiles. This feature suggests that the unobservable fixed differences between Africans and Whites, as long as incomes are concerned, are strongly negatively correlated with education levels. For our analytical purpose, it is reassuring that these unobserved factors seem to be much more associated with education than with segregation. Overall, the residual structure effect, after accounting for the intercept gap, is due primarily to segregation and experience. In both years, segregation contributes positively in the lower halves of the distributions and loses significance for the upper halves. Experience follows a similar pattern, except that it contributes negatively to the income gap above the 6th decile in 2008. A plausible explication for these findings is that in a particularly harsh dualistic labor market for Africans, experience in better jobs represents a signal of reliability and skills for Africans, whereas low-productivity Whites might be protected by discrimination. Alternatively, it is possible that affirmative action legislation adds a premium on experienced African workers. However, this effect disappears in 2014.

VI.C. RIF Regressions

To better understand the structure effect, we now examine the estimation results of the RIF regressions used for the above decompositions. The estimation results are displayed in Figure II for each group, each year, and across quantiles. The partial relationship between education and income is identical in 2008 and 2014 for Africans, indicating that the variation in the structure effect is essentially due to changes in the returns to education of Whites. Moreover, the incomes of Africans display little sensitivity to their education level, whereas the incomes of Whites obey a more complex educational pattern, which is particularly pronounced at the top of their distribution. This might reflect greater heterogeneity in bargaining power for highly educated Whites occupying top positions.

[Figure 2 about here.]

Regarding experience, the pattern for Africans is similar in both years across quantiles. It differs only by its level. In 2008, the linear part is slightly higher, while the quadratic term is slightly lower but only negatively significant from the 4th decile. In 2014, the linear part is not positively significant before the 3rd decile, while the quadratic part is negatively significant after the median. Therefore, Africans enjoyed some small linear experience premium in the bottom of the distribution in 2008, while it vanished for the first quartile by 2014, presumably due to the 2009-2010 economic crisis. At the top of the distribution, the marginal returns to experience are decreasing with quantiles, but slightly less in 2014 than in 2008. In both years, Whites always experienced a better marginal return to experience, the only exception is the reversal of the linear component of experience at the top of the distribution in 2008, which explains the negative and significant contribution to the structure effect.

The most interesting lesson from these RIF regressions concerns the relationship between segregation and income. Segregation is negatively associated with income only for Africans at the bottom of the distribution and in the lower-middle class (up to the median in 2008 and to the 6th decile in 2014) in both years. On the other hand, it is positively associated with income for Whites in 2014 in the upper half of the distribution. It appears to have a positive effect for all Whites in 2008. Hence, the structural effect of segregation is substantial below the median because the gap between the quantile effects of the two groups is at its maximum. Then, it loses significance as the quantile effect for Africans fades away for the upper quantiles. This suggests that the economic mechanisms at work behind the effect of segregation are most likely different for Africans and Whites, and thus, policies addressing this concern for segregation should also differ. For the intercept coefficients, which reflect unobserved group-specific (or not) fixed effects, they are estimated significantly positive at almost all quantiles. For Africans, these intercepts are very stable across quantiles and virtually identical in 2008 and 2014. For Whites, the intercept coefficients steadily increase across quantiles. They are often higher and more irregular in 2014 than in 2008. In the next section, we exploit the 2018 minimum wage reform to shed some lights on potential explanations of the effect of segregation.

Disaggregating the data by gender[16] (Figure III) confirms that segregation increases the earning gap across races. Moreover, although the coefficient of the education and experience variables, and the intercept, differ a lot between genders, approximately the same lessons emerge when comparing races. This is another argument for regrouping the genders in the analysis. However, gender interacts with segregation, as the negative effect of segregation is mostly supported by African women at the bottom of the distribution (Left panel), while the pattern of this effect changed over time. In 2008, African women around the median and slightly above suffered the worst experience of segregation, whereas, in 2014, the most affected stood slightly below the 20th quantile. For White women, the effect of segregation is not significant, although this may result from their small sample. The positive effect of segregation at the top of the distribution mostly corresponds to White men (Figure III, right panel). For African men, both in 2008 and 2014, the segregation effect is almost constant and close to zero, except for a small positive effect between the 60th and 70th quantiles in 2014.

[Figure 3 about here.]

We test the robustness of the segregation measure by using the dissimilarity index[17] instead of the isolation index (Figure IV). The results are very similar. The main difference is now that segregation is also negatively associated with income at the top of the African distribution. Whites still experience a positive association at the top of their income distribution.

[Figure 4 about here.]

VI.D. Sorted effects

In Table IV, we describe the main winners and losers from segregation: the "winners" (respectively "losers") are defined as being the 10 percent Africans most positively (respectively negatively) affected by segregation. Following Chernozhukov et al. (2018), we explore in this way the heterogeneity of the partial association of segregation and incomes by scanning the characteristics of the winners and losers from segregation.

[16] We describe the other factors and the decompositions in the Online Appendix since these effects do not change qualitatively.

[17] We display the results for the other factors and the decompositions in the Online Appendix as they convey the same message.

[Table 4 about here.]

The main African losers from segregation are mostly little educated workers, often female, living in large households with few pecuniary resources. They mostly obtained their job, often in the agriculture and mining sectors, through their extended network, and are ready to work for a wage much below the National Minimum Wage proposed by the African National Congress (ANC), which is the governing and main party in South Africa. This might be because with a little more than 8 years of education on average, they have merely completed mandatory schooling.[18] They are concentrated in KwaZulu-Natal, Gauteng, Eastern Cape, and Limpopo.

On the opposite, the main African winners from segregation are mostly male workers with a household income more than twice the National Minimum Wage. Almost 40 percent were recruited via their extended network and are unionized. On average, they work more hours than the main losers. They are also more educated, which may explain why they can ask for higher salaries as reflected by their reservation wages. However, they have fewer than 12 years of education, on average, a level which would have earned them only a matriculation diploma. Therefore, they are probably as likely as the main losers to have quitted schooling without any valuable diploma. Contrary to Cutler et al. (2008), this rules out education as an important dimension of heterogeneity for this analysis. Finally, African winners are mostly living in Gauteng, KwaZulu-Natal, and Mpumalanga, but are only overrepresented in Gauteng and Mpumalanga. They work mainly in Social services, Finance, Manufacturing, and Wholesale and Retail trade.

These results provide another perspective on the above-discussed mechanisms of the association of segregation and incomes. The main losers rely excessively more on their personal connections (both close or distant) to find a job than the main winners. This would suggest a negative influence of within-group networks backed up by local segregation, at least for some disadvantaged workers. There is also a large difference in unionization rates in favor of the main winners. This may be indicative of some segmentation of the labour market but could also point at the quality of the private network of the latter.

The regional and industrial differences between losers and winners fit well the electoral constituency of the ANC and the political strategy of the EFF (Economic Freedom Fighters, the main far left opposition party). The traditional ANC strongholds are the northern provinces of Limpopo, Mpumalanga, and North West, and the Eastern Cape in the south. Gauteng, the Western Cape, the

[18]Schooling is mandatory in South Africa from age six turning seven to the age of 15 or completion of grade 9, whichever comes first.

Northern Cape, and, to a lesser extent, the Free State have always been disputed with the Democratic Alliance (the main opposition party, centre-right). KwaZulu-Natal is currently administered by the ANC, while challenged by the IFP (Inkatha Freedom Party, a traditionalist opposition party dominated by the Zulu). The mostly rural Eastern Cape, Limpopo, and KwaZulu-Natal, simultaneously with the agricultural sector partially correspond to the ANC's constituency since the ANC largely relied on the rural vote from its inception.[19] Moreover, workers in the agricultural sector are rarely unionized, which coincides with the low unionization of the main losers.[20] Workers in the mining sector, on the other hand, are close to the main South African union (Congress of South African Trade Union, COSATU), of which the National Union of Mineworkers is a member, founded by the currently sitting president Cyril Ramaphosa in the 1980s. The COSATU has always been supportive of the ANC.

Before January 2019, the legal minimum wage varied across activity sectors, ranging in 2015 from 1813 rands per month for Domestic Workers to 2844 rands per month for Contract Cleaners Bhorat et al. (2016). In 2015, it was implemented to 39 percent of formal employees.

In 2018, the National Minimum Wage Bill was passed (for its enforcement in 2019) with the support of the ANC members of parliament and the opposition of the other parties. The minimum wage was set at 3500 rands per month, for 40 worked hours per week. While there are doubts about its practical implementation, given the limited capacity of the monitoring agency, it is still a major economic and political shock. For comparison, the median salary of workers covered by sector agreements was approximately 2447 rands per month, and 3400 rands per month for all workers in the formal sector Bhorat et al. (2016). According to the COSATU, 47 percent of the South African labor force should have benefited from the reform.[21]

The policy target group regroups all individuals with an income below the National Minimum Wage. This is consistent with the ANC's supporters profiles found in a recent survey citizensurveys.net (2018) that shows that low-income individuals are overrepresented when compared to supporters of the other two main parties.

We examine how changes in the minimum wage accord with the intervals of quantiles, for each group, in the graphs of quantile decomposition. It can be argued that, post reform, only the truncated

[19] See Afrobarometer.org (2015) , last accessed on the 14th of November 2019.

[20] Thompson (2001) describes the early formation of trade unions in South Africa consecutive to the rise in the cost of urban living. In 1945, 40% of the unionized workers were employed in commerce and manufacturing, and "the crucial terrain for labor relation was, as ever, the mining industries." (p.179). Unionization in agriculture, far from the urban centers, and being heavily mechanized or of the subsistence type, cannot easily develop.

[21] See AllAfrica.com (2018), last accessed on the 14th of November, 2019

section of the curves that exceed the quantiles corresponding to the considered minimum wage should apply. This provides us with a quick and simple graphical diagnostic device.

Of course, some caution must be taken. In particular, if the reform deeply changes the data generating processes of incomes in the country, it may be harder to deduce insights from the curves that may themselves change with the reform. However, if it is assumed that this is not the case and that, overall, the current relationship of incomes with education, experience and, specifically in our case, local segregation will not be substantially affected by the reform, then the graphs can be used to identify the populations most likely to be affected by the reform.

This identification strategy can be compared with the first identification assumption in Chernozhukov et al. (2013, pp. 2236-2237). These authors assume, for the US, that the conditional density of wages below or at the minimum wage depends only on the value of the minimum wage; that the minimum wage has no effect on unemployment; and that there are no spillover effects onto wages above the minimum. While all of these assumptions are debatable, they provide a benchmark for minimum wage effects. Our approach can be seen as another simplifying perspective that nonetheless assumes less stringent rigidity of the studied phenomena across quantiles.[22]

Under these diagnostic rules, the government reform would lead to the elimination of precisely the cases in which the socio-economic mechanisms involving the segregation variable significantly affects the racial income gap. This can be seen in Figure V right panel for the structure effects, and especially in Figure V left panel for the coefficients of segregation in the RIF regressions for the Africans. Although causal studies would be needed to confirm them, these results hint at the possibility that the minimum wage reform might cancel, or at least substantially reduce, the factors that make local segregation contribute to the wage gap between Whites and Africans through harming African workers

[Figure 5 about here.]

The EFF's strategy is to outbid the ANC and the trade unions. As a consequence, their minimum wage proposal (12500 rands) much exceeds realistic earnings for the poorest workers, especially in the agricultural and mining sectors. It should speak more to individuals working in the sectors of manufacturing, wholesale and retail trade, and construction, hence, mostly urban workers within the EFF's constituency.[23] Indeed, most other political movements made minimal wage proposals too low

[22] In another approach, simulations based on Computable General Equilibrium models have been attempted for assessing the effect of minimum wages in South Africa (Pauw and Leibbrandt 2012). They seem to display a negligible impact of these policies on poverty.

[23] See citizensurveys.net (2018), last accessed on the 14th of November, 2019.

to affect many of these workers directly. On the whole, although there are many other determinants of political programs, the ANC-led reform benefits the ANC's constituency in several ways. On the one hand, it raises the wages of the workers employed in the modern and public sectors, for income categories that predominantly vote for the ANC. On the other hand, by removing wage intervals corresponding to especially harmful associations of segregation and wages, it should contribute to protect some of the main African losers of segregation. However, the reform is also likely to badly damage the employment prospects of low-wage workers.

VII. Conclusion

We proposed a new approach to analyzing the contribution of segregation to socio-ethnic income gaps, as a general preliminary descriptive stage to more specific structural or causal analyses. Performing a thorough distribution decomposition and scrutinizing the heterogeneity of segregation effects allow us to uncover patterns that remain hidden in mean analyses and traditional Mincer equations. The highlighted contribution of segregation to the distribution of the racial income gap not only promotes the integration of segregation mechanisms in earnings models, but also generates hints at the socio-economic mechanisms that explain income differences between socio-ethnic groups.

Segregation was found to be the main contributor to the structure effect, ahead of education and experience, in the case of the distribution of African-White income gap in South Africa. More precisely, segregation affects negatively African bottom incomes, but positively White top incomes. Hints about a few operating mechanisms became apparent, notably through the identification of the subpopulations that suffer and benefit the most from segregation through their wages. The worst impaired are the low-education workers in agriculture and mining, often female, and locked-up in their local informal networks. Finally, we also examined the 2018 minimum wage reform, which, beside its direct effect on their incomes, was found likely to attenuate the harmful impact of segregation on poor Africans by wiping out intervals of the African distribution where most of this impact takes place.

Policies can be often specified with regard to whether they alternatively address composition or structure effects of segregation. Policies against the composition effect of segregation should correspond to measures that change the levels of segregation locally, while keeping relatively fixed the levels of the other factors. For example, since segregation is observed to be a sizeable contributor to the composition effect, aids to migration and to settlement off own-group areas, subsidies to leaving own-group areas, lower tax rates in ethnically mixed subplaces, public investment in mixed residential real estates, should diminish segregation levels without affecting education and experience levels. The most emblematic example of such policies is the Moving To Opportunity experiment (Katz et al. 2001) in the United States which offered vouchers to inhabitants in disadvantaged neighbourhoods to move to a more affluent neighbourhood. In South Africa, the Department of Human Settlements is in charge of urban planning at the national level. It supports the creation of more inclusive areas and offer housing subsidies to the poor. The much awaited land reform could also fall in this category.

In contrast, policies addressing the structure effects of segregation should change the mechanisms through which segregation determines earnings, but not the level of segregation, at least initially. Referring to the evoked mechanisms in this work, such policies may change: racial preferences, neighborhood and peer effects, segmented job networks and labor markets, and spatial mismatch between workers and firms. Racial preferences can be made more tolerant by public propaganda, including anti-racism campaigns in the media, or by deontological rules and bonus to realtors, encouraging them to promote mixed neighborhood. The Truth and Reconciliation Commission headed by Desmond Tutu and the numerous gestures of Nelson Mandela toward the White community (his support to the national rugby team during the 1995 World Cup for instance) are prominent examples. To break the noxious influence of peers transmitting bad habits locally one may foster the entry of 'role models' that would counteract them. Moreover, external information spread through modern communication tools (internet, cell phones) may mitigate the impact of noxious information disseminated by peers. Open recruitment practices in local firms can also be stirred by government inspections to reduce the role of peers in the labor market. Finally, laws against discriminating recruitment or on-the-job discrimination may be passed and enforced particularly in misbehaving neighborhoods. This could defeat labor market segmentation. In this regard, South Africa passed several Affirmative Action bills (Employment Equity Act (1998), Broad-Based Black Economic Empowerment Act (2003), completed by Codes of Good Practice (2007).). Banning union monopoly may also help to avoid close shops supporting bad working habits confined to some groups. Better dissemination of information about the labor market and job openings, including in township and from non-local areas, would also limit segmentation processes.

Of course, there may be policies that change both segregation levels and segregation mechanisms. In particular, endogeneity and selection issues may arise when conducting some policies, as they may also affect other dimensions (e.g., access to school), which would complicate matters.

Finally, the estimated relationships in this work can assist in designing more effectively these diverse kinds of policies, for example by better targeting social programs towards the disadvantaged segregated categories, either on the basis of the segregation level of the subplace where they reside, or of the predicted negative effect of segregation on their wage.

How could the findings of this work be extended to other contexts? Directly leaping to generalization in other countries is always bold in empirical analyses. In other contexts, segregation effects may or may not be essential for wage determination, or may or may not share the features that have been found in South Africa.[24] Obviously, specific data for these contexts are needed for serious investigations. Moreover, there may be several reasons to be cautious when trying to generalize the results of this study. First, segregation was systematic, severe and present everywhere for a large part of the South African history and is still much widespread. This is not so for many countries. Second, South Africa is a country in which one of the groups, the Africans, is large compared to others. In that case, the proportion of Africans in most districts, municipalities, subplaces, is high, and they are therefore more likely to be observed as segregated than the Whites, even under random assignment. This has consequences for some of the evoked mechanisms. For example, individual preferences may be more sensitive to the presence of Whites since the later are rarer. On the contrary, peer effects may be more powerful among White communities since their peers are more special in that case. Finally, rare valuable job information may be more useful in White communities, than common and basic information in large African communities. It may be that in other countries where the Blacks are a minority and the Whites a majority, such as in the US, other facets of these mechanisms would instead emerge. A general dominance of one group in the population may be common to some other countries, but not to all them, and many complex phenomena may occur with balanced group population sizes, such as political polarizations.

To go further in the analyses, the proposed approach should make room for structural and causal studies of incomes. In that case, concerns typically encountered are the endogeneity of human capital factors and selection by labor market participation.[25] In our case, these issues would extend to the potential endogeneity of segregation and differential migrations of racial groups to their preferred specific neighborhoods. Addressing these issues is an important challenge for future research, for which the current analysis has already provided valuable clues.

[24]For example, insignificant neighborhoood effects on self-sufficiency are found in the United States in Kling et al. (2007).

[25]Leibbrandt et al. (2010a) provide evidence of the potential importance of selection as an explanation of the declines in real incomes after the end of Apartheid.

DECLARATIONS OF INTERESTS

Declarations of interest: none.

REFERENCES

Adato, Michelle, Michael R. Carter, and Julian May, "Exploring Poverty Traps and Social Exclusion in South Africa Using Qualitative and Quantitative Data," *Journal of Development Studies*, 42 (2006), 226–247.

Afrobarometer, (https://afrobarometer.org/fr/publications/ad65-south-africans-increasingly-dissatisfied-their-elected-leaders-performance), 25 (2015), last accessed on the 14th of November, 2019.

Agüero, Jorge, Michael R. Carter, and Julian May, "Poverty and Inequality in the First Decade of South Africa's Democracy: What can be Learnt from Panel Data from KwaZulu-Natal?" *Journal of African Economies*, 16 (2007), 782–812.

Alesina, Alberto, Reza Baqir, and William Easterly, "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics*, 114 (1999), 1243–1284.

AllAfrica, (https://allafrica.com/stories/201804250851.html), (2018), last accessed on the 14th of November, 2019.

Altonji, Joseph, and Rebecca Blank, "Race and Gender in the Labor Market," in *Handbook of Labor Economics*, Orley Ashenfelter and David Card, eds. (Elsevier Science, 1999).

Bayer, Patrick, and Kerwin Kofi Charles, "Divergent Paths: A New Perspective on Earnings Differences between Black and White Men Since 1940," *Quarterly Journal of Economics*, 133 (2018), 1459–1501.

Bebbington, Anthony J., D. Mitlin, J. Mogaladi, M. Scurrah, et al., "Decentring Poverty, Reworking Government: Social Movements and States in the Government of Poverty," *Journal of Development Studies*, 46 (2010), 1304–1326.

Becker, Gary, *The Economics of Discrimination*, 2nd ed. (Chicago: University of Chicago Press, 1957).

Bénabou, Roland, "Workings of a City: Location, Education, and Production," *Quarterly Journal of Economics*, 108 (1993), 619–652.

Bhorat, Haroon, Tara Caetano, Benjamin Jourdan, Ravi Kanbur, et al., "Investigating the Feasibility of a National Minimum Wage for South Africa," DPRU Working Paper Series, WP201601, 2016.

Bhorat, Haroon, and Safia Khan, "Structural Change and Patterns of Inequality in the South African Labour Market," DPRU Working Paper Series, WP201801, 2018.

Blumer, Herbert, "Race Prejudice as a Sense of Group Position," *The Pacific Sociological Review*, 1 (1958), 3–7.

Chernozhukov, Victor, Iván Fernández-Val, and Blaise Melly, "Inference on Counterfactual Distributions," *Econometrica*, 81 (2013), 2205–2268.

Chernozhukov, Victor, Iván Fernández-Val, and Ye Luo, "The Sorted Effects Method: Discovering Heterogeneous Effects Beyond their Averages," *Econometrica*, 86 (2018), 1911–1938.

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz, "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment," *American Economic Review*, 106 (2016), 855–902.

Chetty, Raj, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter, "Race and Economic Opportunity in the United States: An Intergenerational Perspective," *Quarterly Journal of Economics*, 135 (2020), 711–783.

Citizen Surveys, (https://citizensurveys.net/citizen-surveys-how-many-white-eff-supporters-exist (2018), last accessed on the 14th of November, 2019.

Cutler, David M., and Edward L. Glaeser, "Are Ghettos Good or Bad?," *Quarterly Journal of Economics*, 112 (1997), 827–872.

Cutler, David M., Edward L. Glaeser, and Jacob L. Vigdor, "When Are Ghettos Good or Bad? Lessons from Immigrant Segregation in the United States," *Journal of Urban Economics*, 63 (2008), 759–774.

Dickens, William T., and Kevin Lang, "A Test of Dual Labor Market Theory," *American Economic Review*, 75 (1985), 792–805.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux, "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, 64 (1996), 1001–1044.

Edin, Per-Anders, Peter Fredriksson, and Olof Aslund, "Ethnic Enclaves and the Economic Success of Immigrants: Evidence from a Natural Experiment," *Quarterly Journal of Economics*, 118 (2003), 329–357.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux, "Unconditional Quantile Regressions," *Econometrica*, 77 (2009), 953–973.

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo, "Decomposition Methods in Economics," in *Handbook of Labor Economics*, David Card and Orley Ashenfelter, eds., vol. 4, chapter 1 (Amsterdam: Elsevier, 2011), 1–102.

Frankel, David M., and Oscar Volij, "Measuring School Segregation," *Journal of Economic Theory*, 146 (2011), 1–38.

Giliomee, Hermann, "The Making of the Apartheid Plan, 1929-1948," *Journal of Southern African Studies*, 29 (2003), 373–392.

Gradín, Carlos, "Race, Poverty and Deprivation in South Africa," *Journal of African Economies*, 22 (2012), 187–238.

———, "Race and Income Distribution: Evidence from the USA, Brazil and South Africa," *Review of Development Economics*, 18 (2014), 73–92.

Ioannides, Yannis M., and Linda Datcher Loury, "Job Information Networks, Neighborhood Effects, and Inequality," *Journal of Economic Literature*, 42 (2004), 1056–1093.

Jann, Ben, "The Blinder-Oaxaca Decomposition for Linear Regression Models," *The Stata Journal*, 8 (2008), 453–479.

Kain, John F., "Housing Segregation, Negro Employment, and Metropolitan Decentralization," *Quarterly Journal of Economics*, 82 (1968), 175–197.

Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman, "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment," *Quarterly Journal of Economics*, 116 (2001), 607–654.

Kepe, Thembela, and Ruth Hall, "Land Redistribution in South Africa," Commissioned report for High Level Panel on assessment of key legislation and the acceleration of fundamental change, Cape Town, 2016.

Kling, Jeffrey R., Lawrence F. Katz, and Jeffrey B. Liebman, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 75(1), 2007, 83–119.

Leibbrandt, Murray, James A. Levinsohn, and Justin McCrary, "Incomes in South Africa after the Fall of Apartheid," *Journal of Globalization and Development*, 1 (2010a).

Leibbrandt, Murray, Ingrid Woolard, Arden Finn, and Jonathan Argent, "Trends in South African Income Distribution and Poverty Since the Fall of Apartheid," OECD Social, Employment and Migration Working Papers, 101, 2010b.

Leibbrandt, Murray, Arden Finn, and Ingrid Woolard, "Describing and Decomposing Post-Apartheid Income Inequality in South Africa," *Development Southern Africa*, 29 (2012), 19–34.

Levy, Brian L., "Wealth, Race, and Place: How Neighborhood (Dis)advantage From Emerging to Middle Adulthood Affects Wealth Inequality and the Racial Wealth Gap" *Demography*, 59 (1), 293–320.

Magnac, Thierry, "Segmented or Competitive Labor Markets?" *Econometrica*, 59 (1991), 165–187.

Magruder, Jeremy R., "Intergenerational Networks, Unemployment, and Persistent Inequality in South Africa," *American Economic Journal: Applied Economics*, 2 (2010), 62–85.

Massey, Douglas S., and Nancy A Denton, "The Dimensions of Residential Segregation," *Social Forces*, 67 (1988), 281–315.

Mhlongo, Niq, *Black Tax: Burden or Ubuntu* (Jonathan Ball Publishers SA, 2019).

Oreopoulos, Philip, "The Long-Run Consequences of Living in a Poor Neighborhood," *Quarterly Journal of Economics*, 118 (2003), 1533–1575.

Özler, Berk, "Not Separate, Not Equal: Poverty and Inequality in Post-Apartheid South Africa," *Economic Development and Cultural Change*, 55 (2007), 487–529.

Pauw, Karl, and Murray Leibbrandt, "Minimum Wages and Household Poverty: General Equilibrium Macro-Micro Simulations for South Africa," *World Development*, 40 (2012), 771–783.

Schelling, Thomas C., "Dynamic Models of Segregation," *Journal of Mathematical Sociology*, 1 (1971), 143–186.

Sherer, George, "Intergroup Economic Inequality in South Africa: The Post-Apartheid Era," *American Economic Review Papers and Proceedings*, 90 (2000), 317–321.

SALDRU, National Income Dynamics Study (NIDS) Wave 1, Database, Version 7.0.0, Pretoria: SA Presidency [funding agency], Cape Town: SALDRU [implementer], Cape Town: DataFirst [distributor], 2018, DOI: https://doi.org/10.25828/e7w9-m033.

SALDRU, National Income Dynamics Study 2014-2015, Wave 4, Database, Version 2.0.0, Pretoria: Department of Planning, Monitoring, and Evaluation, SA Presidency [funding agency], Cape Town: SALDRU [implementer], Cape Town: DataFirst [distributor], 2016, DOI: https://doi.org/10.25828/f4ws-8a78.

Sonneborn, Liz, *The End of Apartheid in South Africa* (Chelsea House, 2010).

The South African Constitution, Chapter 7 Local Government, 1996.

Spaull, Nicholas, "South Africa's Education Crisis: The Quality of Education in South Africa 1994–2011," Technical Report, 2013.

Statistics South Africa, Census 2001 Community Profiles in Super Cross, Database, Version 1, Pretoria: Statistics South Africa, 2003.

Statistics South Africa, South African Census 2011, 10% sample, Census 2011 Geography Metadata, 2015.

Statistics South Africa, "Poverty Trends in South Africa: An Examination of Absolute Poverty Between 2006 and 2015," Technical Report, Johannesburg, 2017.

Thompson, Leonard, *A History of South Africa*, 3rd ed. (New Haven, CT: Yale University Press, 2001).

Wilson, William Julius, *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policies* (Chicago: University of Chicago Press, 1987).

Yinger, John, "Measuring Racial Discrimination with Fair Housing Audits: Caught in the Act," *American Economic Review*, 76 (1986), 881–896.

Zimmerman, Frederick J., "Barriers to Participation of the Poor in South Africa's Land Redistribution," *World Development*, 28 (8), 1439–1460.

TABLE I

DESCRIPTIVE STATISTICS

	2008				2014			
	Whites		Africans		Whites		Africans	
	Mean	Std D.	Mean	Std D.	Mean	Std D.	Mean	Std D.
Income	12909.28	13268.02	3534.29	5777.53	14012.4	13514.05	4542.55	5065.90
Isolation	0.65	0.12	0.93	0.09	0.64	0.12	0.93	0.08
Dissimilarity	0.84	0.08	0.85	0.07	0.85	0.07	0.86	0.07
Years of schooling	12.82	2.03	8.92	4.36	13.28	1.83	10.55	3.66
Experience	21.48	11.69	22.68	13.23	24.03	13.07	20.19	12.73
Male	0.51	0.50	0.57	0.50	0.57	0.50	0.50	0.50
Age	40.31	11.68	37.60	11.12	43.31	13.23	36.74	11.01
Weak link	0.28	0.45	0.35	0.48	0.38	0.49	0.45	0.50
Strong link	0.06	0.24	0.07	0.25	0.04	0.19	0.04	0.20
Union membership	0.20	0.40	0.26	0.44	0.23	0.42	0.27	0.44
Household size	3.08	1.25	4.16	3.11	3.25	1.60	4.48	3.22
Household income	23931.78	20486.24	5868.31	8556.82	27279.05	26478.98	7411.73	9150.55
Hours worked weekly	42.82	13.64	36.36	20.41	41.56	14.70	41.59	15.89
Reservation wage	.	.	.	.	12580.69	14510.48	4774.53	5633.57
Firm size (50+)	.	.	.	.	0.30	0.46	0.25	0.43
Western Cape	0.42	0.49	0.06	0.23	0.37	0.48	0.05	0.22
Eastern Cape	0.03	0.18	0.09	0.29	0.07	0.25	0.08	0.28
Northern Cape	0.06	0.24	0.05	0.22	0.07	0.26	0.04	0.20
Free State	0.00	0.07	0.09	0.28	0.01	0.09	0.08	0.27
KwaZulu-Natal	0.05	0.22	0.26	0.44	0.09	0.29	0.27	0.45
North West	0.02	0.15	0.09	0.28	0.01	0.11	0.08	0.27
Gauteng	0.27	0.44	0.20	0.40	0.30	0.46	0.22	0.41
Mpumalanga	0.12	0.32	0.10	0.30	0.06	0.23	0.10	0.30
Limpopo	0.03	0.16	0.07	0.25	0.03	0.16	0.08	0.27
Agriculture	0.05	0.22	0.16	0.37	0.05	0.22	0.09	0.28
Mining	0.03	0.17	0.07	0.25	0.03	0.16	0.04	0.21
Manufacturing	0.15	0.35	0.17	0.38	0.14	0.34	0.11	0.32
Energy	0.01	0.10	0.01	0.09	0.02	0.12	0.01	0.11
Construction	0.03	0.17	0.07	0.25	0.05	0.21	0.08	0.27
Trade; Hotels	0.18	0.39	0.14	0.35	0.18	0.39	0.19	0.39
Transport	0.04	0.20	0.04	0.20	0.05	0.21	0.05	0.22
Finance; Real estate	0.17	0.38	0.08	0.26	0.21	0.41	0.10	0.31
Social services	0.34	0.47	0.26	0.44	0.29	0.45	0.31	0.46
Observations	440		2922		229		5291	

The variable *Income*, *Household income*, and *Reservation wage* are deflated to November 2014 rands. The dummy variables *Male*, *Weak link*, *Strong link*, *Union membership*, *Firm size*, and sectoral and provincial dummy variables are expressed as a share of the population. *Firm size* and *reservation wage* were only available in 2014. The variable *Strong link* corresponds to individuals getting their job via a household member. The variable *Weak link* corresponds to individuals getting their job via a friend/relative not in the same household. The sample comprises all individuals employed in a formal job and older than 15 who report a positive total personal monthly real income. Data: National Income Dynamics Study Wave 1 and 4.

TABLE II
OLS REGRESSIONS

	2008		2014	
	Whites	Africans	Whites	Africans
Experience	0.08***	0.04***	0.03**	0.02***
	(6.84)	(10.82)	(2.47)	(7.46)
Experience squared	-0.0014***	-0.0004***	-0.0003	-0.0002***
	(-5.87)	(-5.99)	(-1.26)	(-2.90)
Years of schooling	-0.17	-0.05***	-0.31	-0.06***
	(-1.53)	(-4.03)	(-1.13)	(-5.04)
Years of schooling squared	0.01***	0.01***	0.02*	0.01***
	(3.28)	(16.85)	(1.87)	(17.32)
Isolation	1.08***	-0.40**	0.55	-0.46***
	(3.70)	(-2.57)	(1.37)	(-3.22)
Constant	7.32***	6.66***	8.83***	7.29***
	(10.07)	(42.09)	(5.00)	(47.91)
Observations	440	2922	229	5291
R^2	0.355	0.346	0.279	0.259

The sample comprises all individuals employed in a formal job and older than 15 who report a positive total personal monthly real income. The dependent variable is the logarithm of the total personal monthly real income. *Isolation* is the average probability to interact with someone of the same racial group. Data source: National Income Dynamics Study Wave 1 and 4.
t-statistics in parentheses: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE III

OAXACA DECOMPOSITIONS

	2008	2014
Differential		
Prediction_1	9.09***	9.20***
	(214.76)	(161.92)
Prediction_2	7.70***	8.01***
	(455.18)	(626.79)
Difference	1.39***	1.20***
	(30.56)	(20.55)
Composition		
Experience	-0.01	0.06***
	(-0.57)	(4.13)
Education	0.64***	0.46***
	(20.78)	(14.81)
Isolation	0.02	0.11***
	(0.47)	(3.22)
Total	0.65***	0.63***
	(12.75)	(12.82)
Structure		
Experience	0.16	0.11
	(1.34)	(0.91)
Education	-1.13*	-1.76
	(-1.93)	(-0.99)
Isolation	1.05***	0.68**
	(4.80)	(2.53)
Constant	0.66	1.54
	(1.02)	(0.85)
Total	0.75***	0.57***
	(12.69)	(9.02)
Observations	3362	5520

Prediction_1 is the mean of the predicted logarithm of the real monthly income of Whites (in 2014 (Nov.) rands). *Prediction_2* is the mean of the predicted logarithm of the real monthly income of Africans (in 2014 (Nov.) rands). These predictions also correspond to each subgroup unweighted sample mean when the coefficients are estimated using OLS for each subsample. *Experience* and *Education* regroup both the initial variable and its square. *Isolation* is the average probability to interact with someone of the same racial group. The sample comprises all individuals employed in a formal job and older than 15 who report a positive total personal monthly real income. Data source: National Income Dynamics Study Wave 1 and 4.

t-statistics in parentheses: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE IV

CLASSIFICATION ANALYSIS - DIFFERENCE IN THE AVERAGE CHARACTERISTICS OF THE AFRICAN MAIN WINNERS AND LOSERS FROM SEGREGATION

	Main losers		Main winners		Difference		
	Mean	S.E.	Mean	S.E.	Mean	S.E.	P-value
Isolation	.942	.003	.929	.003	.013	.004	.001
Male	.3	.02	.64	.02	-.34	.03	< 0.001
Household size	4.95	.14	4.2	.14	.75	.19	< 0.001
Household income	2635	136.77	8693.49	320.81	-6058.49	347.09	< 0.001
Strong link	.07	.01	.02	.01	.04	.01	.0011
Weak link	.65	.02	.39	.02	.26	.03	< 0.001
Union	.05	.01	.37	.02	-.32	.02	< 0.001
Hours worked weekly	38.09	.67	43.91	.49	-5.82	.84	< 0.001
Reservation wage	2247.69	107.77	5583.54	185.32	-3335.85	209.68	< 0.001
Years of Schooling	8.32	.17	11.35	.13	-3.03	.21	< 0.001
Western Cape	.04	.01	.06	.01	-.03	.01	.0465
Eastern Cape	.13	.01	.06	.01	.08	.02	< 0.001
Northen Cape	.02	.01	.05	.01	-.03	.01	.0053
Free State	.09	.01	.08	.01	.01	.02	.6597
KwaZulu-Natal	.32	.02	.2	.02	.11	.03	< 0.001
North West	.08	.01	.07	.01	.01	.02	.646
Gauteng	.15	.02	.28	.02	-.13	.02	< 0.001
Mpumalanga	.06	.01	.13	.01	-.08	.02	< 0.001
Limpopo	.12	.01	.06	.01	.06	.02	< 0.001
Agriculture	.29	.02	.02	.01	.27	.02	< 0.001
Mining and Quarrying	.12	.01	.02	.01	.1	.02	< 0.001
Manufacturing	.01	.004	.09	.01	-.08	.01	< 0.001
Energy supply	.084	.012	.11	.01	-.02	.02	.2479
Construction	.002	.002	.02	.01	-.01	.01	.023
Wholesale and Retail trade	.05	.01	.09	.01	-.04	.02	.0213
Transport	.14	.02	.16	.02	-.02	.02	.4625
Finance	.02	.01	.07	.01	-.06	.01	< 0.001
Social services	.03	.01	.12	.02	-.1	.02	< 0.001

The "winners" (respectively "losers") are defined as being the 10 percent Africans most positively (respectively negatively) affected by segregation according to the value of their RIF-regression coefficient on isolation. The variables Household income, and Reservation wage are deflated to November 2014 rands. Isolation is the average probability to interact with someone of the same racial group. The variable Strong link corresponds to individuals getting their job via a household member. The variable Weak link corresponds to individuals getting their job via a friend/relative not in the same household. The sample comprises all the Africans employed in a formal job and older than 15 who report a positive total personal monthly real income and whose RIF-regression coefficient on isolation is among the 10% lowest or the 10% highest values. Data: National Income Dynamics Study Wave 4.

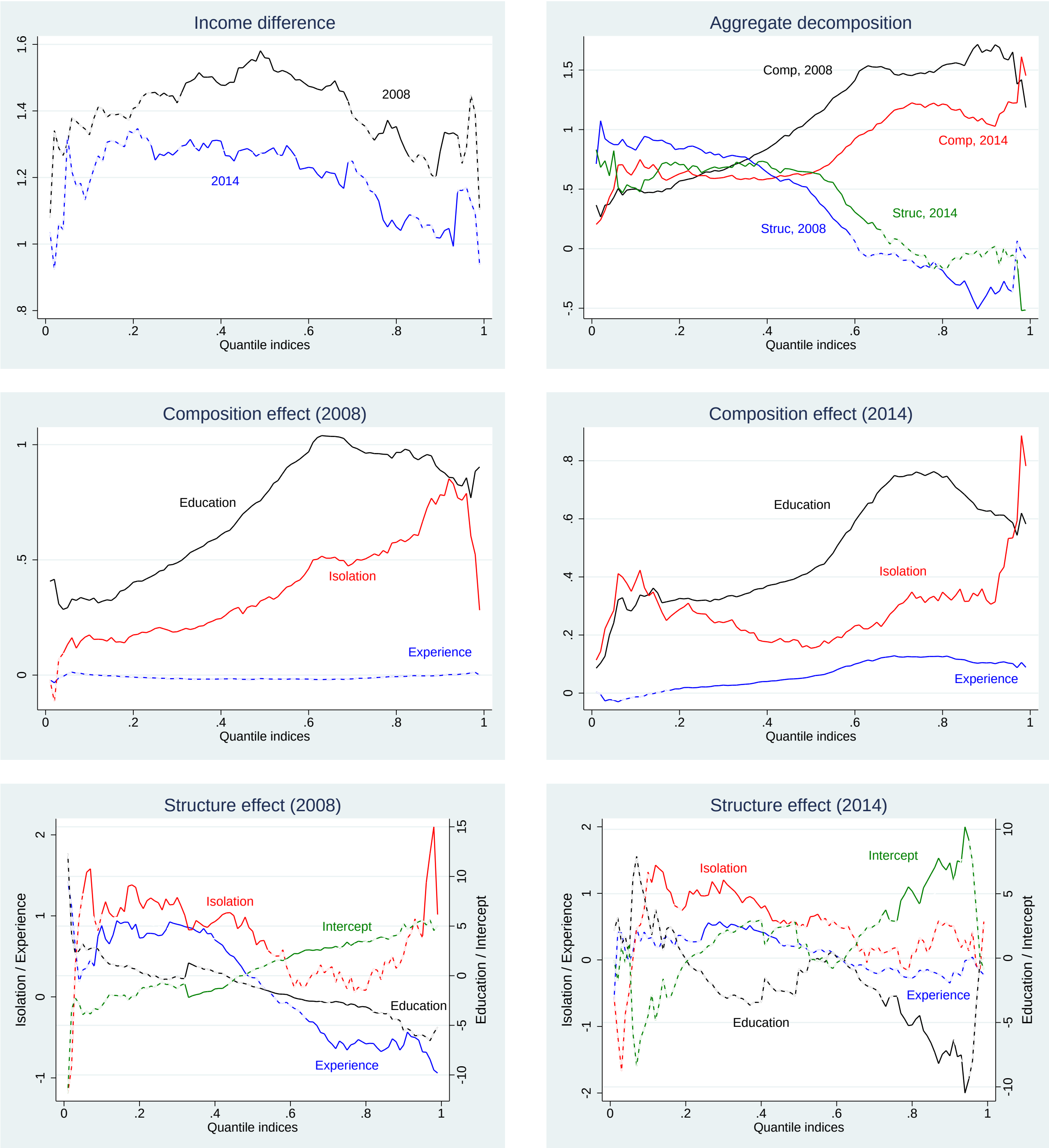

FIGURE I
Aggregate and Detailed Decompositions in 2008 and 2014
The top left panel represents the income quantile gaps in 2008 and 2014. The top right panel depicts the aggregate structure and composition effects for the two years. The middle panels show the factor-specific contributions to composition effects for each year. The bottom panels plots the factor-specific contributions to structure effects for each year. The dashed parts of the curves represent the effects not significantly different from zero at the 5 percent level, except for the income differences, where they represent the quantiles for which the income difference in 2008 is not significantly different from the income difference in 2014. "Comp" and "Struc" refer to the composition effect and the structure effect, respectively. The sample comprises all individuals employed in a formal job and older than 15 who report a positive total personal monthly real income: Whites (2008): n=440; Africans (2008): n=2922; Whites (2014): n=229; Africans (2014): n=5291. Data: National Income Dynamics Study Wave 1 and 4.

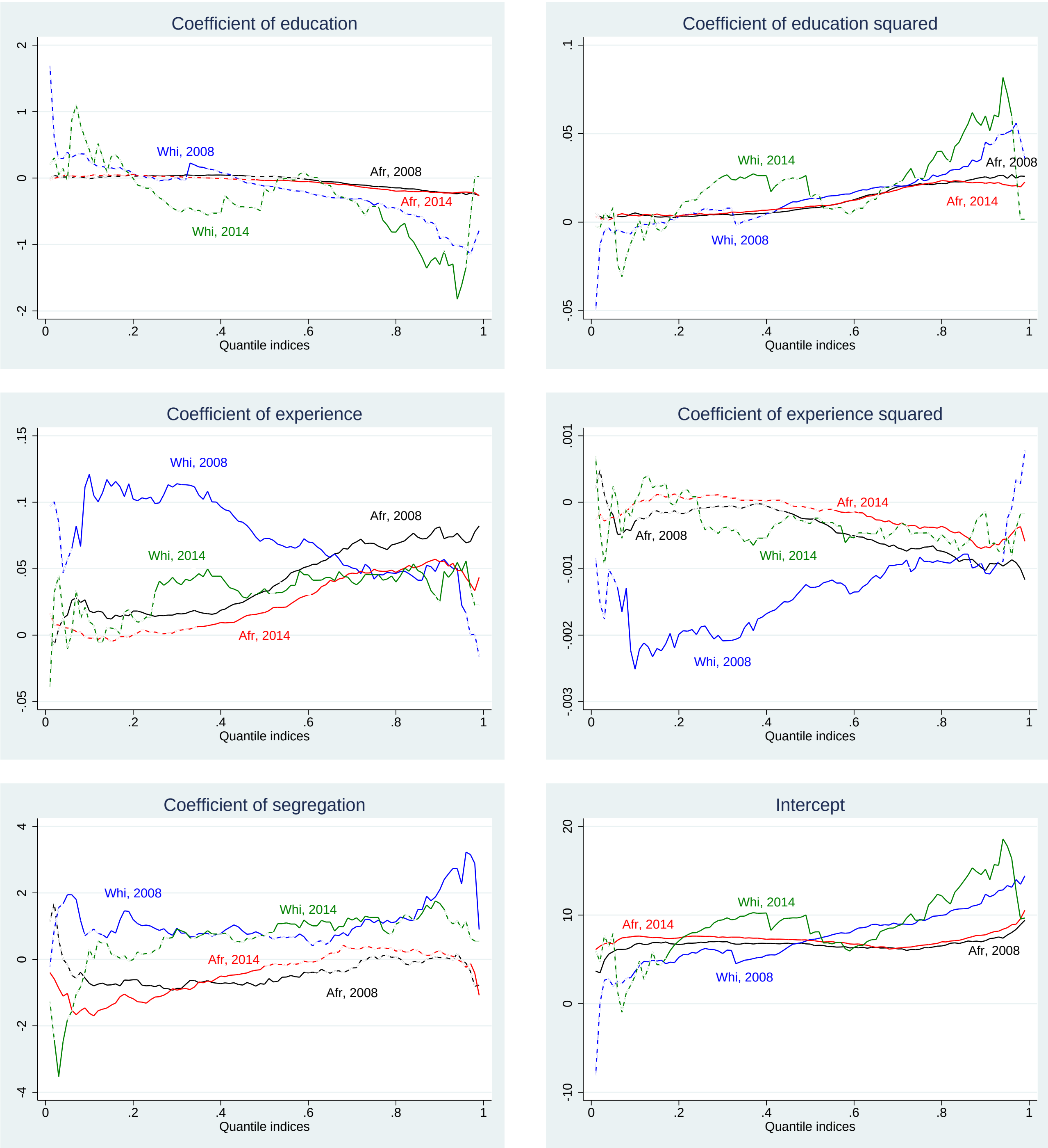

FIGURE II

RIF Regressions of Income by Racial Group and Year

The dashed parts of the curves represent the effects not significantly different from zero at the 5 percent level. "Afr" and "Whi" refer to Africans and Whites, respectively. The sample comprises all individuals employed in a formal job and older than 15 who report a positive total personal monthly real income: Whites (2008): n=440; Africans (2008): n=2922; Whites (2014): n=229; Africans (2014): n=5291. Data: National Income Dynamics Study Wave 1 and 4.

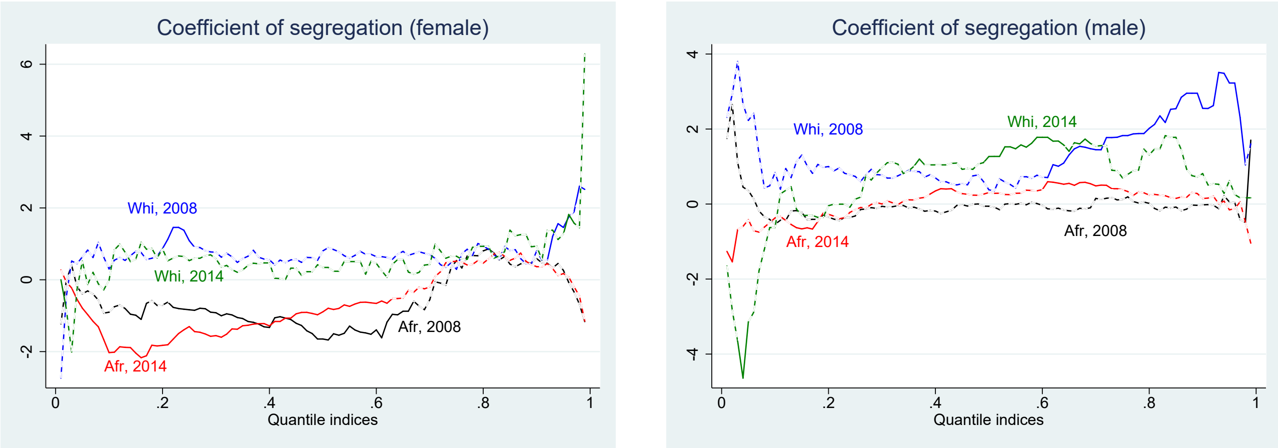

FIGURE III
RIF Regressions of Income by Racial Group, Year and Gender
The dashed parts of the curves represent the effects not significantly different from zero at the 5 percent level. “Afr” and “Whi” refer to Africans and Whites, respectively. The sample comprises all individuals employed in a formal job and older than 15 who report a positive total personal monthly real income. For females, we have the following sample sizes: Whites (2008): n=214; Africans (2008): n=1256; Whites (2014): n=99; Africans (2014): n=2627. For males, we have the following sample sizes: Whites (2008): n=226; Africans (2008): n=1666; Whites (2014): n=130; Africans (2014): n=2664. Data: National Income Dynamics Study Wave 1 and 4.

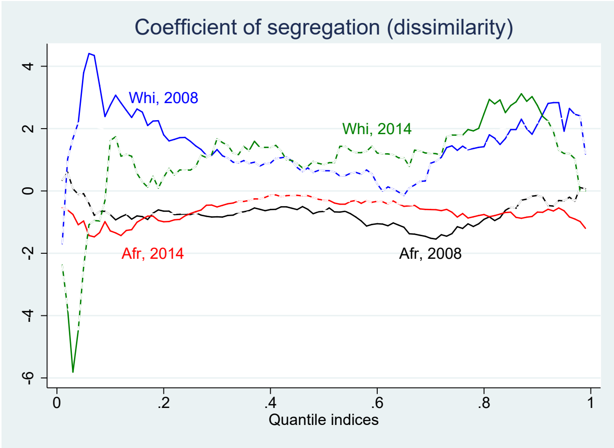

FIGURE IV
RIF Regressions of Income by Racial Group and Year (Dissimilarity index)
The dashed parts of the curves represent the effects not significantly different from zero at the 5 percent level. "Afr" and "Whi" refer to Africans and Whites, respectively. Segregation is measured with the dissimilarity index instead of the isolation index here. The sample comprises all individuals employed in a formal job and older than 15 who report a positive total personal monthly real income: Whites (2008): n=440; Africans (2008): n=2922; Whites (2014): n=229; Africans (2014): n=5291. Data: National Income Dynamics Study Wave 1 and 4.

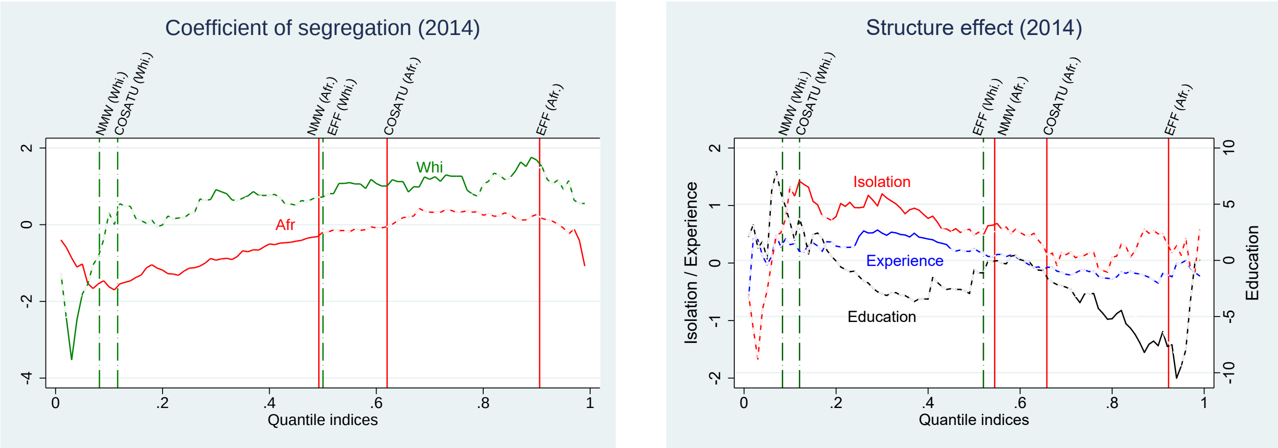

FIGURE V

RIF Regression (Segregation Coefficient) and Detailed Decomposition (Structure) in 2014, and Minimum Wage Proposals

The dashed parts of the curves represent the effects not significantly different from zero at the 5 percent level. “Afr” and “Whi” refer to Africans and Whites, respectively. The plain (respectively dashed) vertical lines represent the minimum wage proposals in the African (resp. White) income distribution of the government (NMW), the Conference of South African Trade Unions (COSATU) and Economic Freedom Fighters (EFF). The contribution of the intercept to the structure effect is not shown for the sake of clarity. See Figure I for the intercept contribution. The sample comprises all individuals employed in a formal job and older than 15 who report a positive total personal monthly real income: Whites (2014): n=229; Africans (2014): n=5291. Data: National Income Dynamics Study Wave 4.