Decomposition of income inequality by attributes: does the race matter in the US?

Frédéric Chantreuil * Thérèse Rebière †‡

Preliminary Version
Do not quote without the consent of the authors

.

October 3, 2016

Abstract

Using the Shapley income decomposition methodology, we isolate the pure racial contribution to income inequality in regards with the population of Blacks and Whites in the United-States over the period 2009-2014. Such a contribution can be seen as the direct perceived racial discrimination on earnings. We show that the pure racial contribution to income inequality is somewhat low, with about 1% to 4%, depending on the US administrative divisions used. The race tends to contribute more to total observed inequality in the West and South part of the US.

Keywords: inequality, conditional decomposition, Shapley value, racial discrimination.

JEL Codes: C71, D63, J15, J71.

*CREM-CNRS, University of Caen Normandy, MRSH, Esplanade de la Paix,, 14032 Caen cedex, France, and TEPP-CNRS; frederic.chantreuil@unicaen.fr

†Corresponding Author, LIRSA-CNAM and IZA, Department of Economics, 40 rue des Jeûneurs, Case 1D2P30, 75002 Paris, France; therese.rebiere@cnam.fr

‡We thank Isabelle Lebon for helpful remarks and comments.

1 Introduction

Income inequality in the USA, particularly regarding racial matter, is a hot topic in the United-States. A nationwide poll conducted in 2015 by CNN and the Kaiser Family Foundation, found that 49% of US-citizens think racism is "a big problem" in society today while they where only 28% in 2011. Since the 60s, the country administration multiplied measures in favor of racial equality especially on the labor market. Title VII of the Civil Rights Act of 1964 prohibits employment discrimination based on race, color, religion, sex and national origin. In 1965, President Johnson signed Executive Order 11246 promoting "affirmative action" that would ensure that all applicants may be employed, and that employees would be treated disregarding their race, creed, color, or national origin. Since then, these positive discrimination measures have been strongly criticized by the Conservative movement and several states decided to take actions against positive discrimination in favor of minority groups in public institutions.

Racial inequalities are often treated as a fact underlined by the wage gap between racial/ethnical groups. As noted by Sites and Parks (2011), and Couch and Daly (2002), racial income inequalities in the United States diminished significantly after World War II following the passage of the Civil Rights Act of 1964 and other measures aimed at reducing labor market discrimination, but have not changed much after 1974, the black-white wages gap remaining around 30% through the end of the 1980s, until a new convergence was observed in the 1990s. As for the recent years, the median annual income of a family in 2014 was of $76,658 for whites, $45,114 for hispanics - the first ethnic minority of the country - and $43,151 for blacks - the second ethnic minority. As a matter of poverty, 10.1% of whites individuals were below the poverty level against 23.6% of hispanics individuals and 26.2% for blacks (Economic Report of the President, 2016). As a matter of allocation, decades of sociological research showed that black-white inequality in local areas is greater where the black population is relatively large (see for instance Huffman and Cohen (2004)).

Contrary to the sociological literature, the economic literature on racial inequality in the USA is sparse. Using Juhn *et al.* (1991) decomposition techniques, Couch and Daly (2002) examine the role of individual characteristics, the employment structure, and overall wage inequality in reducing the racial income gap in the 90s. They find that greater occupational diversity and reductions in unobserved or residual differences explain partly this trend. The convergence therefore is partly due to equalization in the attainment of education and experience, and to the distribution of employment across industries and occupations rather than a pure ethnical matter. A fringe of the literature explain part of the observed income inequality by the reduction in unionization rates (see for instance Koniger *et al.* 2007 and Rosenfeld 2014). Some authors find major impact of unionization rates on the male Black-White gap resulting in a wage premium for blacks (Jones and Shmitt 2014), while others indicate that unionization would be beneficial to women (Rosenfeld and Kleykamp 2012). Chantreuil and Trannoy (2011, 2013), and Shorrocks (2013) used the Shapley value to determine the exact contribution of different

income sources to overall inequality. Our work is a contribution to this latter literature.

Defining racial inequality as the contribution of race to the overall inequality, the field of decomposition methods of income inequality measures appear to be an attractive framework appraising racial inequality. Among income decomposition methods, those inspired by the Shapley value seem particularly interesting[2], since they allow to explain income inequality determining the contribution of the various income sources or the contribution of different sub-populations to overall inequality. However, the Shapley decomposition methods developed so far do not permit the estimation of the share of overall wage inequality due to race. Actually, if the two sub-populations are made up of black and white respectively, the results of the decomposition may show that the contribution of black to the overall income inequality is 35 percent. In such a case, a policy equalizing the incomes of black people would reduce by 35 percent total income inequality. Hence, this decomposition framework does not allow the determination of the contribution of race to income inequality. In order to solve this drawback of the "classical" Shapley decomposition rule, Chantreuil and Lebon (2015) extended this framework to a third dimension, namely the decomposition of in- come inequality by attributes. Defining the wage received by an individual as the sum of several elements, each element representing the part of income resulting from each individuals' attribute, the Shapley decomposition rule offers a simple way to determine the contribution of race as well as all other individuals' attributes to the overall income inequality.

We concentrate on the discriminational part of observed inequalities, that is inequalities associated to pure ethnical concerns. Observed income difference between blacks and whites are linked to racial discrimination (pure ethnical concern) but also, for instance, to differences in education levels of these two racial populations. The Shapley decomposition enables us do distinguish inequality arising from age, gender, education to the share of inequality induced by a racial affiliation. We find that the contribution of race to the total observed inequality is quite low - about 1% to 4% depending on the 9 United States Census Bureau designated divisions. Such a result indicates that the real inequality issue is more correlated to characteristics to which ethnic minorities are associated (low education, poverty etc...) rather than a matter of race in itself.

The paper is organized as follows: Section 2 presents some empirical evidence on racial income inequality in the United States over the period 2009-2014. Section 3 presents the Shapley decomposition methodology according to the conditional decomposition. Section 4 analyses the preliminary results. Finally, section 5 contains some concluding comments.

[2]Chantreuil and Trannoy (2011, 2013) defined the Shapley decomposition rules determining the contribution of different income sources or different sub-populations to the overall income inequality.

2 Racial inequality: some empirical evidence

We use the IPUMS-CPS given by the University of Minnesota from 1989 to 2015. This database allows for tracking individuals over time, each individual having a unique identifier. Earnings are given as weekly earnings[3]. We keep income declared during the 4th and the 8th month of interview and drop duplicates[4]. We keep only individuals declared as pure black or pure white[5]. Individuals are described according to gender: male or female; education level: less than high-school degree, high-school degree plus some college without a degree, undergraduate degree, and graduate degree or more; and age class: below 20 years old, one group every 4 years from 20-24 years old up to 60-64 years old. Individuals over 65 years old, military, out of the labor force, or with a weekly pay below one hour of the federal minimum wage for each year are dropped from the data set. Part-time workers are also dropped from the data set so that we concentrate on full-time workers[6]. Finally, US States are regrouped into 9 regional sub-sets according to the United States Census Bureau designated divisions. The total data set over 26 years from 1989 to 2015 is constituted of 2,140,007 observations among which 11.62% are defined as pure black.

Table 1 and 2 shows the distribution of black and white population per administrative division together with the average weekly income over time.

Table 1: Descriptive statistics on black/white distribution

Region	Data set in 2015	Black share* in 1990	2000	2010	2015
New England	6,054	2.75	4.51	4.24	4.82
Middle Atlantic	6,479	12.89	14.42	15.54	14.94
East North Central	8,835	10.07	12.47	11.12	9.59
West North Central	8,439	2.93	3.64	4.47	5.55
South Atlantic	14,800	21.30	24.34	25.15	23.45
East South Central	6,033	18.67	19.71	18.25	22.49
West South Central	7,979	13.53	15.66	19.93	18.40
Mountain	8,488	2.63	2.85	3.58	3.03
Pacific	6,773	6.65	8.40	6.72	7.12
TOTAL	73,880	11.02	12.53	12.38	13.01

$*$ in the black and white population, excluding other ethnical groups.

[3]The Merged Outgoing Rotation Groups (MORG) extract provided by the Bureau of Labor Statistics enables to reconstitute the hourly pay as the weekly earnings (variable coded 'earnwke') divided by the number of hours worked during the week (variable coded 'uhourse') but such a method can be applied only for a much shorter period of time. Besides the number of hours worked is not given that often.

[4]Individual observed twice: once in month 4 and once in month 8. In case of duplication, we keep only the occurrence of month 4.

[5]Note that for some years the CPS identifies most Hispanic as whites and some as black, we drop these individuals.

[6]It is worth noting that the proportion of part-time workers is around 15 to 17% on average and that the white population is more concerned.

Table 2: Descriptive statistics on weekly earnings

Region	Population	1990	2000	2010	2015
New England	White	1037.20	1093.88	1181.52	1140.72
	Black	797.70	809.31	852.96	814.31
Middle Atlantic	White	1024.45	1111.04	1163.24	1137.27
	Black	795.21	813.68	832.96	821.01
East North Central	White	937.30	1032.75	1006.38	1002.25
	Black	774.12	803.25	821.01	755.48
West North Central	White	853.57	953.72	958.88	968.99
	Black	694.41	796.99	752.28	705.86
South Atlantic	White	934.51	1016.97	1079.50	1055.97
	Black	684.54	754.23	809.27	783.97
East South Central	White	784.29	904.33	901.55	914.97
	Black	598.71	674.50	709.77	676.95
West South Central	White	928.59	995.92	1021.86	1039.67
	Black	624.71	742.86	738.28	774.94
Mountain	White	899.86	996.17	1042.70	1031.17
	Black	713.82	716.38	757.72	909.65
Pacific	White	1103.73	1151.90	1232.97	1201.81
	Black	901.14	920.91	973.89	925.34

$*$ in the black and white population, excluding other ethnical groups.

According to Table 1, the black population is more important in South Atlantic, followed by East and South Central and Middle Atlantic. Compared to the country average, Blacks are under-represented in New England, Mountain, West North Central, Pacific and East North Central. This distribution hides large disparities since administrative divisions cover many different states (see Appendix A for a list of states per division). Regarding real average weekly earnings (base 2015)[7], they differ: while Pacific offers high income for both ethical groups, East South Central offers very low incomes. Racial differences consistently disadvantage the black population. The larger difference is observed for New England with about $326 more for whites in 2015 and the smallest in Mountain with $122 of difference in 2015.

3 Income decomposition framework

The decomposition of the income inequality into appropriate component contributions falls usually into two main cases. The first one studies situations in which different components of total income are examined[8], while the second one considers the influence of

[7]Frequency weights have been applied.

[8]See Fei, Ranis and Kuo (1978), Shorrocks (1982), and Lerman and Yitzhaki (1985).

population subgroups[9]. For both types of decomposition, the Shapley value has been proved to be useful in many applications[10], nevertheless the use of the Shapley decomposition rule by population subgroups such as race or gender does not lead to a clear cut answer looking for the "real" contribution of such individuals' attributes. Chantreuil and Lebon (2015) solved this problem proposing a solution "*assimilating the different dimensions of the status of individuals to a particular wage source in order to assess the contribution of each status*". We call this proposed framework the income decomposition by attributes.

Let consider an income distribution $X = (x_1, x_2, ..., x_n)$ among a set of individuals $N = \{1, ..., i, ..., n\}$ and a set of attributes $A = \{1, ..., j, ..., a\}$ such as age, level of education or race. If the overall income inequality is measured by an inequality index I, such that the value of zero is assigned to an equal income distribution, the contribution of attribute j to the overall inequality $I(X)$ computed with the Shapley decomposition rule is defined by the following formula:

$$Sh_j(X, A, I) = \sum_{S \subseteq A, j \in S} \frac{(s-1)!(a-s)!}{a!} \Big[I\big(Y(S)\big) - I\big(Y(S - \{j\})\big) \Big] \tag{1}$$

with s the cardinality of S, a the cardinality of A and $S \in 2^A$ a subset of the set of attributes A.

By convention, for $S = \emptyset$, $Y(S) = 0$ and for all $S \in 2^A$, $S \neq \emptyset$, $Y(S)$ is defined as follows:

$$Y(S) = \left(\sum_{j \in S} y_1^j + \sum_{\substack{h \notin S \\ h \in A}} \frac{\sum_{i=1}^n y_i^h}{n}, ..., \sum_{j \in S} y_i^j + \sum_{\substack{h \notin S \\ h \in A}} \frac{\sum_{i=1}^n y_i^h}{n}, ..., \sum_{j \in S} y_n^j + \sum_{\substack{h \notin S \\ h \in A}} \frac{\sum_{i=1}^n y_i^h}{n} \right) \tag{2}$$

$Y(S)$ is thus the income distribution obtained from the income distribution $Y(A)$ when the shares of income related to the attributes $h \notin S$ are equally distributed among individuals.

The distribution $Y(A)$ according to the set of attributes A can be derived from the distribution of income X using the conditional decomposition by attributes. Such a method is a generalization of the framework proposed by Chantreuil and Lebon (2015) to the case within which more than two attributes have to be accounted for.

In order to present formally the two approaches, let us introduce the following notations. We consider that each attribute $j \in A$ has $m(j)$ modalities, such that $1 \leq k_j \leq m(j)$, where k_j is the k-th modality of the attribute j.

[9] See Bourguignon (1979), Cowell (1980), Shorrocks (1980, 1984), and Foster and Shorrocks (1988).

[10] A non-exhaustive list of applications of the Shapley value to inequality decomposition includes Sastre and Trannoy (2002), Israeli (2007), Bargain and Callan (2010), Devicienti (2010) or Charpentier and Mussard (2011).

The number of individuals who have the k-th modality of the attribute j is denoted by n_{k_j} such that

$$\sum_{k_j=1}^{m(j)} n_{k_j} = n \tag{3}$$

The number of individuals who have the k-th modality of the attribute 1 and the k-th modality of the attribute 2 is denoted by n_{k_1,k_2} such that

$$\sum_{k_2=1}^{m(2)} n_{k_1,k_2} = n_{k_1} \tag{4}$$

The number of individuals who have the k-th modality of all attribute 1 to a is denoted by $n_{k_1,...,k_j,...,k_a}$ such that

$$\sum_{k_a=1}^{m(a)} n_{k_1,...,k_j,...,k_a} = n_{k_1,...,k_j,...,k_{a-1}} \tag{5}$$

The income of individuals $i \in N$ who has the k-th modality of the attribute j is denoted by $x_i^{k_j}$.

The income of individuals $i \in N$ who has the k-th modality of the attribute 1 and the k-th modality of attribute 2 is denoted by $x_i^{k_1,k_2}$.

The income of an individual $i \in N$ who has the k-th modality of all attributes 1 to a is denoted $x_i^{k_1,...,k_j,...,k_a}$.

The income distribution $Y(A)$ is based on the assumption that the set of attributes is ranked by order of importance from 1 to a. Given this ranking of the individuals' attributes, the share of income of an individual i coming from the attribute 1 is defined as the average income of individuals who have the same attribute 1's modality and the share of income of an individual coming from attribute j is defined as the average income of individuals who have the same sequence of modalities for all attributes from 1 to j.

Thus the distribution $Y(A) = (y_1, ..., y_i, ..., y_n)$ is such that for all $i \in N$

$$y_i = \sum_{j=1}^{a} y_i^j + \left[y_i - \sum_{j=1}^{a} y_i^j \right] \tag{6}$$

where

$$y_i^1 = \left[\frac{\sum_{k_1=1}^{m(1)} x_i^{k_1}}{n_{k_1}} \right] \tag{7}$$

$$y_i^2 = \left[\frac{\sum_{k_2=1}^{m(2)} x_i^{k_1,k_2}}{n_{k_1,k_2}} - \frac{\sum_{k_1=1}^{m(1)} x_i^{k_1}}{n_{k_1}} \right] \tag{8}$$

$$y_i^j = \left[\frac{\sum_{k_j=1}^{m(j)} x_i^{k_1,...,k_j}}{n_{k_1,...,k_j}} - \frac{\sum_{k_{j-1}=1}^{m(j-1)} x_i^{k_1,...,k_{j-1}}}{n_{k_1,...,k_{j-1}}} \right] \tag{9}$$

and

$$y_i^a = \left[\frac{\sum_{k_a=1}^{m(a)} x_i^{k_1,\ldots,k_a}}{n_{k_1,\ldots,k_a}} - \frac{\sum_{k_{a-1}=1}^{m(a-1)} x_i^{k_1,\ldots,k_{a-1}}}{n_{k_1,\ldots,k_{a-1}}} \right] \tag{10}$$

From the previous equations, for all $i \in N$ we thus have

$$\begin{aligned} y_i \;=\; & \left[\frac{\sum_{k_1=1}^{m(1)} x_i^{k_1}}{n_{k_1}} \right] + \ldots + \left[\frac{\sum_{k_j=1}^{m(j)} x_i^{k_1,\ldots,k_j}}{n_{k_1,\ldots,k_j}} - \frac{\sum_{k_{j-1}=1}^{m(j-1)} x_i^{k_1,\ldots,k_{j-1}}}{n_{k_1,\ldots,k_{j-1}}} \right] \\ & + \ldots + \left[\frac{\sum_{k_a=1}^{m(a)} x_i^{k_1,\ldots,k_a}}{n_{k_1,\ldots,k_a}} - \frac{\sum_{k_{a-1}=1}^{m(a-1)} x_i^{k_1,\ldots,k_{a-1}}}{n_{k_1,\ldots,k_{a-1}}} \right] + \left[y_i - \frac{\sum_{k_a=1}^{m(a)} x_i^{k_1,\ldots,k_a}}{n_{k_1,\ldots,k_a}} \right] \end{aligned} \tag{11}$$

4 Preliminary results

Observations have been grouped by 3 years so that the number of blacks in each geographic division is sufficiently high.

The relative contribution of each attributes in percentage of the total contribution according to the conditional decomposition is given in Tables 3 to 6. Figures 1 and 2 show the relative contribution of race to total observed inequality on the administrative division map of the USA for this conditional decomposition for years 2013 to 2015 and years 1989 to 1991.

Table 3: Weekly earnings inequality - Gini index and its conditional decomposition

	Gini	AGE	EDUC	GENDER	RACE	RESIDUAL
Federal 1991-90-89	0.3157	13.22	18.78	11.91	2.43	53.66
Federal 1994-93-92	0.3217	13.98	19.20	10.43	2.25	54.14
Federal 1997-96-95	0.3248	13.28	19.23	10.58	2.47	54.44
Federal 2000-99-98	0.3357	12.43	19.77	10.58	2.34	54.89
Federal 2003-02-01	0.3357	11.27	19.51	10.04	2.37	56.81
Federal 2006-05-04	0.3377	11.85	19.29	9.53	2.58	56.74
Federal 2009-08-07	0.3380	11.37	19.33	9.19	2.68	57.43
Federal 2012-11-10	0.3404	11.24	19.73	8.63	2.35	58.05
Federal 2015-14-13	0.3402	11.26	18.60	8.29	2.61	59.25

Table 4: Weekly earnings inequality - Gini index and its conditional decomposition

	Gini	AGE	EDUC	GENDER	RACE	RESIDUAL
New England 1991-90-89	0.2952	13.66	20.20	12.12	1.10	52.92
New England 1994-93-92	0.3032	13.46	20.36	10.71	1.33	54.14
New England 1997-96-95	0.3122	13.35	21.82	10.18	1.31	53.33
New England 2000-99-98	0.3314	11.61	22.54	10.51	1.57	53.78
New England 2003-02-01	0.3290	10.67	21.23	10.41	1.50	56.20
New England 2006-05-04	0.3328	12.52	20.70	10.00	1.50	55.28
New England 2009-08-07	0.3322	12.79	20.41	9.52	1.73	55.55
New England 2012-11-10	0.3330	12.15	20.62	9.14	1.81	56.28
New England 2015-14-13	0.3325	12.15	19.90	8.41	1.80	57.74
Middle Atlantic 1991-90-89	0.3083	12.13	19.94	11.52	2.18	54.24
Middle Atlantic 1994-93-92	0.3202	12.71	21.14	9.88	2.10	54.18
Middle Atlantic 1997-96-95	0.3262	12.16	21.74	9.76	2.39	53.94
Middle Atlantic 2000-99-98	0.3395	11.26	21.81	9.98	2.52	54.43
Middle Atlantic 2003-02-01	0.3432	10.06	21.84	9.78	2.42	55.90
Middle Atlantic 2006-05-04	0.3398	10.52	21.59	8.37	2.98	56.54
Middle Atlantic 2009-08-07	0.3384	9.94	20.36	9.02	3.30	57.38
Middle Atlantic 2012-11-10	0.3392	10.64	21.06	7.86	2.85	57.59
Middle Atlantic 2015-14-13	0.3379	9.84	19.97	8.11	2.85	59.23
East North Central 1991-90-89	0.3064	14.75	16.78	13.85	1.70	52.93
East North Central 1994-93-92	0.3130	14.84	18.16	12.09	1.53	53.37
East North Central 1997-96-95	0.3170	13.45	18.06	12.64	1.84	54.01
East North Central 2000-99-98	0.3246	12.82	18.99	12.17	1.88	54.14
East North Central 2003-02-01	0.3237	11.28	19.10	10.85	1.91	56.86
East North Central 2006-05-04	0.3256	13.01	18.79	10.91	2.12	55.18
East North Central 2009-08-07	0.3277	11.76	19.27	9.95	2.04	56.99
East North Central 2012-11-10	0.3318	12.63	19.51	8.85	1.72	57.28
East North Central 2015-14-13	0.3314	11.58	18.91	8.85	2.16	58.49

Table 5: Weekly earnings inequality - Gini index and its conditional decomposition

	Gini	AGE	EDUC	GENDER	RACE	RESIDUAL
West North Central 1991-90-89	0.3055	14.63	16.15	12.96	1.21	55.05
West North Central 1994-93-92	0.3096	16.17	15.18	12.41	1.16	55.08
West North Central 1997-96-95	0.3077	14.59	16.33	11.26	1.16	56.67
West North Central 2000-99-98	0.3150	13.07	17.25	11.56	1.37	56.74
West North Central 2003-02-01	0.3190	12.24	16.67	11.14	1.14	58.81
West North Central 2006-05-04	0.3228	13.38	16.21	10.27	1.43	58.70
West North Central 2009-08-07	0.3236	13.05	17.19	10.14	1.34	58.28
West North Central 2012-11-10	0.3233	13.53	17.27	9.37	1.31	58.51
West North Central 2015-14-13	0.3242	12.62	15.99	9.50	1.61	60.28
South Atlantic 1991-90-89	0.3142	11.75	20.90	11.01	4.12	52.22
South Atlantic 1994-93-92	0.3215	12.92	21.05	9.65	3.65	52.74
South Atlantic 1997-96-95	0.3263	12.56	21.26	9.58	3.86	52.74
South Atlantic 2000-99-98	0.3329	11.74	20.86	9.99	3.69	53.72
South Atlantic 2003-02-01	0.3350	10.63	20.61	9.38	3.70	55.69
South Atlantic 2006-05-04	0.3391	10.91	20.57	9.06	4.05	55.40
South Atlantic 2009-08-07	0.3380	10.96	20.40	8.18	3.94	56.52
South Atlantic 2012-11-10	0.3436	10.56	20.67	8.35	3.46	56.95
South Atlantic 2015-14-13	0.3472	10.84	20.21	6.96	3.75	58.25
East South Central 1991-90-89	0.3147	12.27	17.27	13.49	4.27	52.71
East South Central 1994-93-92	0.3222	13.17	17.52	12.65	4.55	52.12
East South Central 1997-96-95	0.3212	12.60	17.58	12.96	4.54	52.32
East South Central 2000-99-98	0.3308	11.87	18.59	12.67	3.59	53.29
East South Central 2003-02-01	0.3256	10.90	18.24	12.32	4.06	54.48
East South Central 2006-05-04	0.3283	10.83	18.04	11.55	3.54	56.04
East South Central 2009-08-07	0.3296	9.94	18.76	11.15	3.81	56.34
East South Central 2012-11-10	0.3306	10.09	17.27	10.28	3.49	58.87
East South Central 2015-14-13	0.3288	11.22	16.08	9.85	3.64	59.21

Table 6: Weekly earnings inequality - Gini index and its conditional decomposition

	Gini	AGE	EDUC	GENDER	RACE	RESIDUAL
West South Central 1991-90-89	0.3314	13.14	18.88	12.32	3.76	51.89
West South Central 1994-93-92	0.3299	13.54	18.44	11.51	3.78	52.73
West South Central 1997-96-95	0.3305	13.20	17.86	11.88	3.85	53.21
West South Central 2000-99-98	0.3481	11.57	19.42	11.74	3.33	53.93
West South Central 2003-02-01	0.3407	10.87	18.51	11.17	3.63	55.83
West South Central 2006-05-04	0.3442	11.60	17.65	10.85	3.35	56.55
West South Central 2009-08-07	0.3491	11.37	17.80	11.22	4.06	55.55
West South Central 2012-11-10	0.3441	10.11	17.97	10.22	3.58	58.12
West South Central 2015-14-13	0.3415	11.08	15.65	10.46	4.06	58.76
Mountain 1991-90-89	0.3146	16.16	15.16	12.07	1.03	55.58
Mountain 1994-93-92	0.3189	16.20	15.99	10.51	1.26	56.05
Mountain 1997-96-95	0.3195	15.64	15.51	11.07	0.92	56.86
Mountain 2000-99-98	0.3281	15.13	15.92	10.65	1.23	57.07
Mountain 2003-02-01	0.3285	14.60	16.57	9.85	1.01	57.96
Mountain 2006-05-04	0.3327	14.95	15.52	10.10	1.21	58.22
Mountain 2009-08-07	0.3316	13.52	15.00	9.48	1.70	60.30
Mountain 2012-11-10	0.3363	13.33	16.41	8.98	1.33	59.95
Mountain 2015-14-13	0.3327	13.34	15.30	8.99	1.46	60.92
Pacific 1991-90-89	0.3114	14.75	16.54	11.52	1.48	55.71
Pacific 1994-93-92	0.3122	15.45	17.40	8.75	1.43	56.97
Pacific 1997-96-95	0.3192	14.61	16.79	9.23	1.60	57.78
Pacific 2000-99-98	0.3394	14.33	18.10	8.95	2.11	56.50
Pacific 2003-02-01	0.3388	12.92	17.09	9.18	2.05	58.75
Pacific 2006-05-04	0.3389	12.85	17.83	8.06	2.22	59.05
Pacific 2009-08-07	0.3332	13.17	18.08	8.03	2.20	58.52
Pacific 2012-11-10	0.3376	12.61	19.40	7.76	1.78	58.44
Pacific 2015-14-13	0.3398	13.13	18.12	7.35	1.91	59.49

In the conditional decomposition applied here, the racial attribute has to be understood as the source of income inequality which remains after having considered age, education and gender. The racial contribution to total observed income inequality is quite low: the figure vary between about 0.92% and 4.55% only while about 40% of the inequality can be explained by age, gender and education taken altogether. The residual attribute - that is unconsidered variables - is large, representing on average 55% of the observed inequality. Such a result shows that pure racial discrimination on incomes exists but observed inequality between Blacks and Whites are mostly correlated non-racial characteristics. Note that this pure and direct racial discrimination in terms of income has to be taken with cautious since another indirect discriminative aspect of income differences is possibly correlated to Blacks obtaining for instance less prestigious jobs. Such

correlations are not taken into account here. Meanwhile, education is the attribute which captures the most of the observed income inequality (between 19.43% and 25.02%) when it comes to comparing ethnies.

In line with the sociological literature showing that black-white inequality in local area is greater where the black population is relatively large, such a statement seems to hold at the division level. Apart from one exception, divisions with the highest racial inequality are divisions with the highest black-white ratio. Comparing Tables 1 to tables 3 to 6 shows that divisions with the highest inequality associated to the racial attribute in 2013 to 2015 - that is in this order or importance, West South Central (4.06%), South Atlantic (3.75%), East South Central (3.64%), Middle Atlantic (2.85%) and East North Central (2.16%) - are also divisions with the highest black-white ratio, while divisions with a small black ratio present a small percentage of inequality associated to the racial attribute. Mountain (1.46%), West North Central (1.61%), New England (1.80%) and Pacific (1.91%) indeed present a black share lower than 7%. The only notable exception is West South Central which shows the higher inequality associated to the racial attribute but is not the division with the higher black proportion (13.53% against 21.30% for South Atlantic). It remains that apart from this exception, the highest the share of blacks in the black-white population, the highest the inequality associated to the racial attribute.

Comparing Gini figures over time show that inequalities tend to rise for all divisions between 1989 and 2015. Part of this rise is captured though a small increase in inequality associated to the racial attribute. Such a rise in the Gini index mostly "benefit" to unobserved characteristics (from 53.66% to 59.25% at the federal level). Education equalities remain stable while inequalities associated to age and gender tend to decrease.

Figure 1: Racial contribution to inequality per division in the US, conditional decomposition

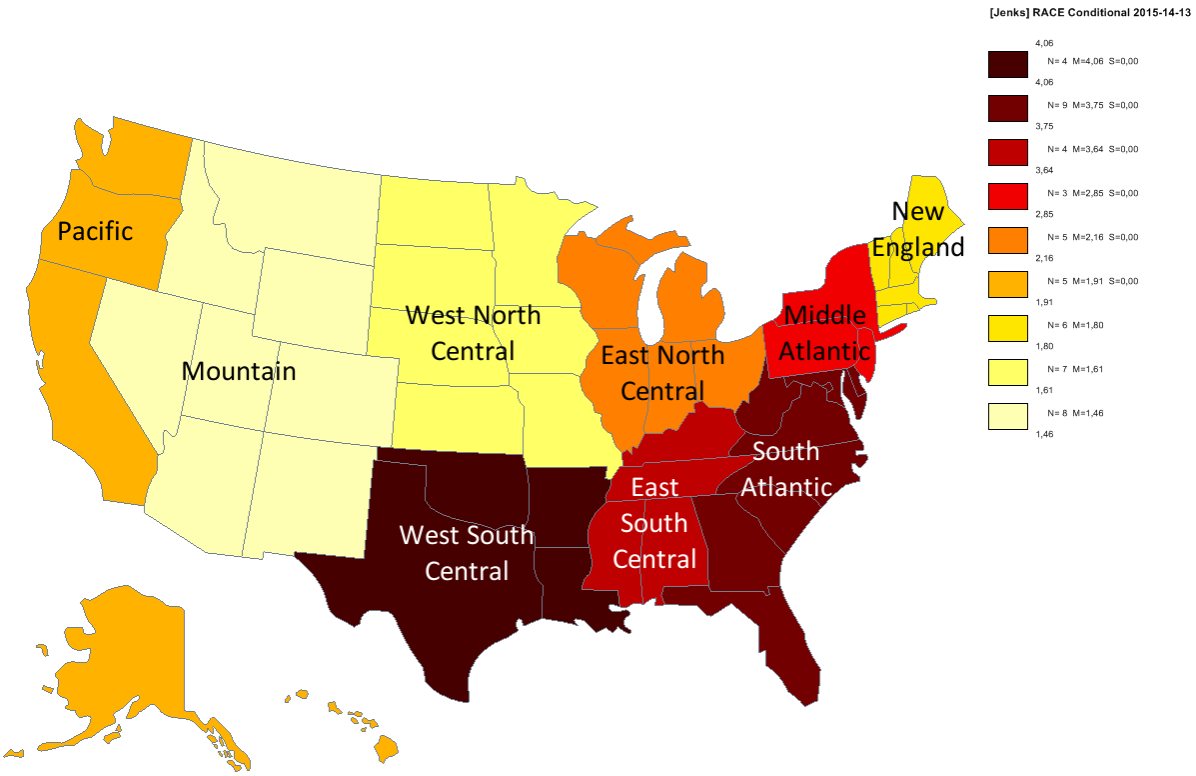
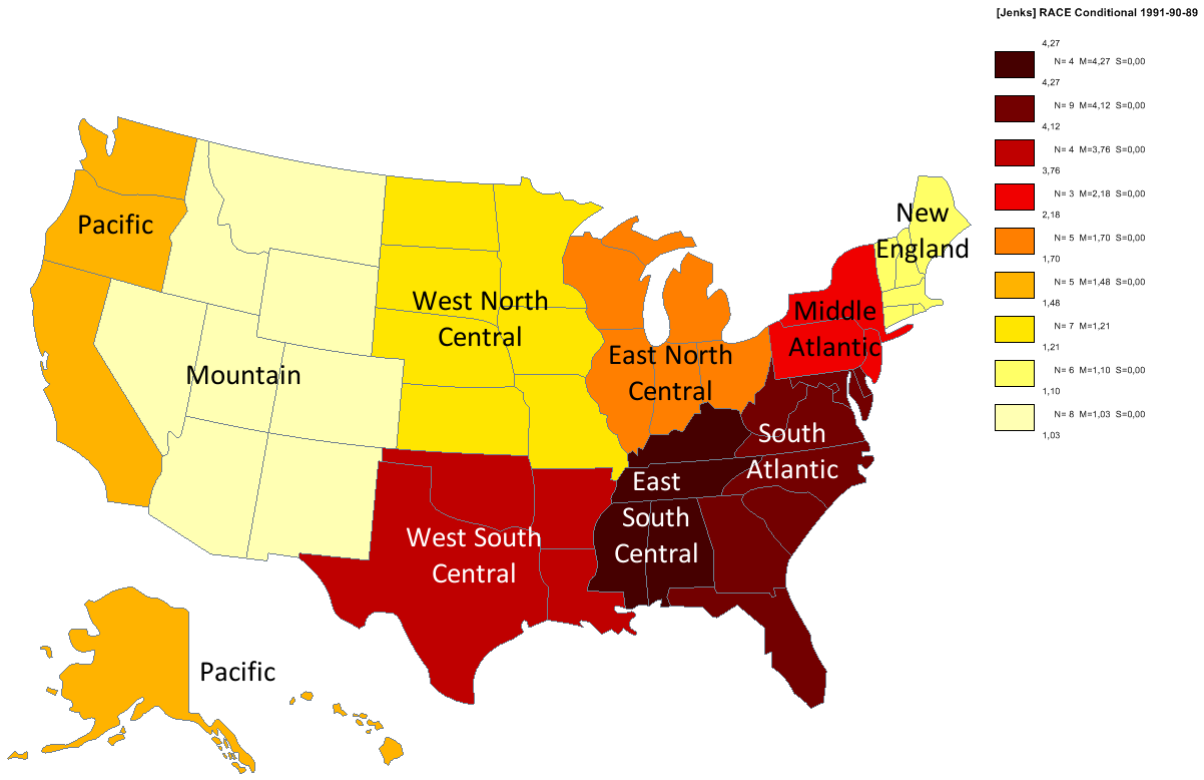


Figure 2: Racial contribution to inequality per division in the US, conditional decomposition



Looking at Figures 1 and 2 shows that racial inequalities are concentrated around the south west divisions but also appears on the west coast. This geographical characteristic does not seems to evolve much between 1989 and 2015.

5 Conclusion

The income inequality decomposition *à la* Shapley (1953) enables us to derive the contribution of an individual characteristic to the total observed income inequality in the US over the period 2009-2014. We concentrate on the discriminational part of observed inequalities, that is inequalities associated to pure ethnical concerns. Observed income difference between blacks and whites are linked to racial discrimination (pure ethnical concern) but also, for instance, to differences in education levels of these two racial populations. The Shapley decomposition enables us do distinguish inequality arising from age, gender, education to the share of inequality induced by a racial affiliation. We find that the racial contribution to income inequality is about 1 to 4% for all of the 9 United States Census Bureau designated divisions. Education accounts for between 20-24% which represent the highest source of income inequality after unobserved characteristics. Such a result indicates that the real inequality issue is more correlated to characteristics to which ethnic minorities are associated (low education, poverty etc...) rather than a matter of race in itself. In terms of public policy, such a result question affirmative actions in public

institutions in general, but would promote them in education. Policies aiming at promoting better access to education for ethnic minorities would indeed impact the observed income inequality associated to education.

References

Bargain, O. and Callan, T. (2010), Analysing the effects of tax-benefit reforms on income distribution: a decomposition approach, *Journal of Economic Inequality* 8, pp.1-21.

Bourguignon, F. (1979), Decomposable income inequality measures. *Econometrica* 47, pp.901-920.

Chantreuil, F., Lebon, I. (2015), Gender contribution to income inequality. *Economics Letters* 133, pp.27-32.

Chantreuil, F., Trannoy, A. (2013), Inequality Decomposition Values: the Trade-off between Marginality and Efficiency. *Journal of Economic Inequality* 11(1), pp.83-98.

Chantreuil, F., Trannoy, A. (2011), Inequality Decomposition Values. *Annales d'Economie et Statistique* 101-102, pp.13-36.

Charpentier, A. and Mussard, S. (2011), Income inequality games. *Journal of Economic Inequality* 9(4), pp.529-554.

Couch, K. and Daly, M.C. (2002), Black-White Wage Inequality in the 1990s: a Decade of Progress, *Economic Inquiry* 40(1), pp.31-41.

Cowell, F.A (1980), On the structure of additive inequality measures. *Review of Economic Studies* 47, pp.520-531.

Devicienti, F. (2010), Shapley-value decompositions of changes in wage distributions: a note, Journal of Economic Inequality 8, pp.35-45.

Economic Report of the President (2016), Transmitted to the Congress on February 2016 together with THE ANNUAL REPORT of the COUNCIL OF ECONOMIC ADVISERS, United States Government Printing Office, Washington.

Fei, J., Ranis, G. and Kuo, S. (1978), Growth and the Family Distribution of Income by Factor Components, *Quarterly Journal of Economics* 92 (1), pp.17-53.

Foster, J. and Shorrocks, A. (1988), Poverty orderings, Econometrica 56(1), pp.173-177.

Huffman, M. and Cohen, P. (2004), Racial Wage Inequality: Job Segregation and Devaluation across U.S. Labor Markets, *American Journal of Sociology* 109(4), pp.902-936.

Israeli, O. (2007), A Shapley-based decomposition of the R-Square of a linear regression, Journal of Economic Inequality 5, pp.199-212.

Juhn, C., K. Murphy, and B. Pierce (1991), Accounting for the Slowdown in Black-White Wage Convergence. In *Workers and Their Wages: Changing Patterns in the United States*, pp.107-143.

Lerman, R. and Yitzhaki, S. (1985), Income Inequality Effects by Income, *Review of Economics and Statistics* 67(1), pp.151-156.

Sastre, M. and Trannoy, A. (2002), Shapley inequality decomposition by factor components: Some methodological issues, *Journal of Economics* 77(1), pp.51-89.

Shapley, L.S. (1953), A value for n-person game, in *Contributions to the Theory of Games*. Vol. 2, Annals of Mathematics Studies, eds. H.W. Kuhn and A.W. Tucker. Princeton University Press, pp.307-317.

Shorrocks, A. (1980), The Class of Additively Decomposable Inequality Measure. *Econometrica* 48(3), pp.613-625

Shorrocks, A. (1982), Inequality Decomposition by Factor Components. *Econometrica* 50(1), pp.193-211.

Shorrocks, A. (1984), Inequality decomposition by population subgroups. *Econometrica* 51, pp.1369-1385.

Shorrocks, A. (2013), Decomposition procedures for distributional analysis: a unified framework based on the Shapley value. *Journal of Economic Inequality* 11(1), pp.99-126.

Sites, W. and Parks, V. (2011), What Do We Really Know About Racial Inequality? Labor Markets, Politics, and the Historical Basis of Black Economic Fortunes, *Politics and Society* 39(1), pp.40-73.

A United States Census Bureau designated divisions

1. New England
- (ME) Maine
- (NH) New Hampshire
- (VT) Vermont
- (MA) Massachusetts
- (RI) Rhode Island
- (CT) Connecticut

2. Middle Atlantic
- (NY) New York
- (NJ) New Jersey
- (PA) Pennsylvania

3. East North Central
- (OH) Ohio
- (IN) Indiana
- (IL) Illinois
- (MI) Michigan
- (WI) Wisconsin

4. West North Central
- (MN) Minnesota
- (IA) Iowa
- (MO) Missouri
- (ND) North Dakota
- (SD) South Dakota
- (NE) Nebraska
- (KS) Kansas

5. South Atlantic
- (DE) Delaware
- (MD) Maryland
- (DC) District of Columbia
- (VA) Virginia
- (WV) West Virginia
- (NC) North Carolina
- (SC) South Carolina
- (GA) Georgia
- (FL) Florida

6. East South Central
- (KY) Kentucky
- (TN) Tennessee
- (AL) Alabama
- (MS) Mississippi

7. West South Central
- (AR) Arkansas
- (LA) Louisiana
- (OK) Oklahoma
- (TX) Texas

8. Mountain
- (MT) Montana
- (ID) Idaho
- (WY) Wyoming
- (CO) Colorado
- (NM) New Mexico
- (AZ) Arizona
- (UT) Utah
- (NV) Nevada

9. Pacific
- (WA) Washington
- (OR) Oregon
- (CA) California
- (AK) Alaska
- (HI) Hawaii