# Race as a Multistate Process

Jerônimo O. Muniz (UFMG)
Aliya Saperstein (Stanford University)
Bernardo L. Queiroz (UFMG)

**Abstract.** Although the existence of racial fluidity in Brazil is generally accepted, changes in racial classification over the life course are not often incorporated into standard demographic estimates. We calculate multistate life tables by drawing on linked data from the largest Brazilian household survey (2017-2019 PNAD-C) to estimate transition probabilities between the White, Brown and Black race categories, which we combine with age and race-specific probabilities of dying. This allows us to ask and answer a series of unconventional questions: How long could one expect to live in a given race? At what ages are transitions between races more likely to occur? How many years can someone born as White, Brown, or Black expect to live in another race? Most surprisingly, our conditional life expectancy estimates show that Brazilians who were born as Black could expect to live almost 15 years of their lives as White, while those who were born White could expect to live as Black for 3 years, on average. The results both provide important new evidence on the scope of racial fluidity in contemporary Brazil and demonstrate the feasibility of accounting for that fluidity in traditional demographic analysis.

## Introduction

To state that race is a social, and thus a variable characteristic is now becoming a truism, a cliché. The mutable nature of race, or the fluidity of racial classifications, has been well documented in both the United States (Dahis, Nix and Qian 2019, Liebler et al. 2017, Saperstein and Gullickson 2013, Saperstein and Penner 2012) and in Latin America (Davenport 2020), where Brazil has been highlighted as the most paradigmatic case due to its high racial fluidity (Carvalho et al 2004, Cornwell et al 2017, Miranda 2014, Muniz and Telles 2021, Senkevics 2017, Silveira 2019).

Here, we approach the variability of race over the life course from a multistate perspective. Multistate life tables allow us to treat race as a non-permanent state, in which individuals may only spend part of their lives. The added virtue of multistate life tables lies in its ability to answer questions that a conventional life table is not able to answer, such as: What is the number of years that a cohort could expect to live in a specific state (race category)? What is the probability of dying in a given state? What proportion of life will be spent in a given racial category? How many years will be lived as White, Brown or Black by those who were born in one of these same states? To our knowledge, this is the first approach to use increment-decrement life tables to describe race as a multistate process over the life cycle.

**Data**

We use linked data from the Continuous National Household Sample Survey (PNAD-C) fielded between 2017 and 2019 to estimate probabilities of transition between self-reported White, Brown and Black Brazilians. Each quarter, about 211,000 households are interviewed, covering approximately 16,000 census sectors of 3,500 municipalities. Selected households are interviewed for five consecutive periods; the households are visited every three months. Unique identifiers allow the linkage of households over consecutive surveys, and the residents' sex, date of birth and years of education allow the identification and linkage of the same person in different waves of the survey (Miranda 2014; Queiroz 2007; Ribas and Soares 2009). Once individuals are linked, a survey of panel data is generated.

We use the algorithm proposed by Ribas and Soares (2009) to link individuals because we cannot match households in the PNAD despite knowing the household identification number and the month of the survey. The matching process using these two unique variables is not perfect because the unit of analysis is the household and not the person. Therefore, if all members of the household move out between the period of the survey, and new members move to that household, these individuals will be in the later survey with the same household identification number of the previous residents. Thus, matching households using household identification number and month of the survey would create false matches in some cases.

We also combine death registries, from the Information System on Mortality from the Ministry of Health (SIM, Datasus) in 2005, and population counts from 2010 and 2020 Brazilian Censuses to estimate race-specific life tables. The probabilities of racial transition (from PNAD-C) and mortality (from Censuses and SIM) are then combined into a single matrix to estimate multistate life tables in which each state is represented by one race (White, Brown or Black) or one absorbing state (death).

**Methods and preliminary results**

Our methodological strategy involves three procedures. In the first one we calculate transition probabilities by age groups between self-reported Whites, Browns and Blacks interviewed two to five times between 2017 to 2019. In the second methodological step we estimate race-specific life tables by age, in order to account for mortality, the only absorbing state in our multistate framework. Finally, in the third step we combine probabilities of racial transition and race-specific mortality to generate the input matrix of our multistate process, which allows us to investigate the amount of time that synthetic birth cohorts would spend in the White, Brown or Black racial state over their life cycles (in addition to other relevant measures). In what follows, we provide further details and some preliminary results.

***Transition probabilities between races.*** Transition probabilities correspond to the relative row frequencies of contingencies tables between the previous and the current reported race. Transition probabilities are thus drawn from all person-years who have self-reported their race over a period of three years. Table 1 shows these transition probabilities derived from 2017-2019 Continuous PNAD data:

**Table 1.** Probabilities of mobility between Whites, Browns and Blacks

| Age group | W--> W | W-> Br | W--> B | Br--> W | Br--> Br | Br--> B | B--> W | B--> Br | B--> B |
|---|---|---|---|---|---|---|---|---|---|
| 0-4 | 0.8647 | 0.1308 | 0.0045 | 0.0903 | 0.8828 | 0.0269 | 0.0342 | 0.2663 | 0.6994 |
| 5-9 | 0.8554 | 0.1394 | 0.0053 | 0.0801 | 0.8897 | 0.0302 | 0.0277 | 0.2861 | 0.6862 |
| 10-14 | 0.8517 | 0.1428 | 0.0056 | 0.0766 | 0.8900 | 0.0334 | 0.0271 | 0.2871 | 0.6858 |
| 15-19 | 0.8528 | 0.1415 | 0.0057 | 0.0778 | 0.8847 | 0.0375 | 0.0242 | 0.2830 | 0.6928 |
| 20-24 | 0.8663 | 0.1275 | 0.0062 | 0.0780 | 0.8832 | 0.0389 | 0.0251 | 0.2580 | 0.7169 |
| 25-29 | 0.8799 | 0.1141 | 0.0060 | 0.0770 | 0.8847 | 0.0383 | 0.0269 | 0.2407 | 0.7324 |
| 30-34 | 0.8816 | 0.1119 | 0.0066 | 0.0745 | 0.8862 | 0.0393 | 0.0272 | 0.2345 | 0.7382 |
| 35-39 | 0.8846 | 0.1091 | 0.0063 | 0.0773 | 0.8823 | 0.0405 | 0.0267 | 0.2395 | 0.7338 |
| 40-44 | 0.8848 | 0.1088 | 0.0064 | 0.0770 | 0.8817 | 0.0413 | 0.0257 | 0.2327 | 0.7416 |
| 45-49 | 0.8881 | 0.1056 | 0.0063 | 0.0802 | 0.8766 | 0.0432 | 0.0259 | 0.2387 | 0.7354 |
| 50-54 | 0.8975 | 0.0967 | 0.0057 | 0.0805 | 0.8756 | 0.0439 | 0.0277 | 0.2359 | 0.7364 |
| 55-59 | 0.9011 | 0.0935 | 0.0054 | 0.0820 | 0.8744 | 0.0435 | 0.0264 | 0.2413 | 0.7322 |
| 60-64 | 0.9042 | 0.0908 | 0.0050 | 0.0846 | 0.8724 | 0.0430 | 0.0261 | 0.2316 | 0.7423 |
| 65-69 | 0.9021 | 0.0929 | 0.0049 | 0.0876 | 0.8690 | 0.0434 | 0.0258 | 0.2261 | 0.7481 |
| 70-74 | 0.9086 | 0.0865 | 0.0049 | 0.0918 | 0.8626 | 0.0455 | 0.0290 | 0.2414 | 0.7296 |
| 75-79 | 0.8996 | 0.0957 | 0.0047 | 0.1003 | 0.8524 | 0.0472 | 0.0310 | 0.2429 | 0.7261 |
| 80+ | 0.9042 | 0.0901 | 0.0058 | 0.1084 | 0.8459 | 0.0457 | 0.0362 | 0.2516 | 0.7122 |
| Observations | | 1,053,476 | | | 1,395,308 | | | 227,139 | |

Note: W= Whites; Br= Browns; B= Blacks

Source: Linked data from the 2017-2019 Continuos PNAD

These transition probabilities yield several notable patterns. Overall, the least stability is observed for identification as Black, followed by Brown and White. However, Black and White identification stability is higher among older Brazilians, while stability in identification as Brown is highest among younger Brazilians. Thus, the probability of a transition from White to Brown appears to decrease with age, while the probability of a transition from Brown to Black increases with age. The most frequent racial transition is from Black identification to Brown, with a probability of about 25% that varies little by age. The least frequent transition is from White to Black, with probabilities on the order of 0.5%, though somewhat higher among working-age Brazilians (reaching 0.7% among respondents in their early 30s). In contrast, transitions to White from either Brown or Black are lowest among working-age Brazilians.

***Race-specific mortality rates.*** Brazil's vital registration system has several limitations, including under registration of deaths, age misreporting and a considerable amount of missing information on the decedent's race, so we use several approaches to estimate race-specific mortality. The first approach, proposed by Merli (1998), triangulates between the methods suggested by Preston and Bennett (1983), Bennett and Horiuchi (1984) – and later improved by Preston et al (1996) – to overcome the bias implicit in the growth rates associated to age distortions, differential coverage and residual intercensal migration present in these earlier procedures and to reconcile their results. These procedures estimate mortality from a set of registered deaths by age and age-specific growth rates derived from two census distributions. We will produce alternative estimates using 2010 census data, that included a question on deaths in the household in the past 12 months, and direct estimates from Ministry of Health Database.

**Table 2.** Race-specific life tables derived from Merli (1998)'s death process after iterations to correct age-specific growth rates

| | Whites | | | | Browns | | | | Blacks | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Age group | dx | lx | qx | ex | dx | lx | qx | ex | dx | lx | qx | ex |
| 0-4 | 30,235 | 1,557,167 | 0.0194 | 72.33 | 27,066 | 908,889 | 0.0298 | 68.22 | 2,134 | 105,836 | 0.0202 | 65.28 |
| 5-9 | 2,284 | 1,526,932 | 0.0015 | 68.76 | 2,154 | 881,824 | 0.0024 | 65.31 | 294 | 103,701 | 0.0028 | 61.62 |
| 10-14 | 2,735 | 1,524,648 | 0.0018 | 63.86 | 2,563 | 879,669 | 0.0029 | 60.46 | 341 | 103,408 | 0.0033 | 56.79 |
| 15-19 | 7,994 | 1,521,913 | 0.0053 | 58.96 | 8,497 | 877,106 | 0.0097 | 55.61 | 1,109 | 103,066 | 0.0108 | 51.95 |
| 20-24 | 10,882 | 1,513,919 | 0.0072 | 54.26 | 11,902 | 868,609 | 0.0137 | 51.12 | 1,573 | 101,957 | 0.0154 | 47.47 |
| 25-29 | 11,537 | 1,503,037 | 0.0077 | 49.63 | 12,971 | 856,707 | 0.0151 | 46.80 | 1,791 | 100,384 | 0.0178 | 43.17 |
| 30-34 | 13,492 | 1,491,500 | 0.0090 | 44.99 | 14,220 | 843,736 | 0.0169 | 42.47 | 2,046 | 98,593 | 0.0207 | 38.91 |
| 35-39 | 18,360 | 1,478,007 | 0.0124 | 40.37 | 17,153 | 829,516 | 0.0207 | 38.15 | 2,690 | 96,547 | 0.0279 | 34.66 |
| 40-44 | 25,098 | 1,459,647 | 0.0172 | 35.84 | 22,528 | 812,362 | 0.0277 | 33.89 | 3,608 | 93,858 | 0.0384 | 30.56 |
| 45-49 | 36,987 | 1,434,549 | 0.0258 | 31.40 | 29,901 | 789,834 | 0.0379 | 29.76 | 4,950 | 90,249 | 0.0548 | 26.64 |
| 50-54 | 53,139 | 1,397,562 | 0.0380 | 27.13 | 38,641 | 759,933 | 0.0508 | 25.80 | 6,347 | 85,300 | 0.0744 | 23.00 |
| 55-59 | 73,797 | 1,344,422 | 0.0549 | 23.07 | 49,414 | 721,293 | 0.0685 | 22.01 | 7,586 | 78,952 | 0.0961 | 19.61 |
| 60-64 | 101,776 | 1,270,625 | 0.0801 | 19.21 | 62,467 | 671,879 | 0.0930 | 18.40 | 8,699 | 71,366 | 0.1219 | 16.39 |
| 65-69 | 146,618 | 1,168,849 | 0.1254 | 15.57 | 86,065 | 609,412 | 0.1412 | 14.93 | 10,973 | 62,667 | 0.1751 | 13.23 |
| 70-74 | 186,920 | 1,022,231 | 0.1829 | 12.34 | 102,380 | 523,347 | 0.1956 | 11.90 | 12,181 | 51,694 | 0.2356 | 10.45 |
| 75-79 | 245,529 | 835,311 | 0.2939 | 9.37 | 127,660 | 420,967 | 0.3033 | 9.03 | 13,610 | 39,513 | 0.3444 | 7.81 |
| 80+ | 589,781 | 589,781 | 1 | 5.77 | 293,307 | 293,307 | 1 | 5.47 | 25,903 | 25,903 | 1 | 4.41 |

Source: Population data from 2000 and 2010 Brazilian Censuses; Death counts from Ministry of Health, 2005 SIM (Sistema de Informações sobre Mortalidade)

***Multistate life tables.*** Increment-decrement life tables allow for the calculation of "the amount of time that individuals of a cohort will spend in a given state of occupancy" (Muniz 2020). The multistate life table estimates combine transition probabilities presented in Table 1 with race-specific probabilities of dying (qx) shown in Table 2. After partitioning these combined probabilities to sum to one, and after ensuring that racial mobility happens only among those who are alive, we obtain the following transition matrix between White (W), Brown (Br), Black (B) and death (D) states:

**Table 3.** Transition probabilities by race

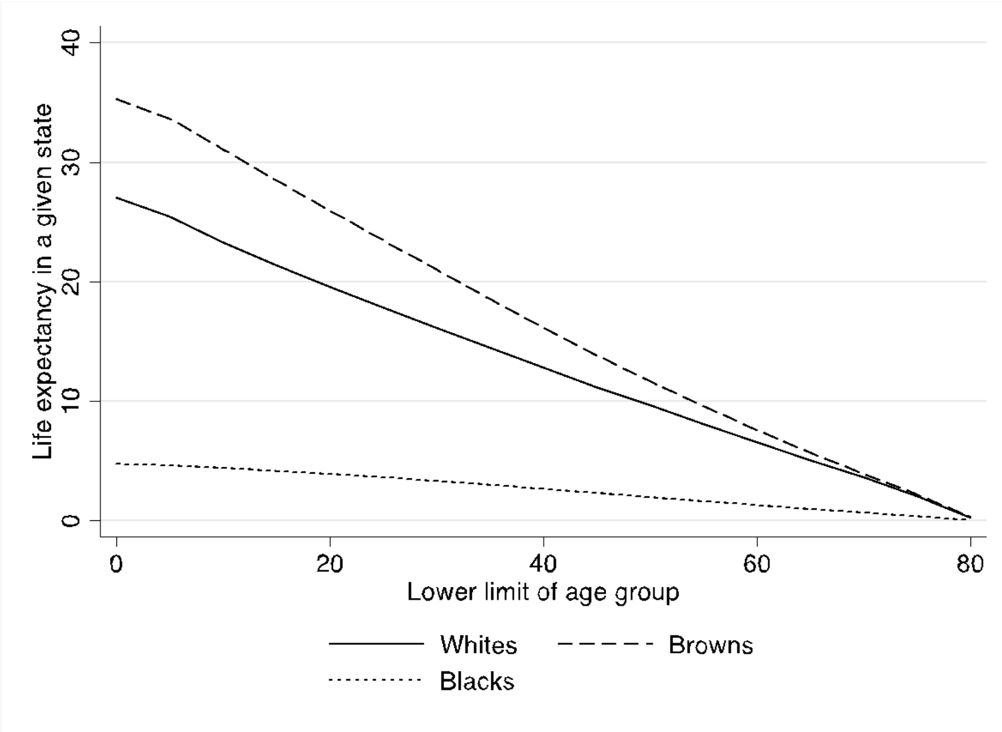
| Age group | W--> Br | W--> B | W--> D | Br--> W | Br--> B | Br--> D | B--> W | B--> Br | B--> D |
|---|---|---|---|---|---|---|---|---|---|
| 0-4 | 0.1283 | 0.0044 | 0.0194 | 0.0876 | 0.0261 | 0.0298 | 0.0335 | 0.2610 | 0.0202 |
| 5-9 | 0.1391 | 0.0053 | 0.0015 | 0.0799 | 0.0301 | 0.0024 | 0.0276 | 0.2853 | 0.0028 |
| 10-14 | 0.1425 | 0.0056 | 0.0018 | 0.0763 | 0.0333 | 0.0029 | 0.0270 | 0.2861 | 0.0033 |
| 15-19 | 0.1407 | 0.0057 | 0.0053 | 0.0771 | 0.0372 | 0.0097 | 0.0239 | 0.2800 | 0.0108 |
| 20-24 | 0.1266 | 0.0061 | 0.0072 | 0.0769 | 0.0383 | 0.0137 | 0.0247 | 0.2540 | 0.0154 |
| 25-29 | 0.1132 | 0.0059 | 0.0077 | 0.0758 | 0.0377 | 0.0151 | 0.0264 | 0.2364 | 0.0178 |
| 30-34 | 0.1108 | 0.0065 | 0.0090 | 0.0733 | 0.0386 | 0.0169 | 0.0267 | 0.2297 | 0.0207 |
| 35-39 | 0.1077 | 0.0063 | 0.0124 | 0.0757 | 0.0396 | 0.0207 | 0.0259 | 0.2329 | 0.0279 |
| 40-44 | 0.1069 | 0.0063 | 0.0172 | 0.0749 | 0.0401 | 0.0277 | 0.0247 | 0.2238 | 0.0384 |
| 45-49 | 0.1029 | 0.0061 | 0.0258 | 0.0772 | 0.0416 | 0.0379 | 0.0245 | 0.2256 | 0.0548 |
| 50-54 | 0.0931 | 0.0055 | 0.0380 | 0.0764 | 0.0417 | 0.0508 | 0.0256 | 0.2184 | 0.0744 |
| 55-59 | 0.0884 | 0.0051 | 0.0549 | 0.0764 | 0.0405 | 0.0685 | 0.0239 | 0.2181 | 0.0961 |
| 60-64 | 0.0835 | 0.0046 | 0.0801 | 0.0767 | 0.0390 | 0.0930 | 0.0229 | 0.2034 | 0.1219 |
| 65-69 | 0.0813 | 0.0043 | 0.1254 | 0.0752 | 0.0373 | 0.1412 | 0.0212 | 0.1865 | 0.1751 |
| 70-74 | 0.0707 | 0.0040 | 0.1829 | 0.0739 | 0.0366 | 0.1956 | 0.0222 | 0.1845 | 0.2356 |
| 75-79 | 0.0676 | 0.0033 | 0.2939 | 0.0699 | 0.0329 | 0.3033 | 0.0203 | 0.1592 | 0.3444 |
| 80+ | 0.0000 | 0.0000 | 1.0000 | 0.0000 | 0.0000 | 1.0000 | 0.0000 | 0.0000 | 1.0000 |

Note: W= Whites; Br= Browns; B= Blacks

This transition matrix shows that, on average, until they are 65 or older, Brazilians have higher probabilities of changing racial identification from White to Brown, or Black to Brown, at each age than their probability of dying. On the other hand, among Brazilians who were identified as White or Brown at birth, or in their first year of life, there is higher probability of child mortality than of being reclassified as Black.

Using the Stata program developed by Muniz (2020) with these data as inputs, and assuming constant force of mortality, we obtain conditional and unconditional life expectancies by age and racial states:

**Figure 1.** Population-based (unconditional) life expectancy at age x in state i




In addition to unconditional life expectancies, we also calculate some key summary measures according to one's state of occupancy:

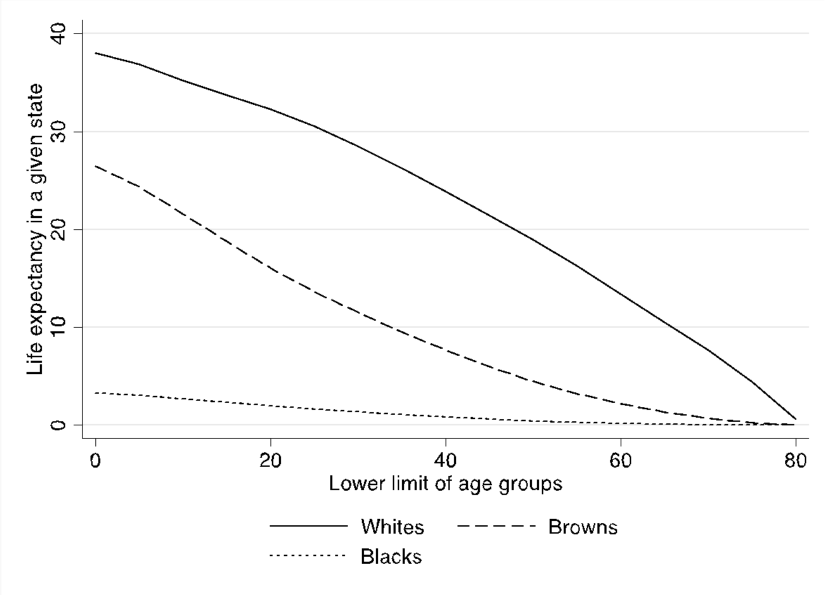
**Table 4.** Summary measures

| | State i<br>White | Brown | Black | Death |
|---|---|---|---|---|
| **Proportion of life spent in state i** | 0.3370 | 0.4400 | 0.0592 | 0.1638 |
| **Probability of dying in state i** | 0.3898 | 0.5073 | 0.0971 | - |
| **Average duration of state i** | 46.7664 | 42.5301 | 16.0225 | - |
| **Mean age of persons in state i** | 35.9561 | 35.4962 | 40.2016 | 61.0117 |

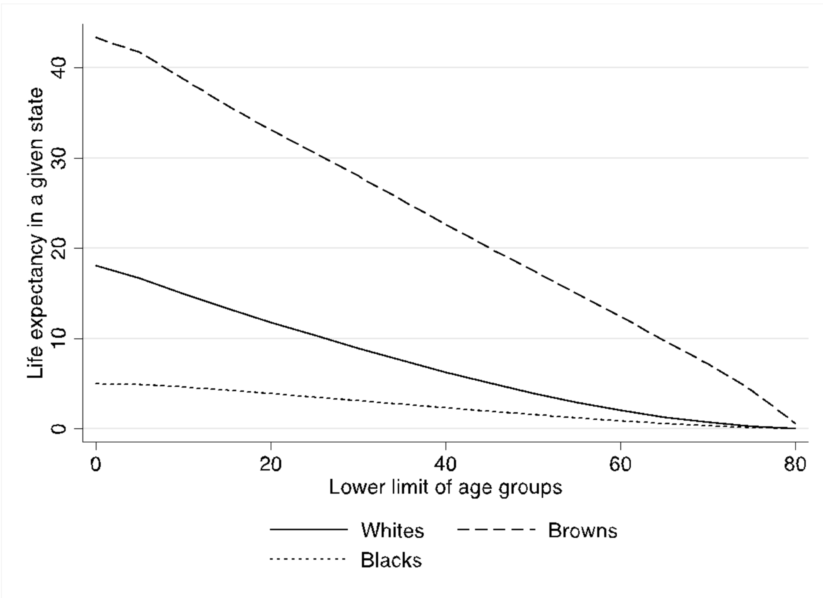
Our results show that, on average, Brazilians will spend about one third of their lives as Whites, 44 percent as Brown and 6 percent as Black. Alternatively, we estimate that a cohort will live about 27 years as Whites, 35 years as Browns, and about 5 years as Blacks, regardless of their state of birth. To obtain the predicted number of years to be lived in a given state *j* by those who are in state *i* at exact age *x,* we also calculated conditional life expectancies, presented in Figure 2.

**Figure 2.** Status-based (conditional) life expectancies at age x in state i

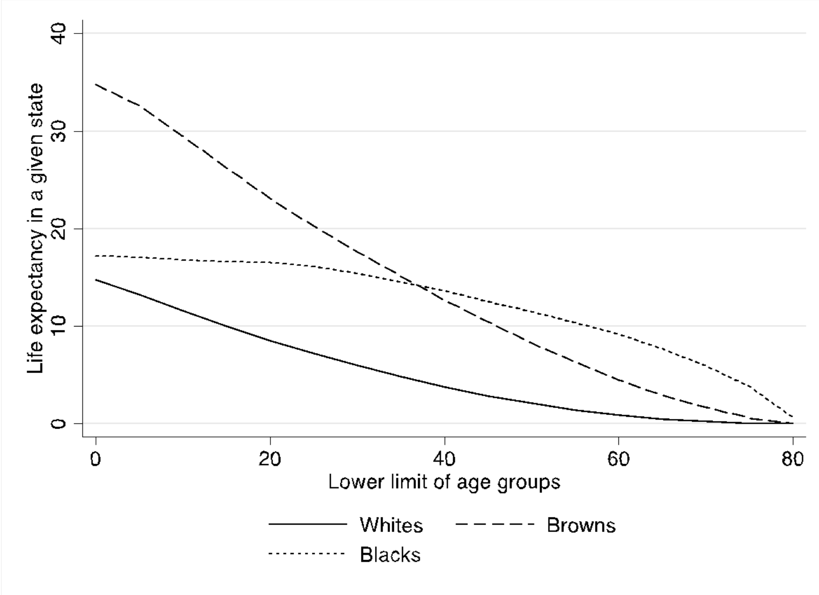
**a. For those who were born White**




**b. For those who were born Brown**



**c. For those who were born Black**




Conditional life expectancies show that individuals tend to spend most of their years in the same state in which they were born. The only exception to this statement are people who were classified as Black at birth, who we estimate spend most of their early years (before age 40) as Brown, and their remaining years of life mostly as Black (see Panel c).

Previous research in the United States has used inconsistent racial classification on birth and death certificates to highlight the fluidity, or unreliability, of racial data (e.g. Hahn et al. 1992). Here, we show that even if someone's racial classification is consistent at birth and death, we should not assume their race remained consistent every year in between. In line with both theory and a growing body of empirical research, increment-decrement life tables allow us to better account for race as a changing propensity rather than a fixed state. They also open up possibilities for measuring temporal and regional variations in racial stability. First, by looking at how these propensities to reclassify from one race to another have shifted over time, would give an indication of how stable or fluid racial boundaries are becoming as time goes by. Second, by examining variations in multistate life tables by race across space would show what regions are more malleable and which ones are more rigid in terms of boundary crossing, even within the same country. Future research should thus incorporate these two dimensions, time and space, into multistate estimates, not only to measure racial mobility over the life course, but also to understand the dynamics of related social processes, such as income mobility and poverty status (Schoen 2020), incarceration, and the lifecycle dynamics of political and religious affiliations.

**References**

Bennett, N. G. and S. Horiuchi. 1984. “Mortality estimation from registered deaths in less developed countries.” *Demography*, v.21, 217-234

Carvalho, J. A. M., Wood, C. H., and Andrade, F. D. (2004). Estimating the sta-bility of census-based racial/ ethnic classifications: the case of Brazil. Population Studies, 58(3), 331–343.

Cornwell, C., Rivera, J., and Schmutte, I. M. (2017). Wage discrimination when identity Is subjective: Evidence from changes in employer-reported race. Journal of Human Resources, 52(3), 719-755.

Davenport, L. (2020). The fluidity of racial classifications. *Annual Review of Political Science*, *23*, 221-240.

Hahn, R.A., Mulinare, J., and Teutsch, S.M. (1992). "Inconsistencies in coding of race and ethnicity between birth and death in US infants: a new look at infant mortality, 1983 through 1985." *Jama* 267(2): 259-263.

Liebler, C. A., Porter, S. R., Fernandez, L. E., Noon, J. M., and Ennis, S. R. (2017). America’s churning races: Race and ethnicity response changes between Census 2000 and the 2010 Census. Demography, 54(1), 259-284.

Merli, G. 1998. “Mortality in Vietnam, 1979-1989.” *Demography* 35(3), 345-360.

Preston, S. H. and N. G. Bennett. 1983. “A census-based method for estimating adult mortality.” *Population Studies* 37(1), 91-104.

Miranda, V. (2014). Measuring racial self-identification over the life course in Brazil, 1940-2013 (Doctoral dissertation). University of Pennsylvania, Philadelphia.

Muniz, J. 2020. Multistate life tables using Stata. The Stata Journal, 20(3), 721-745.

Muniz, J. and Telles, E. 2021. How Migration and Income Affect Racial Classification in Brazil, Presented at the 2021 Meeting of the Population Association of America.

Dahis, R., Nix, E., and Qian, N. (2019). Choosing Racial Identity in the United States, 1880-1940. NBER Working Paper Series, 26465.

Preston, S. H., I. T. Elo, I. Rosenwaike and M. Hill. 1996. “African-American Mortality at Older Ages: Results of a Matching Study.” *Demography* 33(2): 193-209.

Queiroz, B. L. (2007). The determinants of male retirement in urban Brazil. Nova Economia, 17, 11–36.

Ribas, R. P., and Soares, S. S. D. (2009). Sobre o painel da pesquisa Mensal de Emprego (PME) do IBGE: Problemas e soluções para o emparelhamento usando microdados. Revista Brasileira de Estatística, 70(232), 75–108.

Saperstein, A., and Gullickson, A. (2013). A mulatto escape hatch? Examining evidence of U.S. racial and social mobility in the Jim Crow era. Demography, 50(5), 1921–1942.

Saperstein, A., and Penner, A. M. (2012). Racial fluidity and inequality in the United States. American Journal of Sociology, 118(3), 676–727.

Senkevics, A. S. (2017). De brancos para negros? Uma análise longitudinal da reclassificação racial no ENEM 2010-2014. 38ª. Reunião Nacional da ANPEd.

Schoen, R. 2020. Dynamic Multistate Models with Constant Cros-Product Ratios: Applications to Poverty Status. Demography 57, 779-797.

Silveira, L. S. (2019). Reclassificação racial e desigualdade: Análise longitudinal de variações socioeconômicas e regionais no Brasil entre 2008 e 2015 (Unpublished doctoral dissertation). Universidade Federal de Minas Gerais, Belo Horizonte, MG.