

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

**Sobre as Cinzas do Cativo: Educação, Trabalho e Terra entre Ex-escravos e
Descendentes Após a Abolição no Oeste Paulista**

Lucas Cavalcanti Rodrigues

Orientador: Prof. Dr. Renato Perim Colistete

São Paulo

2021

Prof. Dr. Vahan Agopyan
Reitor da Universidade de São Paulo

Prof. Dr. Fábio Frezatti
Diretor da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. José Carlos de Souza Santos
Chefe do Departamento de Economia

Prof. Dr. Wilfredo Fernando Leiva Maldonado
Coordenador do Programa de Pós-Graduação em Economia

LUCAS CAVALCANTI RODRIGUES

Tese apresentada ao Departamento de Economia, Administração e Contabilidade da Universidade de São Paulo como requisito parcial para a obtenção do título de Doutor em Ciências.

Área de concentração: Economia do Desenvolvimento

Orientador: Prof. Dr. Renato Perim Colistete

Versão Corrigida

São Paulo - Brasil

2021

Catálogo na Publicação (CIP)
Ficha Catalográfica com dados inseridos pelo autor

Rodrigues, Lucas Cavalcanti.

Sobre as Cinzas do Cativo: Educação, Trabalho e Terra entre Ex-escravos e Descendentes Após a Abolição no Oeste Paulista / Lucas Cavalcanti Rodrigues. - São Paulo, 2021.

152 p.

Tese (Doutorado) - Universidade de São Paulo, 2021.

Orientador: Renato Perim Colistete.

1. pós-abolição. 2. educação. 3. trabalho. 4. terra. I. Universidade de São Paulo. Faculdade de Economia, Administração e Contabilidade. II. Título.

AGRADECIMENTOS

Historiadores econômicos têm o hábito de dizer que “nada é óbvio”, querendo com isso dizer que qualquer afirmação a respeito da economia e da história deve ser submetida ao escrutínio dos dados e da documentação existente. É uma insistência salutar para a profissão, mas as primeiras palavras deste documento são, sim, uma obviedade: essa tese jamais teria chegado às mãos do leitor sem o suporte e apoio que recebi da minha família e amigos. Uma outra obviedade é que as limitações da memória e o espaço exíguo tornam impossível que eu faça justiça nas citações nominais. De todo modo, aqui está uma tentativa.

Sou grato aos meus pais, Amarildo Rodrigues e Rejane Cavalcanti Rodrigues, por terem sempre apoiado meus estudos e acreditado na minha capacidade. Para economistas, a escolha pelo investimento em educação em famílias pobres deveria ser óbvia: um corolário natural da combinação de imperfeição no mercado de crédito, educação pública disponível e prêmio salarial elevado que caracteriza a economia brasileira. Na vida real, porém, essa escolha não é nada óbvia e não são poucas as pressões e preocupações imediatas que impedem ou dificultam pessoas pobres no Brasil de fazerem as escolhas vistas pelo tecnocrata médio como mais racionais a longo prazo. Por essa razão, contar com o apoio dos meus pais foi fundamental em vários momentos. Tanto mais quando se considera que, em suas vidas escassas em oportunidades educacionais, o propalado “retorno da educação” sempre foi pouco mais que uma abstração.

Meus irmãos, Priscila Cavalcanti Rodrigues e Samuel Cavalcanti Rodrigues, também serviram como fonte permanente de carinho e apoio. Ao longo da minha tese, dois sobrinhos alargaram modestamente em altura e adicionaram dedos às suas contagens de anos. Quando o terceiro sobrinho viu o nascer do Sol pela primeira vez, eu já concluía meu primeiro semestre de doutorado. O amor de cada um deles foi fonte e motivação para meu trabalho. Amo muito vocês Arthur, Sarah e Roney!

É preciso agradecer aos amigos. Felizmente não há poucos nomes para citar aqui. Quero agradecer meus amigos de Juiz de Fora: Bruna Teixeira da Silva, Carolina Moraes Sarmiento de Assis, Henrique Franco, Leonardo Silva Fernandes, Leonardo Cordeiro de Farias Viggiano, Vinícius Thomas de Campos, Vilton Bonifácio e Rogério Silva de Mattos. Os amigos paulistas: Bruna Pugialli da Silva Borges, Lúcia Regina Centurião, Lucas de Arruda Câmara, Matias Rebello Cardomingo, Matheus Assaf, Michael França, Rafael Tavares, Ricardo Sabadini, Tales Rozenfeld e Tiago Ferraz. Entre os amigos paulistas, devo um especial agradecimento a André Lanza e Bruno Witzel que, além de amigos, são colegas de profissão. Essa tese seria um documento completamente diferente, e necessariamente mais pobre, sem a colaboração do André e seu conhecimento sobre imigração europeia em São Paulo. Agradeço imensamente ao Bruno pela leitura e comentários do capítulo sobre educação, além de esclarecimentos sobre o colonato. Ainda na dimensão profissional, estou em eterna dívida também com Bruna Polachini, Wellington Santos da Silva e Juliana Caldas, pelas revisões gramaticais.

Sou grato também a dois economistas que são duas grandes referências intelectuais para mim. Gustavo Barros, por ter sido meu professor e me ensinado muitas coisas, ditas

e não ditas. Serei sempre grato ao Gustavo por ter me incentivado a fazer doutorado e me apresentar ao Renato Perim Colistete. Renato que, fazendo jus à recomendação do Gustavo, foi um orientador de “mão cheia” com quem aprendi muito e continuo aprendendo.

Agradeço também aos professores Dante Mendes Aldrighi e Felipe Pereira Loureiro pelos comentários e sugestões durante as avaliações de progresso. Estou em dívida também com os agentes e instituições de Estado que com seu trabalho, colaboração e financiamento viabilizaram essa pesquisa. Assim, agradeço ao pesquisador Karl Monsma por ter gentilmente me ajudado a sanar minhas dúvidas sobre o acesso ao censo de São Carlos de 1907. Sou grato também aos funcionários do Arquivo Público do Estado de São Paulo e da Fundação Pró-Memória de São Carlos, além de todos os funcionários que fazem a USP funcionar, especialmente aos servidores das bibliotecas da FEA e da FFLCH e dos funcionários do Restaurante Universitário. Agradeço também ao Conselho Nacional de Pesquisa Científica (CNPq) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) que apoiaram financeiramente a pesquisa.

The most common method of direct measurement in history is counting. My reference to counting as a rigorous method of measurement is not to be taken derisively. I use portentous language for what appears to be an elementary operation partly because I want to emphasize the dramatic change in interpretation that may result merely by moving from an impression to an actual count.

Robert William Fogel

RESUMO

No início do século XX, o Oeste paulista prosseguia em sua expansão econômica acompanhando as crescentes exportações de café, que se traduziam em elevada demanda por mão de obra nacional e imigrante. A combinação de expansão econômica com o arrefecimento do ritmo da imigração europeia, em relação aos níveis observados na década de 1890, criou um cenário de relativa escassez de trabalho em São Paulo. A posição de barganha mais favorável dos trabalhadores criou oportunidades de contratos mais vantajosos e melhores condições no mercado de trabalho, inclusive para ex-escravos e descendentes. Esta tese investiga a população negra em São Carlos, um importante município do Oeste Paulista, em meio à expansão econômica da primeira década do século XX e da chegada em massa de imigrantes na região. A tese mostra que a população negra dessa localidade e possivelmente a de municípios vizinhos foi capaz de matricular seus filhos em escolas primárias, ocupar melhores posições no mercado de trabalho, realizar contratos como colonos nas fazendas e até mesmo conseguir acesso à propriedade de terra. Vinte anos após a Abolição, os indicadores socioeconômicos da população negra não divergiam significativamente dos observados para imigrantes de primeira geração. A tese mostra, ainda assim, que permaneciam diferenças substanciais entre as duas populações, particularmente em termos demográficos e educacionais. Famílias imigrantes eram maiores que as famílias negras brasileiras e a mulher imigrante possuía maior taxa de fertilidade. Tais características demográficas representavam maior capacidade de acumulação de recursos pelas famílias imigrantes, mesmo que aparentemente não houvesse diferenças de produtividade entre imigrantes e negros. Por outro lado, a despeito da pequena diferença das taxas de matrícula na educação primária de crianças negras e imigrantes, havia disparidade maior em favor de mães e pais imigrantes em relação aos de famílias negras. Aqui, a herança da escravidão, que praticamente deixou toda a população cativa sem acesso à educação, significou um grande desvantagem para pais e mães das famílias negras. Em uma economia essencialmente rural, a vantagem educacional de imigrantes parece ter produzido um efeito econômico relativamente limitado inicialmente, dada a baixa complementaridade entre capital humano e trabalho agrícola. No entanto, nas décadas seguintes a combinação de urbanização e industrialização mudou o perfil da demanda por mão de obra, beneficiando desproporcionalmente populações com maiores níveis educacionais e favorecendo em especial os imigrantes. Desta forma, ao lado de fatores já destacados pela literatura, como racismo e violência, as evidências reunidas nesta tese sugerem que as disparidades demográficas e dos níveis educacionais tiveram um papel central nas desigualdades acumuladas entre descendentes de imigrantes e descendentes de escravizados em São Paulo ao longo do século XX.

Palavras-chaves: pós-abolição; ex-escravos e descendentes; educação; trabalho; terra

ABSTRACT

At the beginning of the 20th century, the Western region of São Paulo continued to experience the prosperity of its coffee economy, maintaining a high demand for national and immigrant labor. The combination of economic expansion and a slowdown in the pace of European immigration, compared to the levels observed in the 1890s, created a scenario of relative shortage of labor in São Paulo. Workers' increased bargaining position led to better contracts and greater opportunities in the labor market, including for former enslaved persons and descendants. This thesis investigates the black population who lived in São Carlos, a major municipality in Western São Paulo, amidst the rapid economic expansion and the arrival of immigrants in the first decade of the 20th century. The thesis shows that São Carlos's black population and possibly that of neighboring municipalities could enroll their children in primary schools, obtain better positions in the labor market, establish contracts as *colonos* on plantations, and even become landowners. Twenty years after Abolition, the socio-economic indicators of the black population did not significantly diverge from those observed for first-generation immigrants. Even so, the thesis shows that there remained substantial differences between the two populations, particularly in terms of demographic and educational conditions. Immigrant families were larger than black Brazilian families, and immigrant women had a higher fertility rate. Such demographic characteristics represented a higher capacity for capital accumulation by immigrant families, even though there were apparently no differences in productivity between immigrants and blacks. On the other hand, despite the small difference in schooling between black and immigrant children, immigrant mothers and fathers were better educated than their counterparts of black families. Here, the legacy of slavery, which left practically the entire enslaved population without access to education, meant a significant disadvantage for fathers and mothers from black families. In an essentially rural economy, the educational advantage of immigrants may have been of a limited economic effect at first, given the low complementarity between human capital and agricultural labor. In the following years, however, the combination of urbanization and industrialization changed the demand for labor, disproportionately benefiting populations with higher educational levels and favoring immigrants. Thus, alongside factors already highlighted in the literature, such as racism and violence, the data presented in this thesis suggest that the considerable disparities in demography and educational levels played a key role in explaining the inequalities between descendants of immigrants and descendants of enslaved people in São Paulo in the 20th century.

Keywords: post-abolition; former slaves and descendants; education; labor; land

Lista de Figuras

1.1	Divisão por regiões e municípios do estado, São Paulo, 1900	10
1.2	Crescimento demográfico e mudança na distribuição espacial da população, São Paulo, 1836-1900	12
1.3	População por cor e região, São Paulo, 1872/1890	15
1.4	Produção cafeeira, em milhões de arrobas, São Carlos, 1860-1937 . .	20
1.5	Distribuição da população por grupo demográfico e idade, São Carlos, 1907	25
1.6	Pirâmide etária de brasileiros brancos, imigrantes e negros	29
1.7	Razão sexual por grupo etário, São Carlos, 1907	30
1.8	Tamanho da família por grupo populacional, São Carlos, 1907	32
2.1	Exemplo de página de livro de chamada de escola primária de São Carlos, Julho de 1909	41
2.2	Distribuição de cor por bairro, São Carlos, 1907	45
2.3	Alunos de escolas de Ibaté, São Carlos, 1917	53
3.1	<i>IER</i> para imigrantes e negros em ocupações domésticas, São Carlos, 1907	73
4.1	Municípios limítrofes de São Carlos, 1900	97
4.2	Matriz de Confusão	108
4.3	ROC - <i>Multi-layer Perceptron Classifier</i>	109
4.4	Distribuição acumulada da área total entre Brasileiros brancos, Imigrantes e Negros, São Carlos, 1907	114
D.1	Distribuição de nacionalidade por bairro, São Carlos, 1907	152

Lista de Tabelas

1.1	Percentual de nacionais e imigrantes europeus por região de destino, São Paulo, 1890-1910	17
1.2	Casamentos inter e intra-grupos, São Carlos, 1907	24
1.3	Taxa de alfabetização maiores de 15 anos (em percentual), São Carlos, 1907	27
1.4	Percentual de proprietários de terra, São Carlos, 1907	27
2.1	Número de crianças matriculadas por escola e ano, São Carlos, 1906-1909	40
2.2	Cor de alunos e da população até 19 anos, São Carlos, 1907	43
2.3	Cor de estudantes por escola, São Carlos, 1907	43
2.4	Índice de Levenshtein por cor dos estudantes, em percentual, São Carlos, 1907	46
2.5	Profissão dos Pais dos estudantes e da população masculina, São Carlos, 1907 (em porcentagem da respectiva população)	48
2.6	Profissão das Mães dos estudantes e da população feminina, São Carlos, 1907 (em porcentagem da respectiva população)	49
2.7	Alfabetização entre pais de estudantes e na população, São Carlos, 1907	51
2.8	Taxas de matrícula por cor, São Carlos, 1907	56
2.9	Resultados	58
2.10	Efeitos marginais - <i>logit</i> e <i>probit</i>	61
3.1	IER das ocupações dos grupos 1 a 3, População negra masculina, São Carlos, 1907	71

3.2	IER das ocupações dos grupos 1 a 3, População negra feminina, São Carlos, 1907	75
3.3	IER das ocupações dominadas por não-negros, População masculina negra, São Carlos, 1907	78
3.4	IER das ocupações em que negros são excluídos, População masculina negra, São Carlos, 1907	80
4.1	Dados demográficos de proprietários de terra, São Carlos, 1905/1907	87
4.2	Dados de propriedades de negros, São Carlos, 1905/1907	89
4.3	Estatísticas Descritivas da área total de propriedades rurais de Brasileiros brancos, Imigrantes e Negros, São Carlos, 1905/1907	92
4.4	Indicadores relativos à propriedade rural de Brasileiros brancos, Imigrantes e Negros, São Carlos, 1905/1907	92
4.5	Evolução demográfica, Região de São Carlos, 1890-1900	98
4.6	Estatísticas fundiárias, Região de São Carlos, 1905	99
4.7	Nomes e sobrenomes mais comuns de negros, percentuais e números absolutos, São Carlos, 1907	102
4.8	Sobrenomes tradicionais por grupos demográficos, São Carlos, 1907	103
4.9	Resultados de modelos preditivos selecionados	104
4.10	Variáveis preditivas utilizadas nos modelos	105
4.11	Métricas de desempenho da classificação com <i>Multi-layer Perceptron Classifier</i>	107
4.12	Distribuição da área de propriedades segundo o melhor modelo de classificação, Municípios limítrofes de São Carlos, 1905	112
A.1	Classificação dos <i>matchings</i>	135
A.2	Seleção do matching do proprietário “Manoel Bangoiro dos Santos” listado na Estatística Agrícola, São Carlos 1905/1907	139
C.1	Profissão dos Pais negros dos estudantes e da população masculina negra, São Carlos, 1907 (em porcentagem da respectiva população) .	144
C.2	Profissão dos Mães negras dos estudantes e da população feminina negra, São Carlos, 1907 (em porcentagem da respectiva população) .	145

C.3	Distribuição etária entre estudantes, São Carlos, 1906-1909	145
C.4	Resultados - adicionando variável de alfabetização do pai	146
C.5	Resultados - usando <i>dummies</i> para preto e mulato	147
C.6	IER das ocupações dos grupos 1 a 3, População mulata masculina, São Carlos, 1907	148
C.7	IER das ocupações dos grupos 1 a 3, População mulata feminina, São Carlos, 1907	149
C.8	IER das ocupações dos grupos 1 a 3, População preta masculina, São Carlos, 1907	149
C.9	IER das ocupações dos grupos 1 a 3, População preta feminina, São Carlos, 1907	150
C.10	Indicadores relativos à propriedade rural de Brasileiros brancos, Imigrantes e Negros (sem <i>outlier</i>), São Carlos, 1905/1907	150
C.11	Estatísticas Descritivas da área total de propriedades rurais de Bra- sileiros brancos, Imigrantes e Negros (sem <i>outlier</i>), São Carlos, 1905/1907	151

Sumário

Introdução	1
1 Economia e demografia no pós-abolição em São Carlos	7
1.1 A economia e a demografia de São Paulo na segunda metade do século XIX	9
1.2 População e economia de São Carlos no início do século XX	18
1.2.1 A população negra de São Carlos após a Abolição	22
1.3 Conclusões	33
2 A população negra e a educação: nível e determinantes da matrícula escolar	35
2.1 Os dados educacionais: livros de chamadas e o censo municipal . . .	38
2.2 Determinantes da matrícula escolar: uma abordagem a partir das escolas	41
2.2.1 Características das famílias dos estudantes	46
2.3 Determinantes da matrícula escolar: uma análise quantitativa a partir do censo municipal	54
2.3.1 Análise de regressão	56
2.4 Conclusões	62
3 A população negra e o trabalho: o legado da escravidão e as ocupações	65
3.1 Mercado de trabalho e as ocupações em São Carlos	66
3.1.1 Classificação e resultados	69
3.2 Conclusão	81

4 A população negra e a terra: aspectos quantitativos e qualitativos do acesso à propriedade rural	83
4.1 Proprietários negros em São Carlos	85
4.2 Para além de São Carlos: proprietários negros em outros municípios paulistas	95
4.2.1 Nomes e cor da pele	99
4.2.2 Classificação de proprietários	104
4.2.3 Classificação nos municípios limítrofes de São Carlos	111
4.3 Conclusão	115
Conclusão	117
Referências	122
Apêndice A – Algoritmo de <i>matching</i> de nomes	131
A.1 Primeiro estágio	131
A.2 Segundo estágio: <i>matching</i> dos estudantes	135
A.3 Segundo estágio: <i>matching</i> dos proprietários	137
Apêndice B – Índice de Equilíbrio Racial	141
Apêndice C – Tabelas Complementares	144
Apêndice D – Figuras Complementares	152

Introdução

Refletindo sobre o 13 de maio de 1888, evento que presenciou pessoalmente, Lima Barreto escreveria em tom melancólico mais de duas décadas depois: “[o tempo] vai ceifando aspirações, tirando presunções, trazendo desalentos, e só nos deixa na alma essa saudade do passado às vezes composta de coisas fúteis, cujo relembrar, porém, traz sempre prazer. Quanta ambição ele não mata! Primeiro são os sonhos de posição: com os dias e as horas e, a pouco e pouco, a gente vai descendo de ministro a amanuense [...]”.¹

A perspectiva de Lima Barreto sobre os significados da Abolição revela um pessimismo que contrasta com a trajetória de rápida ascensão social de sua família. Nascido no mesmo dia e mês da Lei Áurea, mas sete anos antes, Lima Barreto descendia de escravizados tanto da parte de pai quanto da parte de mãe. Seus pais, no entanto, faziam parte do pequeno grupo de brasileiros alfabetizados, com profissões urbanas, que começavam a ascender nos centros urbanos brasileiros no contexto da modernização e desenvolvimento da sociedade brasileira em fins do século XIX. A mãe, Amália Augusta, tornou-se professora e diretora de escolas para crianças. O pai, Afonso Henriques, exerceu o ofício de tipógrafo e, entre outras coisas, foi responsável pela tradução de uma obra técnica do francês para o português.²

Uma boa parte da historiografia sobre o pós-abolição sugere que a história familiar de Lima Barreto é uma absoluta exceção. Em regra, brasileiros negros teriam falhado em se integrar à nova sociedade baseada no trabalho livre nos anos

1. *Maio* - Crônica publicada em 04/05/1911 no jornal *Gazeta da Tarde*.

2. Lilia Moritz Schwarcz, *Lima Barreto-triste visionário* (Editora Companhia das Letras, 2017), capítulo 1.

que se seguiram à escravidão. Desprovidos de uma ética de trabalho própria de sociedades modernas e sofrendo com o racismo, ex-escravos e descendentes estariam condenados à marginalização social e ao isolamento.³ Histórias como a da família de Lima Barreto, em vez de revelar autonomia e determinação individual seriam, na verdade, subproduto de ações benfeitoras de brancos. Afinal de contas, crucial para o sucesso pessoal de Amélia Augusta e Afonso Henriques foi a patronagem pessoal do médico Manuel Feliciano Pereira Carvalho, no primeiro caso, e do senador Afonso Celso de Assis Figueiredo, no segundo.

Há, porém, limitações grandes na interpretação de parte da historiografia sobre as implicações socioeconômicas da Abolição para a população negra brasileira. A primeira é que a ideia de que a escravidão tenha legado aos descendentes de escravizados vícios morais, transformando-os em indivíduos desprovidos de vontade própria e incapazes de autonomia em uma sociedade livre, encontra pouco respaldo em evidências históricas. Há uma literatura substancial no Brasil que tem mostrado a importância de iniciativas e ações de descendentes de escravizados no pós-abolição.⁴ Pesquisas sobre outras sociedades escravistas também indicam que ex-escravizados ocuparam uma posição importante no mercado de trabalho e como proprietários após o fim da escravidão. No Sul dos Estados Unidos, por exemplo, estudos mostram que o período posterior à Emancipação é marcado por melhoria significativa nos indicadores socioeconômicos da população negra. Mesmo sofrendo com racismo e violência, a população de ex-escravos e descendentes americana foi capaz de se educar, melhorar sua renda e patrimônio e diminuir as disparidades socioeconômicas em relação aos brancos, construída durante o período escravista. Em 1870, o percentual de alfabetizados entre maiores de 9 anos na população negra era de 20%, ao passo que em 1910 esse percentual havia saltado para 70%.⁵ A convergência educacional de negros e brancos

3. Florestan Fernandes, *A integração do negro na sociedade de classes*, vol. 1 (São Paulo: Dominus editora, 1965), cap. 1-6.

4. Sidney Chalhoub, *Visões da liberdade: uma história das últimas décadas da escravidão na corte* (Editora Companhia das Letras, 2011), Cap. 2; Flávio dos Santos Gomes, *Negros e política (1888-1937)* (Zahar, 2005), 27-44; Hebe Mattos, *Das cores do silêncio: os significados da liberdade no sudeste escravista*, 3ª ed. (Editora Unicamp, 2013 [1995]), 297-338.

5. Robert Higgs, *Competition and coercion: Blacks in the American economy 1865-1914* (Cambridge University Press, 1977), 120.

no pós-emancipação refletiu-se em uma convergência de renda, já que entre 1870 e 1900 a renda média de negros em relação a brancos subiu de 24% para 35%.⁶ Há também evidência de convergência, ainda que mais modesta, nos níveis de riqueza.⁷

Uma segunda limitação da historiografia é que existe pouca evidência sistemática para a hipótese de que, após a Abolição, a população negra brasileira estaria excluída das escolas, marginalizada no mercado de trabalho e incapaz de acumular riqueza. Em parte, a escassez de evidências sobre o pós-abolição no Brasil deve-se à ausência de informação de cor da pele nos censo nacionais levados a cabo no período. A informação de cor da pele está disponível no censo de 1890, mas ausente nos censos de 1900 e 1920. Apenas em 1940 o censo brasileiro voltou a apresentar dados desagregados por cor da pele. Na ausência de levantamentos sistemáticos e abrangentes, alguns pesquisadores têm se valido de fontes alternativas, como ilustrado em importantes estudos sobre o mercado de trabalho no período, que fazem uso de dados de firmas individuais ou de documentação criminal.⁸ Essa estratégia, porém, apresenta limitações. Há, por exemplo, problemas de auto-seleção associados ao estudo de firmas individuais, o que limita as possibilidades de inferência dos estudos para além dos casos estudados. Não só não é possível saber se os resultados observados valem para outras firmas, como também não está claro a amplitude regional das conclusões desses estudos. Ao se limitarem a empresas industriais de certas cidades ou a documentos produzidos e mantidos pelo aparelho repressivo do Estado, esses estudos tendem a se circunscrever a centros urbanos, em um período em que a grande maioria dos

6. Higgs, *Competition*, 146.

7. Robert Higgs, "Accumulation of property by Southern blacks before World War I", *The American Economic Review*, 1982, 725-737; Robert A Margo, "Accumulation of Property by Southern Blacks Before World War I: Comment and Further Evidence", *American Economic Review*, 1984, 768-776.

8. George Reid Andrews, *Negros e brancos em São Paulo, 1888-1988* (Editora da Universidade do Sagrado Coração, 1998 [1991]), Cap. 4; Ramatis Jacino, "O negro no mercado de trabalho em São Paulo pós-abolição-1912/1920" (tese de doutorado, Universidade de São Paulo, 2012); Molly Catherine Ball, "Inequality in São Paulo's Old Republic: A Wage Perspective, 1891-1930" (tese de doutorado, University of California, Los Angeles, 2013); Hildete Pereira de Melo, João Lizardo de Araújo e Teresa Cristina de Novaes Marques, "Raça e nacionalidade no mercado de trabalho carioca na Primeira República: o caso da cervejaria Brahma", *Revista Brasileira de Economia* 57, n. 3 (2003): 535-569.

brasileiros vivia na zona rural e trabalhava na agricultura.

Esta tese, ao usar uma fonte censitária produzida para um importante município do interior paulista, oferece uma oportunidade de se testar empiricamente hipóteses da literatura sobre o pós-abolição, ao mesmo tempo em que procura superar as limitações de fontes, encontradas em muitos estudos para o período. Para tanto, usa-se como principal fonte documental do estudo um raro censo produzido cerca de 20 anos após a Abolição que contém, entre outras informações, a cor da pele dos indivíduos. O censo foi realizado pela Câmara Municipal de São Carlos e, assim, constitui-se em fonte detalhada das condições socioeconômicas e demográficas no interior do estado de São Paulo no pós-abolição.⁹

Usando os dados de São Carlos, esta tese apresenta estimativas do acesso à educação básica, mercado de trabalho e à terra para a população de ex-escravos e descendentes. Argumenta-se, no Capítulo 1, que a história da população negra paulista no fim do século XIX guarda similaridades com a história de imigrantes europeus. Procurando novas oportunidades para si mesmos e para seus filhos, parte da população negra foi atraída para a fronteira agrícola onde a elevada demanda por mão de obra, combinada com as novas modalidades de contrato do trabalho agrícola, lograram criar oportunidades econômicas para trabalhadores. Assim como a população imigrante, trabalhadores negros teriam se beneficiado do bom momento do café e experimentado um período de relativa prosperidade econômica.

A expansão da economia teria consequências para as possibilidades de ascensão social da população negra, viabilizando acesso à escola, melhores posições no mercado de trabalho e acesso à terra. Estimativas de acesso à educação básica são apresentadas no Capítulo 2 onde, a partir do cruzamento dos dados do censo municipal de 1907 com listas de chamadas de escolas públicas isoladas, constata-se que crianças pretas e mulatas estavam presentes nas escolas primárias em ta-

9. Outros trabalhos utilizaram pioneiramente o censo de São Carlos de 1907 para estudar os anos após a Abolição, como Oswaldo Mário Serra Truzzi e Maria Silvia Beozzo Bassanezi, "População, grupos étnico-raciais e economia cafeeira: São Carlos, 1907", *Revista Brasileira de Estudos de População* 26, n. 2 (2009): 197-218; Karl Monsma, *A reprodução do racismo: fazendeiros, negros e imigrantes no oeste paulista, 1880-1914* (São Carlos: Editora da Universidade Federal de São Carlos, 2016)

xas similares às de crianças brancas. Embora, entre adultos, o nível de educação de negros fosse inferior ao de brasileiros brancos e imigrantes, as estimativas de acesso à educação primária indicam que, mantidas as taxas de matrícula nos níveis estimados, a população negra lograria diminuir o *gap* de capital humano com descendentes de brancos em poucas gerações.

O Capítulo 3 aborda o mercado de trabalho, examinando a estrutura ocupacional de São Carlos em 1907, considerando número e tipos de trabalho em termos de habilidades para homens e mulheres. Apresenta-se um índice para avaliar a equidade racial em diferentes ocupações no mercado de trabalho da época. Além de revisitar achados de Karl Monsma, o cálculo do índice permite identificar sub e sobre-representação para diferentes subgrupos demográficos em cada ocupação.¹⁰ Além disso, a partir do índice foi criada uma classificação que permite observar as diferenças gerais entre a estrutura ocupacional de brasileiros brancos, negros e imigrantes. Os resultados indicam, entre outras coisas, a posição mais favorável de mulatos em relação a pretos na estrutura ocupacional.

Por fim, no Capítulo 4, apresenta-se uma análise detalhada dos aspectos quantitativos e qualitativos do acesso à terra entre negros em 1907. A análise é implementada com o cruzamento dos dados do censo de São Carlos com a Estatística Agrícola de São Paulo de 1905. Nesse capítulo, o escopo regional é ampliado e abrange também municípios vizinhos de São Carlos, como Araraquara e Rio Claro. A presença de negros entre proprietários de terra, tanto em São Carlos quanto nas estimativas para os municípios vizinhos, é um dos principais resultados do capítulo. Além de trazer novas evidências e conclusões sobre a situação de ex-escravos e descendentes após a Abolição, esta tese também inova nos métodos utilizados, em duas frentes complementares. Em primeiro lugar, foram implementados algoritmos de *matching* para efetuar o cruzamento de fontes nominativas de acordo com um procedimento formal, o que permite a replicação dos resultados do estudo por parte de outros pesquisadores. Em segundo lugar, foram utilizados algoritmos de classificação de *machine learning* para inferir informações no contexto de registros documentais incompletos.

Diferentemente do que seria esperado por parte da literatura sobre o pós-

10. Monsma, *A reprodução*, 92-101.

abolição, esta tese mostra que a população negra de um destacado município da expansão cafeeira do Oeste paulista teve um importante papel no pós-abolição, tanto na educação, nos mercados de trabalho rural e urbano, e até mesmo na categoria de proprietários rurais. Se o legado do passado escravocrata ainda grassava sobre as instituições, ex-escravos e descendentes procuraram, com os meios disponíveis, reescrever sua história no país que se reconstruía sobre as cinzas do cativeiro.

Capítulo 1

Economia e demografia no pós-abolição em São Carlos

Ao longo da segunda metade do século XIX, a rápida expansão cafeeira levou São Paulo a uma posição de destaque na economia brasileira. O dinamismo da atividade econômica na antiga província e depois estado teve como marca uma elevada heterogeneidade regional, com o Oeste paulista materializando, como nenhuma outra região, a prosperidade da economia cafeeira.¹ A consolidação do café no Oeste foi acompanhada por restrições cada vez maiores à oferta de trabalho escravo, que se mantinha como o pilar da produção nas fazendas da província apesar da entrada de imigrantes europeus desde a década de 1840.² A rápida expansão econômica e a simultânea redução da tradicional oferta de trabalho cativo ameaçavam aumentar o custo do trabalho de maneira inédita para fazendeiros habituados a uma oferta abundante de braços para a lavoura.³ A

1. Pierre Monbeig, *Pioneiros e fazendeiros de São Paulo* (São Paulo: HUCITEC, 1984), 167-172; Sérgio Milliet, *Roteiro do café e outros ensaios: contribuição para o estudo da história econômica e social do Brasil*, 3ª ed. (São Paulo: Departamento de Cultura, 1941), 5-72.

2. Warren Dean, *Rio Claro: um sistema brasileiro de grande lavoura, 1820-1920* (Rio de Janeiro: Paz e Terra, 1977), Cap. 4; Maria Tereza Schorer Petrone, "Imigração", em *O Brasil republicano: sociedade e instituições (1889-1930)*, vol. 3 (Bertrand Brasil, 1977); Sérgio Buarque de Holanda, "Prefácio", em *Memórias de um colono no Brasil*, (Autor: T. Davatz) (São Paulo: Livraria Martins, 1941).

3. Thomas H Holloway, *Imigrantes para o café: café e sociedade em São Paulo, 1886-1934* (Rio de Janeiro: Paz e Terra, 1984), cap. 3; Bruno Gabriel Witzel de Souza, "From Bonded Laborers to Educated Citizens? Immigration, Labor Markets, and Human Capital in São Paulo, Brazil (1820-2010)" (tese de doutorado, Georg-August-Universität Göttingen, 2019), cap.2.

conhecida resposta da imigração subsidiada pelo governo de São Paulo a partir de 1884 serviu não apenas para evitar a explosão do custo do trabalho no período posterior à Abolição, mas também teve consequências demográficas para todas as regiões do estado de São Paulo.

São Carlos foi um dos municípios importantes da expansão cafeeira e constitui um exemplo representativo das mudanças observadas na economia e na demografia do Oeste paulista.⁴ A partir de um recenseamento realizado pela administração municipal em 1907, neste capítulo são apresentados dados da economia e demografia de São Carlos e analisadas evidências de que a elevada demanda por mão de obra no Oeste paulista, sendo apenas parcialmente satisfeita pela oferta de trabalho de imigrantes, acabou por criar condições contratuais relativamente favoráveis para trabalhadores nacionais, inclusive para ex-escravos e descendentes após a Abolição. Em especial, o crescimento econômico da região da qual São Carlos fazia parte abriu espaço para que ex-escravos e descendentes tivessem acesso à educação, ao trabalho e à terra em níveis similares àqueles observados para a população imigrante, embora substancialmente aquém do alcançado pela população branca nacional, conforme será visto em outros capítulos desta tese. Não obstante, as oportunidades trazidas pela expansão econômica não parecem ter sido suficientes para atenuar diferenças estruturais entre a população negra e a população branca, manifestadas nos indicadores demográficos e de capital humano.⁵ Dessa forma, ao lado de fatores já destacados pela literatura como racismo e violência, diferenças na demografia e educação ajudam a explicar as desigualdades observadas entre descendentes de imigrantes e descendentes de escravizados em São Paulo.⁶

Este primeiro capítulo está dividido em três seções. Com o objetivo de apresentar o contexto histórico da pesquisa, são resumidos inicialmente fatos e dados gerais acerca da economia e demografia paulista na segunda metade do século

4. No início do século XX, o hoje município de São Carlos era chamado de “São Carlos do Pinhal”. Ao longo desta tese, contudo, será utilizada a denominação atual do município.

5. Nesta tese, define-se *negros* como o grupo de indivíduos classificados como *pretos* ou *mulatos* pelo censo municipal de São Carlos de 1907. Ver seção 1.2.1 para mais detalhes.

6. Sobre trabalhos que documentam o racismo na sociedade e no mercado de trabalho ver Fernandes, *A integração*; Andrews, *Negros e Brancos*. Em relação à violência, ver Monsma, *A reprodução* que também estudou o caso de São Carlos.

XIX e início do século XX. Em seguida, apresenta-se um breve panorama do município de São Carlos no início do século XX, reconstituindo linhas gerais de sua história bem como suas especificidades demográficas e socioeconômicas, com ênfase na população negra local. A seção final conclui o capítulo, resumindo os principais pontos discutidos.

1.1 A economia e a demografia de São Paulo na segunda metade do século XIX

A segunda metade do século XIX no Brasil assistiu à consolidação econômica da região sudeste, à ascensão do café como principal produto de exportação brasileiro e à redução da importância relativa das antigas lavouras de açúcar e algodão do nordeste.⁷ Ao mesmo tempo, a expansão das ferrovias produziu uma substancial redução nos custos de transporte, viabilizando a integração econômica de regiões do interior do país.⁸ Em termos demográficos, consolidou-se uma nova distribuição populacional com peso crescente das regiões do sudeste e sul do país, como efeito da nova hierarquia regional produzida pelas mudanças produtivas.⁹

Em um primeiro momento, a atividade cafeeira concentrou-se na zona costeira dos estados do Espírito Santo, Rio de Janeiro e São Paulo, com destaque para o Vale do Paraíba, região localizada nesses dois últimos estados, onde a produção de café teria sua primeira grande expansão no século XIX.¹⁰ São Paulo, que se tornaria o grande produtor de café do mundo, somente tomou a liderança da produção nacional na década 1880, quando a expansão da fronteira do café levou à ocupação do chamado Oeste paulista, em particular na região tributária das estradas de ferro Mogiana e Paulista.¹¹ A Figura 1.1 apresenta a divisão do

7. Caio Prado Júnior, *História econômica do Brasil* (São Paulo: Brasiliense, 2006 [1945]), 114.

8. William R Summerhill, *Order against progress* (Stanford University Press, 2003), 58.

9. Emília Viotti Da Costa, *Da senzala à colônia*, 5ª ed. (Unesp, 2010 [1997]), 228-243; Luiz Aranha Corrêa do Lago, *Da escravidão ao trabalho livre: Brasil, 1550-1900* (Editora Companhia das Letras, 2014 [1978]), 66.

10. Lago, *Da escravidão*, 64-65.

11. Ao longo de todo o texto, adota-se a divisão regional de Milliet na versão modificada de Holloway.

estado de São Paulo segundo as regiões tributárias das ferrovias e os respectivos municípios no ano de 1900. São Carlos, com destaque na cor branca no mapa, formava um dos municípios da região da Paulista.

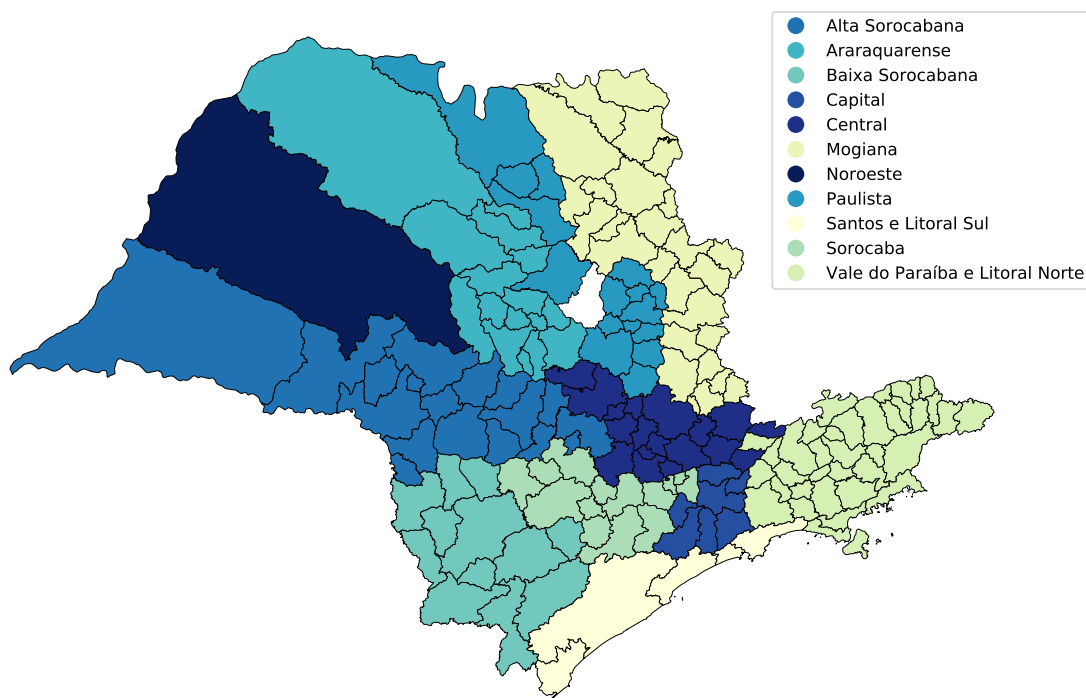


Figura 1.1: Divisão por regiões e municípios do estado, São Paulo, 1900

Fonte: IBGE.

Nota: São Carlos destacado em branco no mapa
Divisão regional baseada em Holloway (1984).

A mudança do eixo econômico do Vale do Paraíba para o Oeste paulista ao longo do século XIX pode ser observada analisando a evolução da distribuição regional da produção cafeeira e as estatísticas populacionais antes e depois da década de 1850. Enquanto em 1836 o Vale do Paraíba respondia por 86,5% da produção de café de São Paulo, com uma produção de 510.406 arrobas, seu peso relativo caiu para 20% da produção cafeeira paulista em 1886. Enquanto isso, as regiões da Mogiana e da Paulista, que conjuntamente respondiam por apenas 1,5% da produção de café em 1836, passaram a responder por 45,5% da produção

paulista em 1886.¹²

O café também tornou-se a principal força por trás da transformação demográfica testemunhada por São Paulo ao longo do século XIX.¹³ A economia cafeeira afetou a população em ao menos três dimensões: seu crescimento, sua distribuição ao longo do território e sua composição por cor em cada região. Em relação ao primeiro aspecto, o crescimento demográfico de São Paulo entre 1836 e 1900 foi da ordem 3,3% ao ano, o que levou sua população, cujo total era de 284.312 habitantes no início do período, a alcançar a marca de 2.279.608 em 1900. Esse crescimento, por sua vez, foi marcado por duas características. Em primeiro lugar, houve claro desbalanceamento na intensidade de crescimento antes e depois de 1886.¹⁴ O crescimento populacional entre os levantamentos de 1836 e 1886 foi de 3% ao ano, enquanto entre 1886 e 1900 a população aumentou à taxa anual de 4,6%. Embora à primeira vista essa diferença possa parecer pequena, ela de fato produz diferenças dramáticas na variação do número absoluto da população entre cada um dos levantamentos considerados. A população paulista aumentou em 937.068 habitantes no meio século entre 1836 e 1886, mas no curto período de quatorze anos, entre 1886 e 1900, o número de habitantes paulistas elevou-se em 1.058.228 pessoas (ver primeiro painel na Figura 1.2).

A segunda dimensão da evolução demográfica paulista no século XIX é a que se refere à redistribuição da população no território do estado. Essa redistribuição seguiu a expansão da produção cafeeira mencionada anteriormente. O Vale do Paraíba, que concentrava 37% da população em 1836, passou a responder por apenas 17% em 1900. No outro extremo, a região de Mogiana, que tinha apenas 7% da população paulista em 1836, somou mais de 20% da população do estado em 1900 (ver segundo painel na Figura 1.2). Essa realocação espacial foi o produto direto dos diferentes ritmos de crescimento observados nas populações de cada uma das regiões do estado ao longo do século XIX, com taxas anuais de crescimento populacional entre 1836 e 1900 bastante variadas. A região de crescimento mais modesto foi justamente o Vale do Paraíba, cuja população aumentou 2,2% ao

12. Milliet, *Roteiro*, 19,20.

13. Jose Francisco de Camargo, *Crescimento da população no Estado de São Paulo e seus aspectos econômicos*, vol. 1 (São Paulo: Instituto de Pesquisas Econômicas, 1981), 269.

14. *Ibid.*, 53.

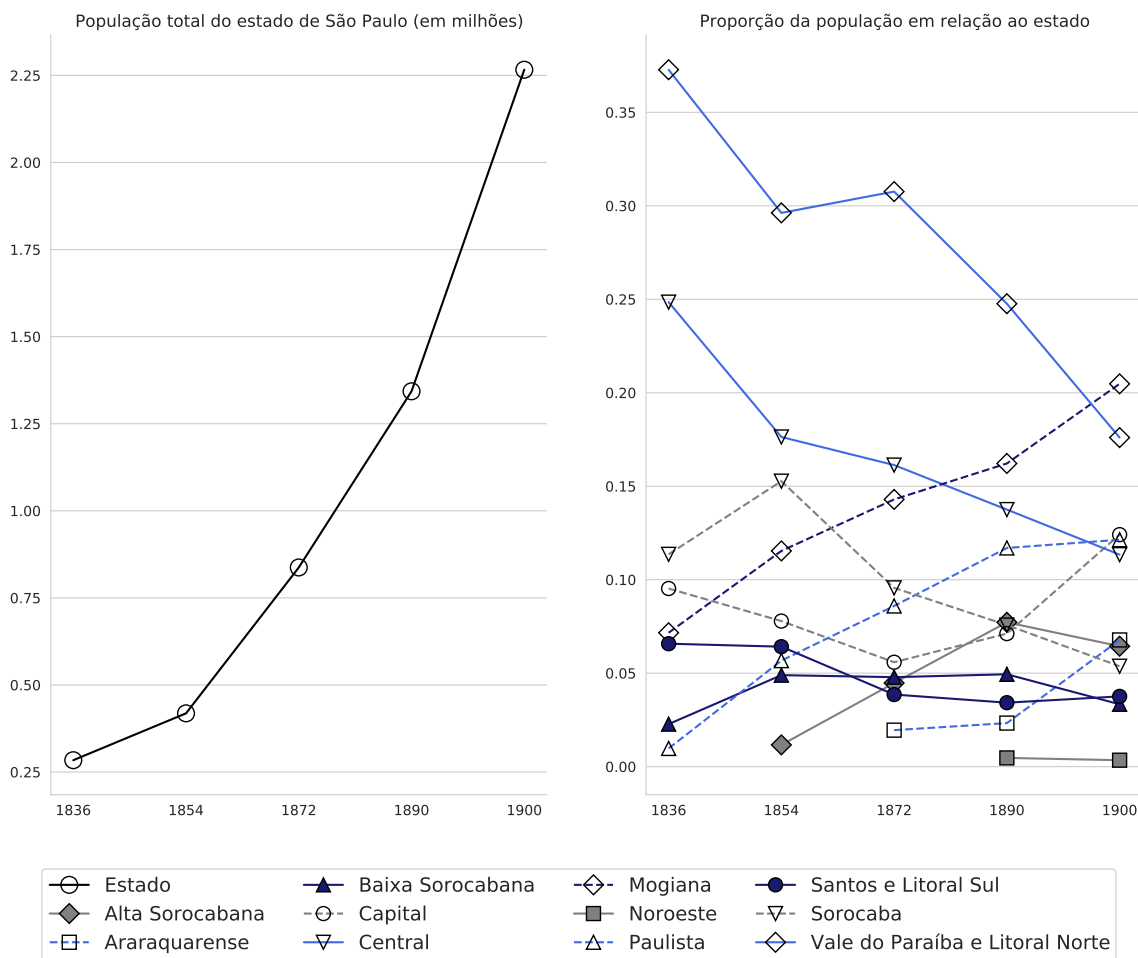


Figura 1.2: Crescimento demográfico e mudança na distribuição espacial da população, São Paulo, 1836-1900

Fonte: Recenseamentos de 1836, 1854, 1872, 1886 e 1900.

ano. A região da Paulista, por outro lado, apresentou a maior taxa de crescimento demográfico no período e viu sua população crescer à taxa de 7,5% ao ano. Esse crescimento teve o efeito de transformar a região em questão, que era a região menos populosa do estado em 1836, na quarta mais populosa em 1900.¹⁵

A terceira dimensão dos efeitos da expansão cafeeira diz respeito à composição por cor da população de São Paulo. No início do século XX, a proporção de

15. O ano de 1836 não inclui regiões de ocupação mais recente, como Araraquarense, Alta Sorocabana e Noroeste

brancos e negros na população foi, em larga medida, resultado dos deslocamentos populacionais para o estado de São Paulo, que tomaram forma ao longo do século XIX. Esses movimentos populacionais incluem o tráfico de escravizados, primeiro pelo tráfico atlântico e, posteriormente, pelo tráfico interprovincial; e a migração de trabalhadores livres, nacionais e estrangeiros. No que diz respeito à escravidão, observa-se ao longo do século XIX uma combinação de aumento da demanda com restrições cada vez maiores à oferta do trabalho cativo. O aumento da demanda por mão de obra escrava teve como consequência um crescimento mais acentuado da população escrava na primeira metade do século XIX, mas os efeitos das crescentes restrições ao tráfico se fizeram sentir após 1850, levando a um ritmo de crescimento mais modesto da população escrava a partir de então. A população escrava cresceu 1,9% ao ano entre 1836 e 1854 e, subsequentemente, desacelerou para um crescimento anual de 1,6% entre 1854 e 1872. Ainda assim, o crescimento da população cativa foi tal que, em 1887, São Paulo, ao lado de Rio de Janeiro e Minas Gerais, concentrava metade da população escrava do país.¹⁶

A migração, por sua vez, dividiu-se entre a entrada de europeus e de brasileiros de outras regiões do país. Ambos movimentos se consolidaram com maior clareza no final do século XIX. Em relação à migração de trabalhadores livres brasileiros, o Relatório da Secretaria de Agricultura de São Paulo de 1928 faz o primeiro registro oficial de entrada de nacionais apenas em 1878, quando migraram 380 brasileiros livres de outros estados, embora a migração de províncias como a de Minas Gerais fosse comum para regiões de São Paulo em épocas anteriores. A imigração europeia, por seu turno, aumentou substancialmente apenas a partir da década de 1880, com seu pico sendo atingido na década de 1890, quando cerca de 140 mil imigrantes entraram no estado em 1895. É provável que muitos dos atrativos da imigração europeia também tiveram o efeito de atrair trabalhadores brasileiros de outros estados.

Em virtude da expansão da economia cafeeira, o impacto regional do tráfico de escravizados e da imigração foi sentido de forma desigual nas diferentes regiões do estado.¹⁷ Regiões onde a produção cafeeira se estabeleceu primeiro

16. Da Costa, *Da senzala*, 67-68; Lago, *Da escravidão*, 81.

17. Samuel H Lowrie, "The Negro Element in the Population of Sao Paulo, a Southernly State of

puderam contar com maior disponibilidade de mão de obra escrava e, assim, a influência africana na população foi maior. Já regiões cujo momento de prosperidade econômica se deu simultaneamente à restrição mais aguda de oferta de escravizados sofreram uma influência demográfica maior da população europeia e dos nacionais de outras partes do país. A heterogeneidade regional do impacto demográfico da escravidão, em particular, pode ser atestada observando os dados da população escrava nos recenseamentos de 1836 e 1854 e, de forma mais precisa, a partir do dados de cor da população nos levantamentos de 1872 e 1890. Em 1836, nas regiões mais populosas do estado, Vale do Paraíba e Litoral Norte e Central, o percentual de escravizados era de 31,1% e 36,3%, respectivamente. Além disso, aproximadamente 70% da população escrava total de São Paulo se concentravam nessas duas regiões. Por outro lado, nas regiões da Baixa Sorocabana e da Paulista, escravizados representavam, respectivamente, apenas 19,2% e 14,4% do total da população local em 1836. Devido à baixa densidade populacional dessas localidades à época, a população escrava da Baixa Sorocabana e da Paulista era apenas 1,9% do total do estado naquele ano.¹⁸

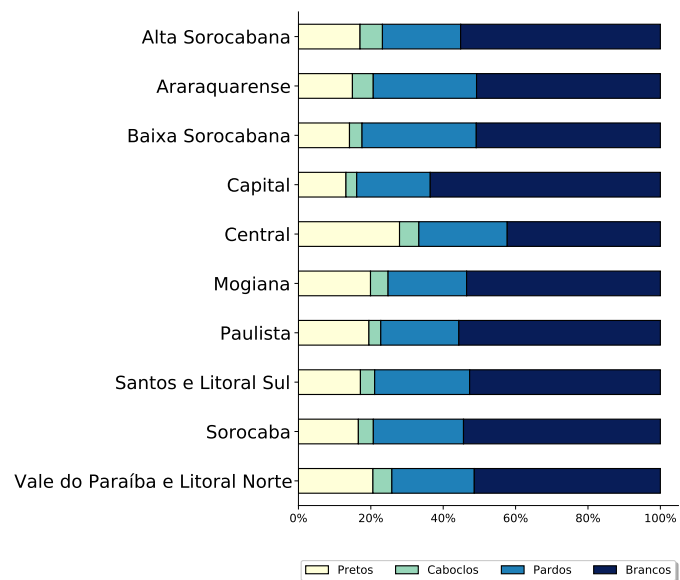
Os dados de cor da população nos censos de 1872 e 1890, apresentados na Figura 1.3, permitem averiguar melhor a relação entre proporção de indivíduos negros e região de São Paulo. É possível identificar três padrões no que se refere à composição por cor das regiões paulistas em fins do século XIX. Primeiro, a região da Capital tornou-se a região mais branca do estado, com mais de 80% dos seus habitantes registrados com essa cor de pele no censo de 1890.¹⁹ A singularidade da situação demográfica da capital, aliada a outras diferenças econômicas e sociais em relação ao restante do estado, sugere que são limitadas as possibilidades de generalizações, para o conjunto do estado, das conclusões de estudos feitos apenas para a cidade de São Paulo.²⁰

Brazil", *Phylon* (1940-1956) 3, n. 4 (1942): 408; Milliet, *Roteiro*, 31,54.

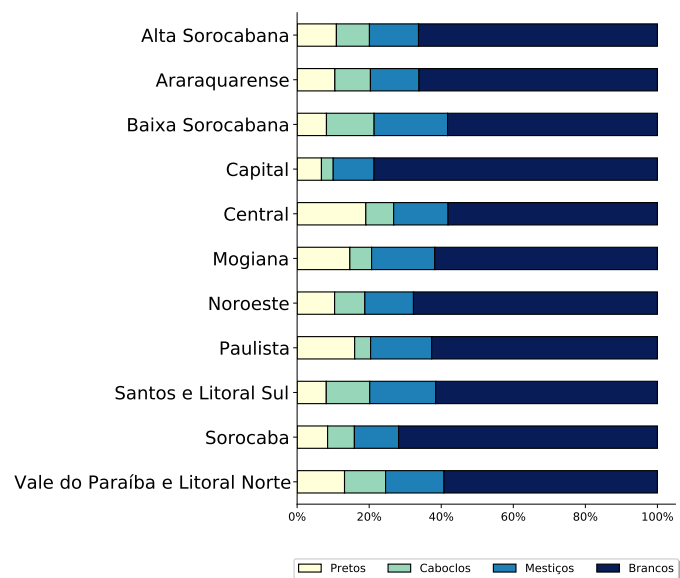
18. Utilizamos a divisão de regiões por áreas tributárias das ferrovias também em períodos anteriores ao seu surgimento na década de 1860, com o propósito de comparar as áreas de São Paulo ao longo do século XIX com um critério único de divisão geográfica.

19. Fernandes, *A integração*, 7.

20. Apesar de se referirem em determinadas passagens ao trabalho do negro no interior no pós-abolição —sugerindo uma possível correspondência entre os padrões observados na capital e o restante do estado —, Florestan Fernandes e George Andrews apresentam dados quantitativos



1872



1890

Figura 1.3: População por cor e região, São Paulo, 1872/1890

Fontes: Censos nacionais de 1872 e 1890.

que referem-se basicamente à capital. Ver Andrews, *Negros e Brancos*, 110-111 e Fernandes, *A integração*, Cap. 1-3

O segundo padrão ocorre nas regiões da fronteira agrícola do estado, como Noroeste e Araraquarense, que concentraram uma proporção expressiva de população branca, mas com percentual menor do que o observado na Capital. Por fim, encontram-se as regiões de ocupação antiga, como Santos e Litoral Sul, o Vale do Paraíba e Litoral Norte e regiões intermediárias do ponto de vista da ocupação do território, como Mogiana e Paulista (Figura 1.3).

O fato de em 1890 a proporção de negros nas regiões da Mogiana (38,2%) e da Paulista (37,3%) ser similar às regiões de Santos e Litoral Sul (38,4%) e do Vale do Paraíba e Litoral Norte (40,7%) está relacionado à realocação de parte da mão de obra escrava entre as regiões do estado no período escravista e à imigração de brasileiros para o Oeste paulista no final do século XIX.²¹ A migração de nacionais, embora em número muito inferior ao da imigração europeia, parece ter seguido um padrão regional similar ao observado por essa última. Como mostra a Tabela 1.1, Mogiana, a região que mais atraiu imigrantes europeus, também foi a que mais recebeu trabalhadores nacionais, absorvendo 26,1% do total de brasileiros que entraram no estado de São Paulo entre 1890 e 1910. A região Araraquarense também foi uma grande receptora de nacionais e estrangeiros (12,6% e 13,3%, respectivamente), embora a correspondência aqui seja menor, uma vez que esse foi o terceiro destino mais comum para estrangeiros e o quarto para brasileiros.²²

A existência de similaridades entre a entrada de estrangeiros e nacionais em São Paulo, contudo, não deve ofuscar diferenças importantes nos dois movimentos migratórios e eventuais omissões decorrentes de observar-se apenas a imigração bruta. Os percentuais da Tabela 1.1 foram calculados a partir de dados de imigração total e, portanto, desconsideram as saídas de população que ocorreram no mesmo período. Assim, localidades com elevada densidade populacional, mas

21. Lowrie, "The Negro", 405.

22. A possibilidade de que houvesse grande mobilidade de trabalhadores nacionais na agricultura paulista já foi aventada por outros pesquisadores e é consistente com a evidência internacional. Para o caso brasileiro, ver Holloway, *Imigrantes*, 107; Mattos, *Das cores*, 49-65; Maria Lúcia Lamounier, "Between slavery and free labour: experiments with free labour and patterns of slave emancipation in Brazil and Cuba c. 1830-1888. 1993" (tese de doutorado, London School of Economics e Political Science, 1993), 183-198. A mobilidade de trabalhadores também era bastante elevada no Sul dos EUA no pós-emancipação. Ver Gavin Wright, *Old South, new South: Revolutions in the southern economy since the Civil War* (Basic Books, 1986), 64-70.

Tabela 1.1: Percentual de nacionais e imigrantes europeus por região de destino, São Paulo, 1890-1910

	Brasileiros	Europeus
Mogiana	26,1	28,8
Paulista	8,7	23,5
Araraquarense	12,6	13,3
Central	3,7	12,1
Capital	2,0	11,0
Alta Sorocabana	14,0	7,2
Noroeste	22,6	1,2
Sorocaba	5,5	0,8
Vale do Paraíba e Litoral Norte	0,1	0,8
Baixa Sorocabana	4,4	0,6
Santos e Litoral Sul	0,2	0,6

Fonte: Valores calculados a partir dos dados dos livros da Hospedaria dos Imigrantes do Estado de São Paulo. Ver “De braços para a lavoura a proprietários rurais? Imigrantes e o acesso à terra em São Paulo, 1886-1920” (Tese de doutorado, Universidade de São Paulo, 2021).

Nota: O percentual foi obtido a partir do total da coluna. Assim, por exemplo, o valor na primeira célula significa que 26,1% dos brasileiros que imigraram para São Paulo entre 1890 e 1910 tiveram como destino a região da Mogiana.

baixa taxa de migração de nacionais, como as regiões da Capital, Santos e Litoral Sul e Vale do Paraíba e Litoral Norte, podem ter de fato testemunhado uma saída líquida de brasileiros. Além disso, a existência de regiões que atraíram, simultaneamente, nacionais e estrangeiros não impede que se observem diferenças importantes nos destinos dos dois grupos de imigrantes. O município de Botucatu, que faz parte da região Noroeste, recebeu uma proporção de brasileiros muito maior do que aquela observada entre estrangeiros. Por outro lado, a Capital foi um destino muito mais comum para europeus do que para nacionais, absorvendo mais de 10% dos indivíduos do primeiro grupo.

De todo modo, os dados da Tabela 1.1 sugerem que o padrão migratório de nacionais imaginado por Florestan Fernandes, vale dizer, o de que a migração de nacionais teve como destino preferencial regiões de baixa produtividade como forma de evitar a competição com o trabalhador estrangeiro, não encontra res-

paldo nos dados da Hospedaria dos Imigrantes.²³ De fato, a proposição de que a migração de nacionais se dirigiu para regiões de menor produtividade marginal do trabalho faz pouco sentido econômico. Outros autores já haviam notado a tendência das novas regiões de atraírem mão de obra nacional e, assim, terem composição populacional relativamente balanceada entre negros e brancos.²⁴ Na seção 1.2.1, mostra-se que os dados demográficos da população branca e negra de São Carlos, município pertencente à região da Paulista, são consistentes com uma migração de nacionais na década de 1880.²⁵

A migração de nacionais para o Oeste paulista teve, possivelmente, implicações que vão muito além da composição da população local por cor da pele. Na verdade, é provável que esse movimento populacional afetasse variáveis demográficas importantes da população negra nas regiões em questão. A continuidade da análise, com o recorte de cor para o período posterior a 1890, fica comprometida dada a ausência dessa informação nos recenseamentos de 1900 e 1920 e da inexistência do recenseamento de 1910. Na seção seguinte, utiliza-se o censo demográfico realizado pelo município de São Carlos no ano de 1907, que possui informação da cor dos indivíduos e permite examinar a estrutura demográfica da população negra no pós-abolição, bem como suas principais características.

1.2 População e economia de São Carlos no início do século XX

Como um município pertencente à região da Paulista, São Carlos é um caso representativo dos eventos econômicos e demográficos narrados anteriormente para o

23. Fernandes, *A integração*, 20.

24. Milliet, *Roteiro*, 54. Embora na passagem o autor esteja referindo-se à migração de nacionais que tomou forma *após* a imigração estrangeira, os dados da Hospedaria sugerem que há razão para acreditar que a composição étnica do Oeste paulista tenha origem mais remota.

25. Estudando o impacto das ferrovias sobre a integração dos mercados de produtos agrícolas, Summerhill mostra que o coeficiente de variação dos preços do café caiu de 0.27 para 0.14 entre 1854 e 1906 em São Paulo. É provável que a redução do custo de transporte não tenha se limitado a integrar os mercados de bens de consumo final. Um menor custo de transporte também teria o efeito de integrar o mercado de trabalho no estado e, por consequência, reduzir a heterogeneidade no custo do trabalho nas diferentes regiões paulistas. Ver Summerhill, *Order*, 62

conjunto de São Paulo. No início do século XX, a maior parte do território de São Carlos encontrava-se na zona rural do município. A tabela de informações agregadas do censo municipal de 1907 mostra que a unidade territorial “Cidade São Carlos”, que corresponde ao núcleo urbano do município, respondia por 20% da população total.²⁶ Ou seja, 80% da população de São Carlos em 1907 distribuíam-se entre os seus 24 bairros da zona rural. Em cada um desses bairros e, em alguns casos, confundindo-se com eles, encontravam-se as fazendas dos municípios. Diferentemente do que ocorre com os dados dos bairros, o censo municipal não traz informação agregada por fazenda da região. De todo modo, a informação está disponível, ainda que de forma incompleta, no nível de indivíduos e proprietários, o que permitiu a identificação de 50 fazendas. Algumas fazendas confundiam-se com bairros rurais, pois a existência de nomes idênticos de bairros e fazendas sugere que, para esses casos, possivelmente a propriedade rural correspondia à designação recebida pela subunidade territorial do município.²⁷

No conjunto das fazendas concentrava-se a maior parte da atividade econômica de São Carlos no início do século XX, com ampla predominância da lavoura cafeeira. Ainda assim, a dominância do café não deve ter sido tão grande como o registrado no Anuário Estatístico do Estado de São Paulo de 1900, em que são detalhados dados sobre a produção agrícola, indústria extrativa e produção zootécnica para todos os municípios paulistas.²⁸ No Anuário, houve registro de informações apenas sobre produção de cachaça e café para o município de São Carlos. Naturalmente, a ausência de estatísticas de produção de outras lavouras não implica sua inexistência e provavelmente houve subregistro de outras atividades produtivas. A ausência de dados de produção agrícola e de outras ativida-

26. São Carlos, Recenseamento da população de S. Carlos do Pinhal no anno de 1907, de accordo com a lei n. 126 votada pela Camara Municipal em sessão de 4 de fevereiro de 1907 (São Carlos: Manuscrito, 1907). Uma análise descritiva do censo municipal de São Carlos de 1907 também pode ser encontrada em trabalhos anteriores, como Truzzi e Bassanezi, “População, grupos étnico-raciais e economia cafeeira: São Carlos, 1907” e Karl Monsma, “Conflito simbólico e violência interétnica: europeus e negros no oeste paulista, 1888-1914”, *História em revista* 10, n. 10 (2004)

27. Os seis casos em que o nome de um bairro também é o nome de uma fazenda são: Can, Capão Preto, Jacaré, Monjolinho, Pinhal, Água Quente.

28. Piza, Antonio de Toledo. *Relatorio do anno de 1900 apresentado em 13 de janeiro de 1902 ao cidadão Dr. Bento Pereira Bueno*, Secretario de Estado dos Negocios do Interior e da Justiça. São Paulo: Typographia do Diário Official 1903, 550-551.

des econômicas na cidade de São Carlos contrasta com a diversidade produtiva apresentada pela maioria dos municípios paulistas retratados no Anuário, para os quais são detalhados números referentes à produção de açúcar, feijão, milho, tabaco, além de informações de extração de madeira e carvão e da criação de gado.

Após a chegada da ferrovia em 1884, São Carlos alcançou certo protagonismo na produção cafeeira.²⁹ Segundo o Anuário de 1900, a produção de café em São Carlos alcançou a marca de 1.355.950 arrobas naquele ano, o que tornou o município o quarto maior produtor do estado. Esse nível de produção ainda aumentaria substancialmente até alcançar o seu auge em 1906, quando foram produzidas 2.214.550 arrobas de café em São Carlos (ver Figura 1.4).

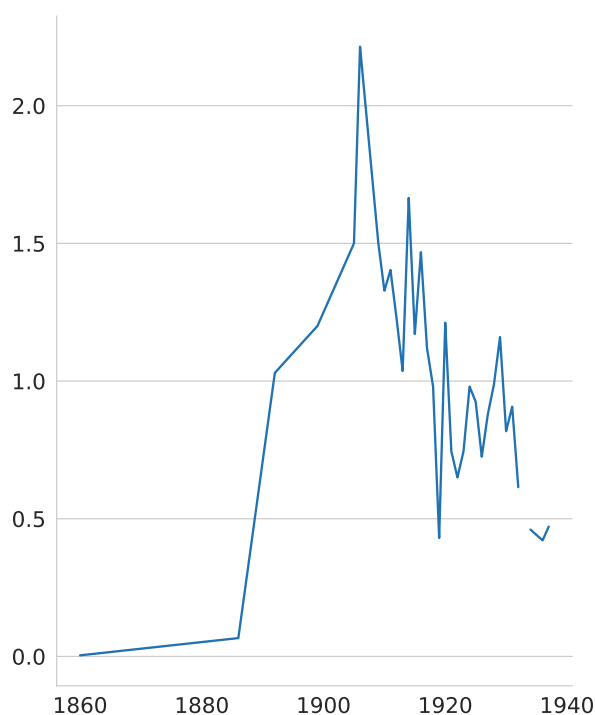


Figura 1.4: Produção cafeeira, em milhões de arrobas, São Carlos, 1860-1937

Fonte: Oswaldo Mário Serra Truzzi, *Café e indústria: São Carlos: 1850-1950* (São Carlos: Imprensa Oficial, 2007), Tabela 1, Cap.3, p.95

29. Oswaldo Mário Serra Truzzi, *Café e indústria: São Carlos: 1850-1950* (São Carlos: Imprensa Oficial, 2007), 94-95.

A expansão da lavoura cafeeira trouxe rápido crescimento demográfico do município. Se em 1872 o município contava com 6.897 habitantes, em 1900 o censo nacional indicou uma população de 55.729 pessoas. Este total aparece como sendo 38.641 habitantes em 1907, segundo o censo municipal realizado neste ano. Em que pese a imprecisão e provável inconsistência de critérios entre os levantamentos, os dados mostram um crescimento anual de 4,1% entre 1872 e 1907. Esse crescimento demográfico parece ter coincidido com uma mudança na composição de cor da população. Segundo a classificação adotada no Recenseamento de 1872, São Carlos já era um município (embora por pequena margem) majoritariamente branco, com 50,6% da sua população registrada com essa cor de pele. A outra metade da população dividia-se em pardos, pretos e caboclos, com predominância dos primeiros, que representavam 27,4% do total. Já em 1907, a população branca constituía 87,6%, tornando o percentual de negros bastante inferior ao observado em outras regiões do Brasil e de outras áreas de *plantations* nas Américas.³⁰

É provável que a imigração europeia tenha sido responsável por grande parte do aumento relativo da população branca, mas inconsistências nos critérios de classificação de cor podem ter levado à superestimação da população branca.³¹ Entre 1884 e 1907, chegaram 20.640 imigrantes em São Carlos, em sua grande maioria vindos da Itália.³² Contudo, como visto na seção anterior, o dinamismo econômico do Oeste paulista não se limitou a servir de imã para o imigrante europeu, pois atraiu também o trabalhador nacional.³³ Além disso, a proximidade da Abolição parece ter se constituído em obstáculo modesto para o crescimento da

30. Para comparação, no sul dos EUA, região em que a população negra é historicamente minoritária, o percentual de negros era de 29,8% em 1910. Ver Campbell Gibson e Kay Jung, *Historical census statistics on population totals by race, 1790 to 1990, and by Hispanic origin, 1790 to 1990, for the United States, regions, divisions, and states* (US Census Bureau Washington, DC, 2002), 22

31. É possível, por exemplo, que em alguns momentos a cor registrada refletia mais a posição social do que pigmentação da pele. Ver Mattos, *Das cores*, Cap. 5

32. Truzzi, *Café e indústria*, 66.

33. É provável que os dados de imigração de brasileiros da Hospedaria de Imigrantes subestimem o total de imigrantes nacionais que se dirigiram à cidade de São Carlos. Isso porque, presumivelmente, a informação assimétrica sobre demanda por trabalho era um problema menor para trabalhadores nacionais —principalmente entre os provenientes de outras regiões paulistas—do que para imigrantes europeus. Dessa forma, a passagem pela Hospedaria, que em parte servia como uma entidade alocadora de trabalho, pode ter sido uma intermediação dispensada por muitos brasileiros.

população escrava local. Entre 1872 e 1887, a população escrava de São Carlos aumentou significativamente de tamanho, passando de 1.568, no início do período, e alcançando 2.982 em 1887.³⁴ Claramente, esses dois últimos movimentos populacionais terão repercussões para a população negra local. Na seção seguinte, essas repercussões serão discutidas ao lado de dados demográficos detalhados da população negra de São Carlos em 1907.

1.2.1 A população negra de São Carlos após a Abolição

O grupo demográfico objeto de análise nesta parte é a população de ex-escravos e descendentes de São Carlos, conforme registrada em 1907, que ao longo do trabalho será chamada simplesmente de *negros*. Assim, os indivíduos classificados pelo censo municipal de São Carlos de 1907 como *pretos* ou *mulatos* serão denominados negros quando analisados em conjunto. Não obstante, a agregação, embora utilizada em alguns momentos, será evitada sempre que possível. Acompanha-se, assim, o entendimento de alguns autores de que a distinção entre ascendência africana integral e mista nos levantamentos demográficos latino-americanos é um indicador de que, para essas sociedades, tal identificação é socialmente relevante.³⁵

A partir da definição do grupo social que será objeto de estudo, dois outros grupos de referência serão utilizados para comparação com os indicadores da população negra. O primeiro grupo é composto por “brasileiros brancos”. Nesse grupo incluem-se famílias luso-brasileiras pertencentes à elite local, como por exemplo os Arruda Botelho e os Camargo Penteado, cujas raízes estão na concessão de sesmarias feitas no final do século XVIII.³⁶ Além das famílias tradicionais, o grupo de brancos brasileiros inclui outros proprietários de terra já estabelecidos na região nas décadas que precederam o fim da escravidão e a imigração europeia, bem como os descendentes da população branca pobre livre

34. Relatório apresentado à Assembléa Legislativa Provincial de São Paulo pelo presidente da província Barão do Parnahyba no dia 17 de janeiro de 1887. São Paulo: Typographia a vapor de Jorge Seckler & Comp. São Paulo, 1887

35. George Reid Andrews, *Afro-Latin America: Black Lives, 1600-2000* (Harvard University Press, 2016), 5.

36. Truzzi, *Café e indústria*, 32-34.

local. Passados mais de cem anos das doações de terra, brasileiros brancos, em que pese sua heterogeneidade, constituíam o grupo populacional com os indicadores mais elevados de riqueza e capital humano e, portanto, servem como um termo de comparação para avaliar a situação socioeconômica de ex-escravos e descendentes. O grupo “imigrantes” constitui outra base de comparação, em virtude de seu peso demográfico e da ampla literatura disponível a respeito de suas condições socioeconômicas e perspectivas de mobilidade social. Como um tema central desta tese é avaliar essas dimensões para a população de ex-escravos e descendentes, a comparação com os imigrantes será também um elemento importante de análise.

Por construção, os imigrantes são um grupo formado exclusivamente por brancos. Apesar de existirem estrangeiros no censo municipal de São Carlos de 1907 que foram classificados como pretos ou mulatos, essa ocorrência possivelmente se explica por erro de preenchimento ou presença de ex-escravos africanos. Esse parece ter sido o caso de três indivíduos com suposta idade superior a 100 anos, cuja nacionalidade foi anotada como “Outra”. Outros casos são em número muito pequeno comparado ao tamanho da nacionalidade a que pertencem, indicando possível erro de preenchimento. Um exemplo são dois italianos, em um universo de 11.305 de pessoas dessa nacionalidade, classificados como pretos no censo de São Carlos. Optou-se por não considerar esses casos no cômputo dos indicadores socioeconômicos analisados. Essa decisão levou à eliminação de quatro observações em um universo de 38.641 indivíduos.

É importante notar que existe uma dificuldade importante associada à separação dos dois grupos de referência, isto é, brasileiros brancos e imigrantes. O censo de São Carlos de 1907 adotou como critério classificar filhos de imigrantes nascidos no Brasil como brasileiros. Dessa forma, a simples tabulação dos dados indica que a população branca brasileira é majoritária, como indica o primeiro painel da Figura 1.5. Fica claro que há superestimação da população brasileira branca por esse critério quando se avalia a distribuição da população entre os grupos por faixa etária, como ilustrado no segundo painel da Figura 1.5. A partir da faixa etária que inclui indivíduos com idade entre 20 e 25 anos, portanto adultos com nacionalidade de origem, há evidente predominância de imigrantes que

superam 60% da população total.

Seria possível, a partir dos dados do censo de 1907, construir uma variável que considerasse a nacionalidade de crianças como sendo equivalente à nacionalidade do chefe da família. Esse procedimento, contudo, apresenta limitações, visto que a informação de situação parental, ou seja, a variável que indica se o indivíduo é chefe de família ou filho(a), não está disponível para todos os indivíduos. Por essa razão, a variável nacionalidade será usada da forma como definida no censo, exceto quando a comparação de indicadores socioeconômicos envolver crianças, como educação e trabalho infantil, como será visto mais adiante.

Além da motivação analítica, a separação dos três grupos encontra respaldo nos dados do censo de São Carlos. A Tabela 1.2, que mostra a presença de casamentos intra e inter-grupos, indica que brasileiros brancos, imigrantes e negros constituíam, em larga medida, grupos demográficos separados. Cerca de 93% dos casamentos se davam intra-grupos, sendo que nos 7% restantes, 3/4 se davam entre brasileiros brancos e imigrantes. Em outras palavras, casamentos entre os grupos analisados eram raros, ainda mais raros eram os casamentos de negros com brasileiros brancos ou imigrantes.³⁷

Tabela 1.2: Casamentos inter e intra-grupos, São Carlos, 1907

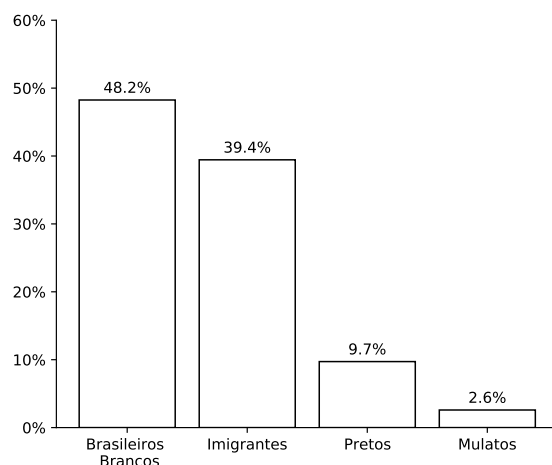
Homem/Mulher	Brasileiros brancos	Imigrantes	Negros
Brasileiros brancos	875	85	16
Imigrantes	222	3704	13
Negros	36	25	627

Fonte: São Carlos, Recenseamento da população de S. Carlos do Pinhal no ano de 1907, de acordo com a lei n. 126 votada pela Câmara Municipal em sessão de 4 de fevereiro de 1907 (São Carlos: Manuscrito, 1907). 8 v.

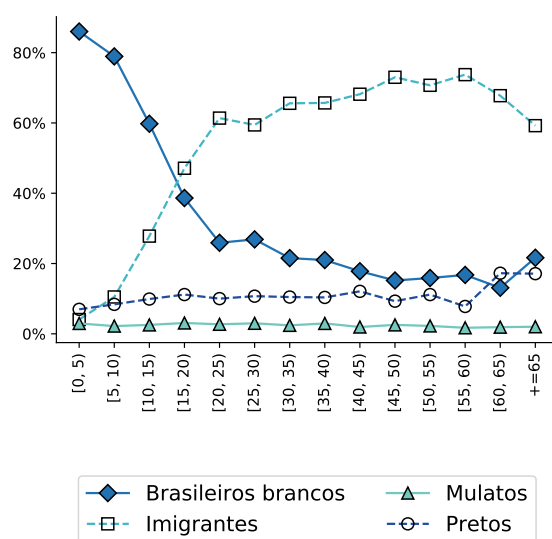
Notas: Casamentos intra-grupos na diagonal principal. Considerou-se um casal a sequência de homem e mulher no censo em que o homem estava registrado como “chefe” e a mulher como “dependente-esposa”.

A distinção socioeconômica entre os três grupos também é clara quando se

37. Embora, sob essa métrica, imigrantes tenham endogamia menor de que brasileiros negros, a tendência de imigrantes se casarem majoritariamente entre si se manteve pelo menos até o final da década de 1920 em São Carlos. Ver Oswaldo Mário Serra Truzzi, “Padrões de nupcialidade na economia cafeeira de São Paulo (1860-1930)”, *Revista Brasileira de Estudos de População* 29, n. 1 (2012): 169–189



População total por grupo



Distribuição por idade e grupo

Figura 1.5: Distribuição da população por grupo demográfico e idade, São Carlos, 1907

Fonte: São Carlos, Recenseamento de 1907.

analisam os dados de educação e acesso à terra. Em relação à educação, pode-se notar ao menos três fatos a partir da observação da Tabela 1.3. Em primeiro lugar, há clara hierarquia entre os três grupos demográficos analisados no âmbito desta tese, com brasileiros brancos apresentando indicadores melhores dos que

os de imigrantes, enquanto os indicadores desse último grupo superam os de brasileiros negros. Em segundo lugar, observa-se que, para todos os grupos, as mulheres aparecem em posição de desvantagem, com a alfabetização sendo sempre mais comum entre homens. Esse resultado é consistente com o fato de que as taxas de matrícula escolar para mulheres foram mais baixas do que as de homens ao longo do século XIX em São Paulo.³⁸ Além disso, mesmo em países mais desenvolvidos, as taxas de alfabetização de homens ainda eram maiores do que as de mulheres em fins do século XIX, o que pode ajudar a explicar o fato de a alfabetização ser mais comum para homens também entre imigrantes.³⁹ Em terceiro lugar, a Tabela 1.3 demonstra que a alfabetização de mulatos era significativamente maior do que a de pretos. A combinação do efeito de cor da pele e sexo sobre alfabetização faz das mulheres pretas o grupo demográfico com os piores indicadores de alfabetização na população.

A posição relativa dos grupos demográficos muda um pouco quando são analisados os dados de acesso à terra, como pode ser visto na Tabela 1.4. Novamente, observa-se a tendência de brasileiros brancos e imigrantes apresentarem indicadores melhores que os da população negra. No entanto, desta vez o indicador analisado mostra que a vantagem de imigrantes é menos pronunciada em relação a brasileiros negros como um agregado (13,5%). Além disso, quando se comparam imigrantes (14,2%) com o segmento de mulatos (16,8%), estes últimos levam vantagem no acesso à terra em relação aos europeus —um resultado que talvez possa ser considerado surpreendente. Entre brasileiros brancos, proprietários de terra compunham 36% da população adulta masculina. Entre imigrantes, esse percentual era de apenas 14,2%, número um pouco maior do que o observado entre pretos e inferior à taxa registrada entre mulatos.

Tomadas pelo valor de face, as taxas de acesso à terra entre imigrantes e negros

38. Renato Perim Colistete, *O atraso em meio à riqueza: uma história econômica da educação primária em São Paulo, 1835 a 1920* (São Paulo, 2016), 109.

39. Na Inglaterra, um dos países pioneiros na massificação da educação básica, a razão das taxas de alfabetização entre homens e mulheres ainda era maior que 1 em 1870 (80,6/73,2). É provável que na Itália, país de origem da maioria dos imigrantes em São Carlos, essa razão ainda fosse mais alta. Para evolução dos níveis de alfabetização entre homens e mulheres na Inglaterra, ver Michael Sanderson, *Education, economic change and society in England 1780-1870* (Cambridge University Press, 1995), 19.

Tabela 1.3: Taxa de alfabetização maiores de 15 anos (em percentual), São Carlos, 1907

Grupo	Sexo	Alfabetizado
Brasileiros brancos	Feminino	51,0
	Masculino	62,6
Imigrantes	Feminino	21,7
	Masculino	40,3
Mulato	Feminino	13,9
	Masculino	28,4
Preto	Feminino	6,1
	Masculino	14,7

Fonte: São Carlos, Recenseamento de 1907.

Notas: Variável de nacionalidade definida de acordo com a nacionalidade do chefe da família. Como pretos eram maioria da população negra, os indicadores da população negra como todo estão mais próximos dos valores observados para pretos do que para mulatos. Dessa forma, alfabetizados na população negra perfaziam 18,1% entre homens e 8,4% entre mulheres.

Tabela 1.4: Percentual de proprietários de terra, São Carlos, 1907

Grupo	Percentual de proprietários
Brasileiros brancos	36,0
Imigrantes	14,2
Mulato	16,8
Preto	12,7

Fonte: São Carlos, Recenseamento de 1907.

Notas: Dados para a população masculina com idade igual ou maior de 20 anos.

parecem bastante elevadas. Para efeito de comparação, os dados de censo municipal indicam que o acesso à terra de brasileiros negros em São Carlos cerca de 20 anos depois do fim da escravidão era maior do que foi observado nos estados do Sul dos Estados Unidos em 1876, 11 anos depois da Emancipação americana. Para os 10 estados americanos para os quais existem dados disponíveis, o percentual de negros proprietários de terra variou de 4% a 8%.⁴⁰ Desta forma, os percentuais de acesso à terra de negros em São Carlos são realmente expressivos e podem ser entendidos como uma evidência relevante da posição alcançada por ex-escravos

40. Higgs, *Competition*, 51-52.

e descendentes após a abolição no município.

Ao mesmo tempo, há ambiguidades nas definições contidas no censo de 1907 que podem relativizar esses resultados, embora seja difícil dizer em que medida. Como já notado por Karl Monsma, a definição de propriedade no censo parece ser muito ampla, dado que o total de proprietários listados supera por larga medida o número obtido na Estatística Agrícola de 1905.⁴¹ Uma série de fatores podem ter contribuído para a incompatibilidade da informação sobre propriedade nos dois levantamentos, tais como a situação legal da posse de terra e a inclusão de ativos distintos da propriedade rural. No Capítulo 4, analisa-se com maior profundidade a questão fundiária e os problemas com as definições da posse da terra, utilizando os dados da Estatística Agrícola de 1905.

Outra dimensão em que é possível caracterizar os grupos demográficos analisados é a composição de sexo-idade, ilustrada pelas pirâmides etárias de cada um dos grupos na Figura 1.6. É possível notar que as três distribuições são caracterizadas por uma base larga em relação ao topo. Essa forma da pirâmide etária é típica de populações que estão atravessando a primeira fase da transição demográfica, marcada por taxas de mortalidade infantil relativamente menores do que as observadas no período pré-transição e elevadas taxas de fertilidade.⁴²

Outra tendência observada na composição de sexo-idade das subpopulações sob consideração é a elevada razão sexual, número de homens para cada 100 mulheres, para negros com idade superior a 40 anos, ilustrada na Figura 1.7. Em parte, essa elevada razão sexual pode ser explicada pela migração de trabalhadores nacionais para a fronteira agrícola nas décadas anteriores, mencionada na segunda seção deste capítulo. Um fato que corrobora essa conjectura é que na capital do estado ocorreu precisamente o inverso, já que entre o censo nacional

41. Monsma, *A reprodução*, 307.

42. Historicamente, a demografia dos países ao longo do século XIX e XX é caracterizada por um movimento de queda nas taxas de fertilidade e mortalidade. Esse movimento, que é conhecido como transição demográfica, ocorre de forma não simultânea, de modo que, em geral, as taxas de mortalidade convergem para patamares baixos *primeiro* que as taxas de fertilidade. No breve período de vigência de elevadas taxas de fertilidade com queda gradual das taxas de mortalidade, os países passam pelo seu período de crescimento demográfico mais agudo. Jennifer Hickes Lundquist, Douglas L. Anderton e David Yaukey, *Demography: the study of human population* (Waveland Press, 2014), 145.

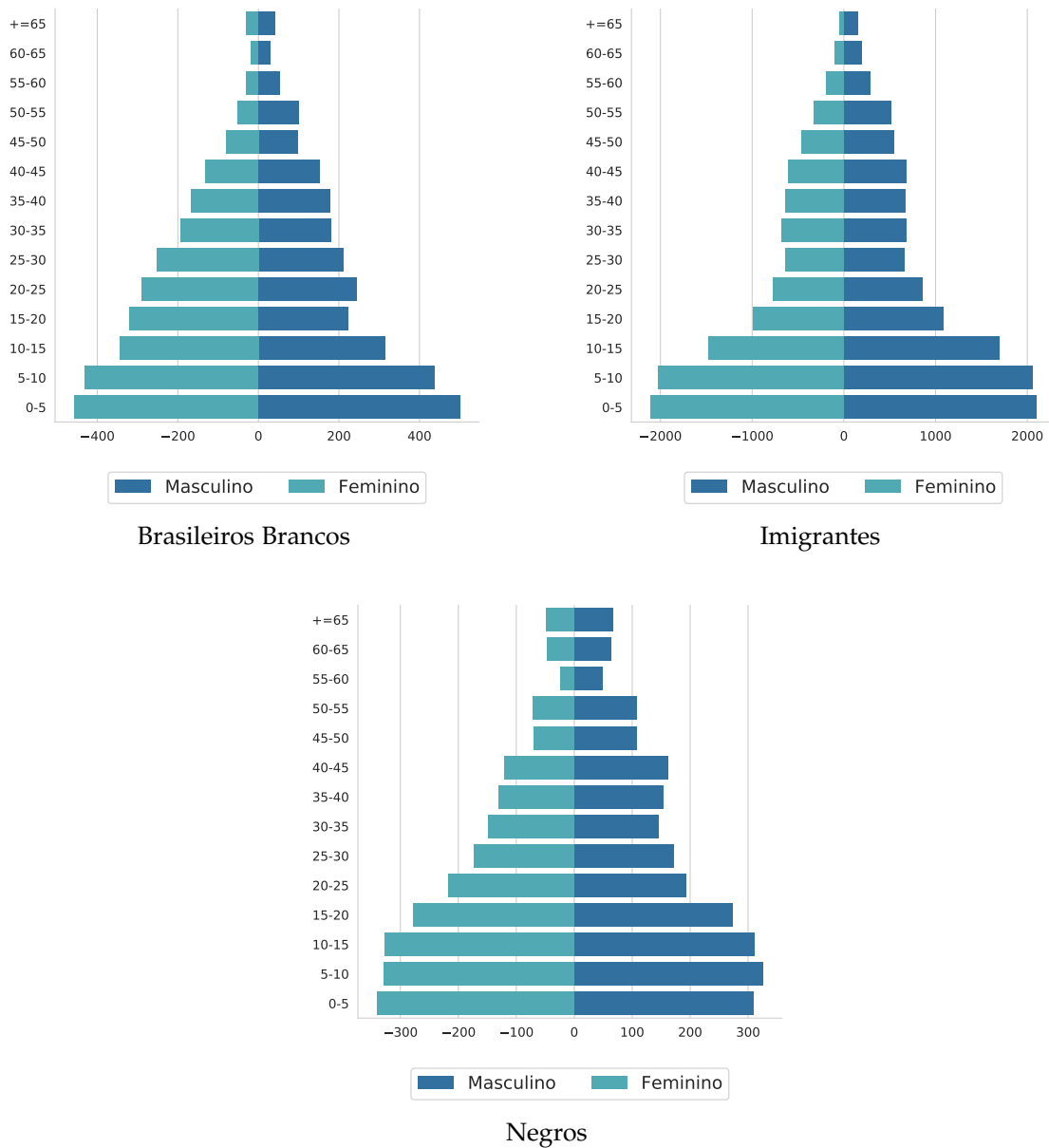


Figura 1.6: Pirâmide etária de brasileiros brancos, imigrantes e negros

Fontes: São Carlos, Recenseamento de 1907.

de 1890 e o censo municipal realizado em 1893 a população de mulheres negras na cidade de São Paulo cresceu 44,1%, contra um crescimento de apenas 13% da

população negra masculina.⁴³ Ou seja, parece haver evidência de que a migração de negros da capital em direção ao interior, no final do século XIX, foi caracterizada por um grande desequilíbrio de gênero. Como sustenta Florestan Fernandes, o fato de que regiões urbanas sejam caracterizadas por elevada demanda por serviços domésticos pode ter contribuído para retenção de amplos contingentes da população negra feminina na Capital.⁴⁴

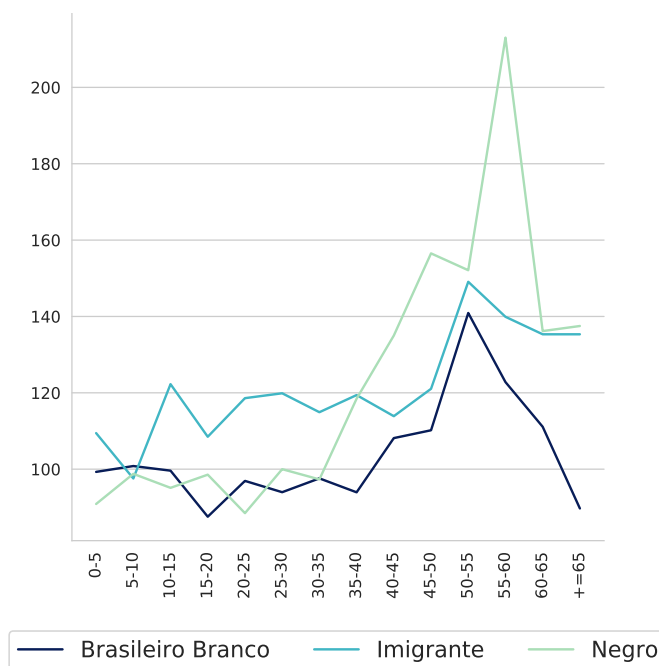


Figura 1.7: Razão sexual por grupo etário, São Carlos, 1907

Fonte: São Carlos, Recenseamento de 1907.

Notas: A razão sexual é dada pela razão entre o número de homens e o número de mulheres multiplicada por 100.

É plausível também que outros fatores podem ter concorrido para o elevado número de homens por mulher entre negros com mais de 40 anos em São Carlos, como, por exemplo, a projeção de padrões demográficos da escravidão, que era caracterizada por um elevado número de homens por mulher, e uma maior mortalidade relativa de mulheres negras no parto.⁴⁵ Cada um desses fatores pode

43. Fernandes, *A integração*, 42.

44. *Ibid.*, 42-43.

45. Monsma, *A reprodução*, 104.

ter contribuído para o fenômeno observado com diferentes graus de intensidade. Para o presente estudo, é suficiente notar que a elevada razão entre homens e mulheres na população relaciona-se com duas hipóteses sobre a evolução socioeconômica da população negra local. Em primeiro lugar, a possível migração de nacionais do sexo masculino para o município de São Carlos sugere uma elevada demanda por trabalho que parece não ter sido inteiramente atendida pela imigração europeia, indicando escassez relativa de trabalho. Essa escassez teria o efeito de aumentar o poder de barganha de trabalhadores nacionais e permitiria condições contratuais melhores do que as alcançadas em outras regiões do estado de São Paulo no pós-abolição.

Em segundo lugar, a combinação de elevada proporção de homens em relação a mulheres com altas taxas de mortalidade materna e infantil parece indicar que a população negra brasileira residente em São Carlos estava vivenciando um momento demográfico distinto daquele experimentado pela população branca. Mesmo que as três sub-populações estivessem crescendo em termos absolutos, como resultado de uma queda prévia da mortalidade infantil, é possível que a mortalidade infantil ainda fosse relativamente mais elevada para negros. Um dado que reforça essa conjectura é a diferença de fertilidade entre as mulheres de cada subgrupo populacional. Enquanto o número de filhos por mulher entre brasileiros brancos e imigrantes era, respectivamente, de 3,16 e 3,54, essa taxa era de apenas 2,5 para mulheres negras. A Figura 1.8 compara a distribuição do número de filhos para cada um dos subgrupos populacionais.

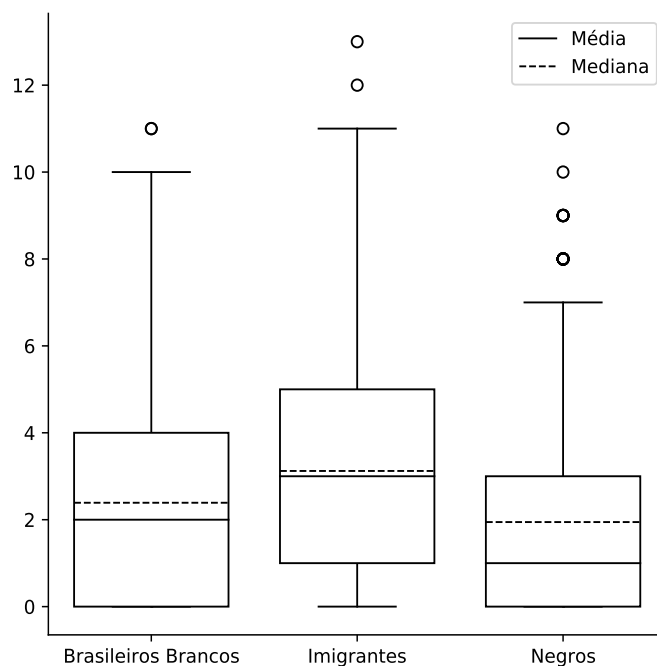


Figura 1.8: Tamanho da família por grupo populacional, São Carlos, 1907

Fonte: São Carlos, Recenseamento de 1907.

Notas: Classificação dos filhos foi feita utilizando-se o critério *jus sanguinis*. Um filho de italiano, por exemplo, foi classificado no grupo Imigrantes, ainda que tenha nascido no Brasil.

As diferenças demográficas entre a população negra e branca podem ter tido efeitos não desprezíveis para o sucesso econômico da população negra no longo prazo. Homens negros que se beneficiaram do dinamismo da economia cafeeira no Oeste paulista podem ter enfrentado possibilidades limitadas de transferir seus ganhos materiais (por exemplo, na propriedade da terra) para gerações posteriores, diante de dificuldades de constituição de vínculos familiares. Além disso, a baixa fertilidade de mulheres negras tinha como resultado a formação de famílias relativamente pequenas. Como já sugerido em outros trabalhos, em uma economia agrária, intensiva em trabalho, a baixa fertilidade pode levar a uma renda familiar menor devido ao menor número de braços para a lavoura, diminuindo as possibilidades de acumulação de capital no âmbito familiar.⁴⁶

46. Holloway, *Imigrantes*, 214; Louis Albert Gaffre, *Visions du Brésil* (São Paulo: F. Alves & Cia., 1912), 267.

1.3 Conclusões

Este capítulo dedicou-se a uma apresentação geral de informações básicas, mas importantes da economia e demografia do estado de São Paulo e do município de São Carlos para o período compreendido entre o final do século XIX e início do século XX. Foi dada particular atenção aos movimentos migratórios de nacionais e estrangeiros e sua relação com a composição demográfica da população negra de São Carlos no pós-abolição. Argumentou-se, inicialmente, que a existência desses movimentos populacionais é consistente com uma elevada demanda por trabalho no Oeste paulista e teve consequências importantes para a composição demográfica da população negra.

Analisando brevemente os dados relativos à educação e demografia da população negra, vários resultados importantes já foram destacados. A população negra em São Carlos possuía nível de educação, medido pelas taxas de alfabetização, inferior ao observado para brasileiros brancos e imigrantes. Além disso, em cada um dos subgrupos populacionais, mulheres estavam em situação de desvantagem em relação aos homens. Por outro lado, decompondo o agregado de negros em mulatos e pretos, um primeiro resultado importante obtido foi que mulatos apresentaram indicadores de educação melhores que de pretos. A situação socioeconômica relativamente mais favorável de mulatos também se reflete nos indicadores de propriedade da terra, já que o percentual de proprietários de terra entre mulatos chega a superar o observado para imigrantes.

Também sugeriu-se que a população negra passava por um momento na transição demográfica distinto daquele que caracterizava a população de brasileiros brancos e de imigrantes, com taxas de mortalidade infantil ainda elevadas e, por consequência, famílias menores. Esse achado corrobora conjecturas feitas por outros autores que viam nas numerosas famílias imigrantes uma vantagem na acumulação de capital no âmbito familiar.⁴⁷ No próximo capítulo apresenta-se uma análise mais detalhada dos níveis de educação e capital humano das populações negra e branca, nacionais e imigrantes, em São Carlos. Para tanto, serão avaliados os dados de matrícula na educação primária a fim de se inves-

47. Holloway, *Imigrantes*, 214; Gaffre, *Visions du Brésil*, 267.

tigar as condições e prováveis desvantagens enfrentadas por negros no acesso à educação.

Capítulo 2

A população negra e a educação: nível e determinantes da matrícula escolar

Em 1910, o Brasil possuía apenas 30 crianças matriculadas na escola para cada 1.000 habitantes. Essa taxa era bastante inferior àquelas observadas entre países que lideravam as estatísticas de matrícula em educação primária da época, como os Estados Unidos e a Alemanha, que alcançavam taxas de 197 e 164, respectivamente. Mesmo quando comparado aos demais países latino-americanos, o desempenho brasileiro empalidece frente às médias alcançadas pelos países andinos e pelos países do Cone Sul. Entre os primeiros, a média de crianças matriculadas por mil habitantes era de 48 em 1910, enquanto essa mesma taxa era de 95 em média para Argentina, Chile e Uruguai.¹

Não está claro o quanto a desigualdade no acesso à educação entre brancos e negros contribuiu para esse quadro geral. Nos Estados Unidos, por exemplo, a diferença de acesso à educação entre esses dois grupos demográficos ao longo do século XIX, um subproduto das restrições ligadas à educação de escravizados, parece ser a principal explicação para os resultados relativamente ruins dos indicadores americanos de alfabetização no final do século XIX. Enquanto na Alema-

1. Colistete, *O atraso*, 28,32.

nha o percentual de analfabetos era de apenas 0,2% em 1890, os Estados Unidos tinham mais de 13% da sua população iletrada no mesmo ano. Em larga medida, a taxa americana era explicada pelo elevado percentual de analfabetos na população negra (56,8%), na medida em que a taxa de analfabetismo da população branca nativa era de apenas 6,2%.² Esse diferencial, contudo, foi rapidamente corrigido nas primeiras décadas de século XX, com o percentual de alfabetizados entre negros americanos alcançando 70% em 1910.³

Assim como nos EUA, no Brasil também houve constrangimento legal à educação de escravizados. Não sendo considerados cidadãos, crianças e adultos escravizados encontravam-se fora do alcance do conteúdo do inciso XXXII do Capítulo 179 da Constituição de 1824, não sendo elegíveis para instrução primária. Esse dispositivo constitucional também esteve presente na legislação ordinária provincial. Os exemplos incluem as províncias de Minas Gerais (Lei nº 13 de 1835), Rio de Janeiro (Regulamento de 1º de Setembro de 1847) e São Paulo (Regulamento da Instrução Pública de 1869).⁴ Além disso, a exclusão de escravizados da vida escolar prevaleceu mesmo na ausência de previsão legal nas províncias, em função das atitudes sociais dominantes em relação às pessoas mantidas sob cativeiro.

Menos claro, contudo, é o que aconteceu com o acesso à educação da população negra no pós-abolição. Neste capítulo, ao estimar taxas de matrícula escolar para a população negra de São Carlos, procura-se lançar luz sobre o tema da educação de ex-escravos e descendentes no período que se seguiu à Lei Áurea. No Capítulo 1, mostrou-se que há evidências de que a prosperidade da economia cafeeira no Oeste paulista criou oportunidades e ganhos materiais para imigran-

2. Colistete, *O atraso*, 37-38.

3. Ver Higgs, *Competition*, 120. De modo geral, ao longo de todo século XX, houve rápida convergência educacional da população negra americana com a população branca. Ver Robert A Margo, *Obama, Katrina, and the persistence of racial inequality*, relatório técnico (National Bureau of Economic Research, 2016), 24-31

4. Sobre essas leis, ver Surya Aaronovich Pombo de Barros, "Negrinhos que por ahi andão: a escolarização da população negra em São Paulo (1870-1920)", 2005, p.81-83. Em geral, as restrições legais à educação eram impostas apenas a escravizados, não a negros livres. Alguns estudos apontam a presença de negros livres em escolas (ver Marcus Vinicius Fonseca, "Pretos, pardos, crioulos e cabras nas escolas mineiras do século XIX." (tese de doutorado, Universidade de São Paulo, 2007); Adriana Maria Paulo Da Silva, "A escola do professor Pretextato dos Passos e Silva: questões a respeito das práticas de escolarização no mundo escravista", *A história da educação dos negros no Brasil*. Niterói: EdUFF, 2016, 141-162)

tes e trabalhadores nacionais nas duas décadas que se seguiram à Abolição, incluindo a população negra. Teria esse ganho material viabilizado investimentos em educação primária por parte dessas populações? Esse tipo de investimento de tempo e recursos também estava ao alcance de ex-escravos e seus descendentes? As taxas de matrícula na educação primária ajudam a responder a essas perguntas.

O cômputo das taxas de matrícula apresentado adiante foi possível mediante o cruzamento dos dados de escolas extraídos de livros de chamada com o censo municipal de São Carlos de 1907.⁵ Os resultados demonstram que a taxa de matrícula da população negra em educação primária não diferia significativamente das taxas observadas para o restante da população nas escolas públicas isoladas. Quando são analisadas todas as escolas —incluindo o grupo escolar do município e escolas particulares —, porém, há pequena vantagem para a população branca. Com efeito, o modelo econométrico estimado mostra que, controlando para variáveis socioeconômicas da família do estudante, uma criança negra tinha em média uma probabilidade de estar na escola que era 10 pontos percentuais menor do que uma criança branca. Os resultados implicam ainda que a baixa alfabetização das mulheres negras no pós-abolição parece ter sido particularmente prejudicial para as chances de matrícula de crianças negras.

Essas questões serão tratadas em 5 seções ao longo deste capítulo. Na seção seguinte, são apresentadas detalhadamente as bases de dados utilizadas. A seção 2.2 resume as estimativas de matrícula escolar para cada um dos grupos demográficos analisados nesta tese, com destaque para os dados de alunos pretos e mulatos. Nas duas partes seguintes, dois exercícios empíricos são implementados com o objetivo de avaliar quais variáveis socioeconômicas das famílias dos alunos tiveram mais influência sobre a probabilidade de matrícula. A seção 2.4 conclui o capítulo, resumindo os principais achados e discutindo suas eventuais implicações.

5. Livros de chamadas, 1906-1909, Livros: E-3910, E-3912, E-3914, E-3916, E-3674, São Paulo, Arquivo Público do Estado de São Paulo.

2.1 Os dados educacionais: livros de chamadas e o censo municipal

Os dados utilizados neste capítulo resultam do cruzamento das informações de livros de chamadas de cinco escolas isoladas primárias públicas de São Carlos na década de 1900 com o censo do município realizado em 1907. Essas escolas estavam localizadas nos seguintes locais: Bairro da Estação, Santa Cruz, Ibaté e Vila Nery. A combinação dos dados das duas fontes permite obter informações demográficas dos alunos e de suas famílias, bem como relacionar essas informações à vida escolar de indivíduos encontrados no censo municipal.

As informações das cinco escolas primárias de São Carlos referem-se a livros de chamadas dos anos letivos de 1906, 1907, 1908 e 1909.⁶ Ao longo dos quatro anos para os quais há dados, as cinco escolas aqui analisadas receberam 456 alunos únicos. O *Anuário de Ensino do Estado de São Paulo* indica que, em 1909, São Carlos possuía 702 alunos matriculados nas escolas estaduais.⁷ Dessa forma, os 456 alunos representariam cerca de 65% de todos os alunos das escolas estaduais no período.⁸ Em parte, a diferença parece ocorrer devido ao fato dos livros de chamadas utilizadas *não* incluírem o grupo escolar “Coronel Paulino Carlos”. Este grupo escolar foi fundado em 1904 e contava naquele ano com 346 alunos.⁹ Se, por um lado, a ausência do grupo escolar é uma limitação dos dados, por outro lado as crianças matriculadas com informação disponível formam um grupo mais homogêneo de filhos e filhas de famílias residentes em bairros rurais e urbanos possivelmente mais distantes da área central da cidade de São Carlos.

Cada livro de chamada contém 50 páginas duplas. Cada uma dessas páginas

6. Os livros de chamadas eram de preenchimento obrigatório por parte dos professores, de acordo com os Capítulos 69 e 70 do Decreto N. 248, de 26 de julho de 1894.

7. Directoria Geral da Instrução Publica, *Anuario do Ensino do Estado de São Paulo, 1909-1910*, Oscar Thompson. São Paulo: Typografia do Diário Oficial, 1910, 150-151.

8. Nas primeiras décadas do regime republicano, as escolas estaduais representavam a maior parte das matrículas em escolas primárias. Em São Carlos, em 1909, o percentual de alunos das escolas estaduais representava 53,4% do total (1315). Sobre a dominância das escolas estaduais na oferta de educação primária, ver Colistete, *O atraso*, p. 222

9. Inspectoria Geral de Ensino, *Anuario do Ensino do Estado de São Paulo, 1907-1908* São Paulo: Typografia Augusto Siqueira & C., 1908, 231-232.

representava um mês de aula e continha espaço para o preenchimento de dados de nome do professor, nome dos alunos, mês, frequência mensal de cada aluno e, em alguns casos, notas de comportamento e desempenho escolar. Como o número de páginas não coincide com o número de meses de aulas e como o número de alunos pode implicar o uso de mais de uma página dupla por mês de atividade escolar, era comum que livros de chamadas incluíssem frações de anos letivos.¹⁰ Essa é a razão pela qual, para alguns anos e algumas escolas, as informações não estão disponíveis para todo o ano letivo. A Tabela 2.1, que resume os dados obtidos com os livros de chamada de cada escola, e a Figura 2.1, que mostra um exemplo de uma página de um livro de chamadas, dão uma ideia mais precisa das informações coletadas.

Para identificar os dados demográficos dos estudantes e de suas famílias, procedeu-se ao cruzamento de nomes nos livros de chamada com os nomes listados no censo municipal de São Carlos de 1907. O cruzamento de informações de fontes históricas nominais apresenta uma série de obstáculos de natureza prática: alterações na grafia dos nomes, mudanças de localidade por parte dos indivíduos e homônimos. Dessa forma, utilizou-se um algoritmo de *matching*, com o qual foi possível identificar 267 estudantes no censo dentro do universo de 456 crianças matriculadas nas escolas. Uma descrição detalhada do algoritmo empregado encontra-se no Apêndice A. Aqui apresenta-se apenas o princípio básico do algoritmo e uma medida que servirá para avaliar a validade dos cruzamentos.

O algoritmo utilizado para a implementação dos cruzamentos de nomes está baseado no conceito de distância de *string*. Uma *string* consiste em qualquer sequência de caracteres que representa uma combinação de letras, números ou símbolos, enquanto *matching* se refere à identificação da ocorrência de uma observação em uma base de dados em outra base. No presente caso, realiza-se um *matching* toda vez que se encontra um estudante registrado nos livros de chamada no censo de São Carlos. A principal medida de distância de *string* utilizada foi a *distância Levenshtein*, que calcula o mínimo de operações (inserções, remoções

10. Um exemplo desse caso é a Escola Feminina do Bairro da Estação, cujo livro de chamada termina em setembro de 1909. A última informação é do mês de setembro de 1909, não porque esse ano letivo acabou de forma prematura, mas porque o número de folhas do livro de chamada terminou antes do fim do ano letivo.

Tabela 2.1: Número de crianças matriculadas por escola e ano, São Carlos, 1906-1909

Escola	Ano	Total	Média mensal	1º mês	Último mês
E.F. Bairro da Estação	1907	56	32,3	Março	Dezembro
	1908	55	34,3	Fevereiro	Dezembro
	1909	52	33,7	Fevereiro	Setembro
E.F. de Santa Cruz	1907	44	29,8	Fevereiro	Dezembro
	1908	51	35,8	Fevereiro	Outubro
	1909	46	32,5	Fevereiro	Novembro
E.F. de Vila Nery	1907	40	28,0	Março	Dezembro
	1908	60	34,1	Janeiro	Dezembro
	1909	58	28,7	Janeiro	Novembro
E.M. de Ibaté	1906	34	28,3	Outubro	Dezembro
	1907	55	31,0	Janeiro	Dezembro
	1908	51	31,6	Janeiro	Dezembro
	1909	41	30,4	Janeiro	Outubro
E.M. de Vila Nery	1907	44	38,8	Julho	Dezembro
	1908	47	39,3	Janeiro	Dezembro
	1909	43	36,5	Janeiro	Julho

Fonte: Livros de chamadas, 1906-1909, Livros: E-3910, E-3912, E-3914, E-3916, E-3674, São Paulo, Arquivo Público do Estado de São Paulo.

Nota: A coluna "Total" corresponde ao número de estudantes únicos que apareceram pelo menos uma vez no livro de chamada no ano em questão.

e permutações) necessárias para transformar uma *string* em outra.¹¹ Quando a distância de Levenshtein é igual a zero, os dois nomes são idênticos. Em todos os outros casos, são necessárias modificações para realizar o *matching*. Em essência, o algoritmo consiste em identificar os potenciais *matchings* de um nome, usar medidas de distância de *string* para ordenar os *matchings* encontrados e aceitar ou rejeitar o melhor *matching* de acordo com um critério pré-estabelecido.

A partir do *matching* realizado entre as duas bases de dados citadas, a próxima seção apresenta estimativas de matrícula escolar para crianças brancas, pretas e mulatas em São Carlos.

11. Vladimir I Levenshtein, "Binary codes capable of correcting deletions, insertions, and reversals", 10, n. 8 (1966): 707-710.

informações escolares sobre os filhos da população adulta que não faz parte do grupo de pais dos estudantes das escolas para as quais se dispõe de dados. É certo, porém, que a maioria dessa população adulta restante tinha filhos que estavam fora da escola, já que no censo populacional existiam 9.631 indivíduos com idade entre 7 e 15 anos. Desta forma, como havia 1.315 alunos nas três redes de ensino (estadual, municipal e particular) de São Carlos, cerca de 8.316 crianças estariam fora da escola. Em uma outra seção, será feita uma tentativa de contornar essa limitação dos dados das escolas, aproveitando informações do próprio censo municipal.

A Tabela 2.2 apresenta a distribuição de cor dos alunos e da população com idade até 19 anos. Utilizou-se o limite de 19 anos porque essa é a idade máxima encontrada entre os estudantes do ensino primário da base de dados. O Decreto n. 248, de 26 de julho de 1894, que ainda regulamentava o ensino primário paulista na primeira década do século XX, previa a possibilidade de indivíduos com idade de 16 anos frequentarem escolas primárias. Como os dados de idade foram obtidos no censo realizado em 1907 e os dados das escolas têm informações até o ano de 1909, era esperado encontrar estudantes com idade de 19 anos. De todo modo, esses alunos constituem uma minoria, uma vez que, dos 267 estudantes, apenas três possuíam 19 anos de idade. A distribuição de idade entre estudantes é apresentada na Tabela C.3 no Apêndice C.

Os resultados mostram que em São Carlos, entre 1906 e 1909, a representação de negros em escolas isoladas era maior que sua representação na população em geral. Como reproduzido na Tabela 2.2, pretos constituíam 9,9% dos alunos nas cinco escolas no período analisado e apenas 8,8% na população geral na mesma faixa etária, enquanto esses percentuais eram de 2,3% e 2,7% para mulatos, respectivamente. Em conjunto, negros eram 11,5% da população até 19 anos e 12,2% entre os alunos da base de escolas.

Os números apresentados na Tabela 2.2, contudo, devem ser interpretados com cautela. Como foi possível obter dados demográficos de cerca de 1/3 (267) das crianças nas escolas estaduais de São Carlos, a adição de novas informações sobre o contingente não conhecido dessa população poderia levar a mudanças na participação de alunos negros nas escolas. Um possível viés é a sobre-represen-

Tabela 2.2: Cor de alunos e da população até 19 anos, São Carlos, 1907

Cor	Estudantes		População até 19 anos	
	Total	(%)	Total	(%)
Branco	230	87,8	19923	88,5
Preto	26	9,9	1914	8,8
Mulato	6	2,3	578	2,7

Fontes: Livros de chamadas - Arquivo do Estado de São Paulo. São Carlos, Recenseamento de 1907.

Notas: Não é possível identificar com exatidão a nacionalidade de todos os pais das crianças brancas. Se for assumido o critério de imigrante de 2ª geração, isto é, que a nacionalidade da criança é a mesma nacionalidade do chefe, informação que não está disponível para todos os casos, então 87,6% (163/186) das crianças brancas eram imigrantes. A situação se inverte quando se considera o local de nascimento. Neste caso, das 230 crianças brancas, 201 eram brasileiras.

tação na amostra de escolas que podiam estar localizadas em bairros com maior presença de famílias negras. Para verificar essa possibilidade, cruzou-se o dado de cor dos alunos com o nome das escolas. O resultado é apresentado na Tabela 2.3.

Tabela 2.3: Cor de estudantes por escola, São Carlos, 1907

Escola	Branco	Mulato	Preto	% de negros
E. F. Bairro da Estação	44	1	1	4,3
E. F. de Santa Cruz	49	0	3	5,8
E. F. de Vila Nery	45	3	12	25,0
E. M. de Ibaté	60	0	5	7,7
E. M. de Vila Nery	32	2	5	17,9

Fontes: Livros de chamadas - Arquivo do Estado de São Paulo. São Carlos, Recenseamento de 1907.

Como se vê na Tabela 2.3, os alunos negros parecem estar concentrados nas escolas, masculina e feminina, de Vila Nery —dos 32 estudantes negros da amostra, 22 estavam nas escolas do bairro de Vila Nery. O viés causado pela localização das escolas com dados disponíveis fica mais claro quando se considera a distribuição racial dos bairros de São Carlos em 1907. A última página do censo municipal de São Carlos apresenta os dados agregados por bairro. Essa informação pode ser

analisada com o *heatmap* abaixo. Na Figura 2.2, a célula fica mais escura quando se aproxima de zero e mais clara quando próxima a 1. Assim, na primeira coluna, que apresenta os percentuais da população branca, vê-se que “Botafogo”, com 97,9% da população registrada como branca no censo, é o bairro mais branco de São Carlos em 1907. Já “Vila Nery” é o terceiro bairro mais negro da cidade, com quase 1/4 da população identificada como “preta” no censo e 5,7% como “mulata”. Ou seja, é razoável supor, com base na localização das escolas e nos residentes dos bairros, que o verdadeiro percentual de alunos negros matriculados em escolas isoladas de São Carlos na década de 1900 seja um pouco menor do que foi apresentado na Tabela 2.2.

Outra possível fonte de erro é o algoritmo utilizado para o cruzamento das informações das escolas com os dados do censo municipal. Supondo que a relação entre a maior distância de *string* entre dois nomes afeta negativamente a probabilidade das duas entradas corresponderem à mesma pessoa, uma maior média de distância de *string* entre estudantes negros significaria que a probabilidade dos cruzamentos dos dados dos estudantes negros serem válidos é menor do que dos brancos. Como a Tabela 2.4 mostra, é precisamente o contrário que ocorre. De fato, 33,3% dos estudantes mulatos e 38,5% dos estudantes pretos recebem índice Levenshtein de zero. Isto é, para esses casos os nomes que aparecem no livro de chamada são idênticos aos que aparecem no censo municipal. Por outro lado, esse percentual é de apenas 12,2% para os alunos registrados como brancos. Em outras palavras, se fossem usados apenas registros idênticos, a tendência seria subestimar o número de crianças brancas, visto que nesse caso o percentual de crianças negras nas escolas seria de 71,8%.

É possível, portanto, identificar duas possíveis fontes de erros nos dados disponíveis, que podem fazer com que a estimativa seja diferente do percentual real de alunos negros nas escolas de São Carlos na década de 1900. O viés de localização tende a superestimar o número de estudantes negros nas escolas, enquanto o viés causado pelo algoritmo de *matching* parece subestimar esse número. Mesmo se se supuser que o viés de localização é dominante, é revelador da demanda da população negra por educação o mero fato de existir um número razoável de crianças negras em escolas primárias em uma localidade marcada

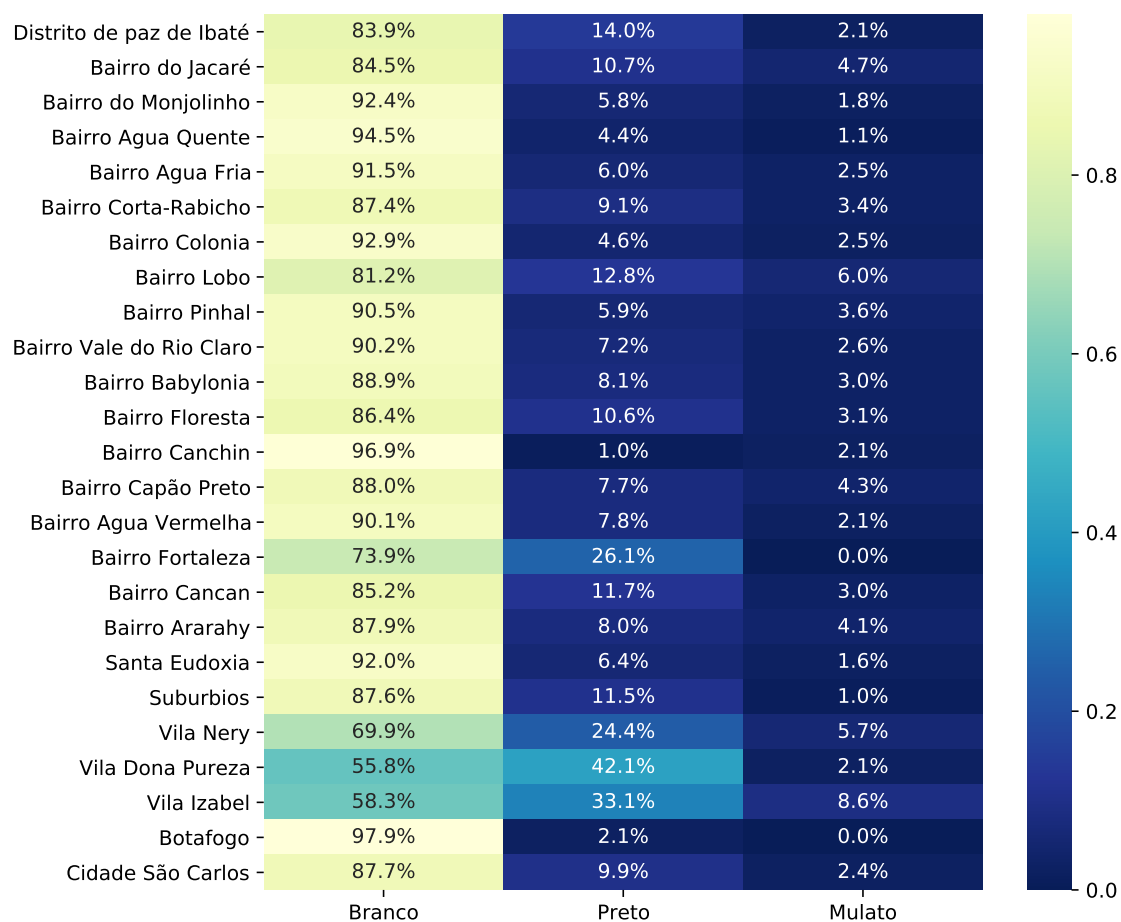


Figura 2.2: Distribuição de cor por bairro, São Carlos, 1907

Fonte: São Carlos, Recenseamento de 1907.

Notas: As linhas representam os bairros de São Carlos em 1907, enquanto as colunas representam os grupos demográficos “Branco”, “Preto” e “Mulato”. Assim, cada célula indica o percentual de pessoas do grupo da coluna no bairro da linha. Conforme escala à direita, as cores indicam se o valor está mais próximo de 0 ou de 100%. Por exemplo, o bairro com maior proporção de pretos é Vila Dona Pureza com 42,1% de pretos em relação ao total de moradores do bairro. Para distribuição por nacionalidade, ver Figura D.1 no Apêndice D.

por enorme predominância demográfica da população branca, em um contexto de expressiva violência racial.¹²

12. Monsma, *A reprodução*, Cap. 5.

Tabela 2.4: Índice de Levenshtein por cor dos estudantes, em percentual, São Carlos, 1907

Levenshtein	Branco	Mulato	Preto
0	12,2	33,3	38,5
1	15,7	0,0	11,5
2	24,3	0,0	15,4
3	15,7	50,0	7,7
4	11,7	0,0	15,4
5	7,8	0,0	0,0
6	4,8	0,0	0,0
7	3,0	16,7	7,7
8	1,3	0,0	3,8
9	1,3	0,0	0,0
10	1,3	0,0	0,0
11	0,4	0,0	0,0
17	0,4	0,0	0,0

Fontes: Livros de chamadas - Arquivo do Estado de São Paulo. São Carlos, Recenseamento de 1907.

Notas: Os valores da Tabela referem-se aos percentuais de *matchings* para cada valor da distância de Levenshtein. Por exemplo, o valor na segunda coluna e primeira linha indica que $1/3$ dos estudantes que identificamos como mulatos têm Levenshtein igual a zero em relação ao seu correspondente no censo. Isso significa que para $1/3$ dos estudantes mulatos o *matching* tem o nome idêntico.

Na próxima seção, apresenta-se uma análise exploratória do perfil demográfico das famílias dos alunos nas escolas primárias. O objetivo é identificar quais fatores no âmbito familiar tornaram a matrícula escolar mais provável e como o perfil específico de famílias negras se relaciona com esses fatores.

2.2.1 Características das famílias dos estudantes

Como a lista nominal dos alunos matriculados em escolas em São Carlos em 1907 é parcial, não é possível realizar uma comparação exata do perfil demográfico dos pais com o do restante da população adulta. Ainda assim, nesta seção, busca-se contornar as limitações dos dados analisando a correlação entre as variáveis das famílias dos alunos e a matrícula escolar, utilizando as médias da população total, formada majoritariamente de não matriculados, como referência para compara-

ção. É importante notar que, nas tabelas abaixo, o grupo de comparação refere-se à população masculina e feminina adulta na mesma faixa etária dos pais e mães dos alunos.

O primeiro passo é observar as profissões dos pais dos alunos. As Tabelas 2.6 e 2.5 listam, respectivamente, as dez profissões mais comuns de pais e mães. Para efeito de comparação, cada uma das tabelas contém a mesma informação para a população feminina e masculina do município na mesma faixa etária.

Entre os homens, fica clara a predominância da profissão de “Negociante” entre os pais dos alunos, uma vez que homens nessa ocupação são apenas 3% do total na população e 9,4% entre os pais dos estudantes. No outro extremo, a sub-representação mais clara entre pais é a de “Colono”, que é a ocupação de metade da população masculina, mas apenas 26% entre os pais dos alunos. Em relação às mulheres, a sub-representação de colonos entre mães de alunos também prevalece (49% na população em geral contra 26,6% entre mães de alunos). Tal sub-representação pode ser efeito do fato de que mulheres de colonos eram normalmente classificadas como colonos no censo municipal de 1907. Outro padrão que se repete é a predominância de negociantes entre as mães dos alunos, quando comparada com o total de mulheres negociantes no conjunto da população (7,1% entre mães de alunos, contra apenas 1,8% na população feminina total). Uma diferença relevante dos dados de homens e mulheres é o número de casos não informados (n.d.) para as mulheres. É digno de nota que essa ausência de informação, que pode significar inexistência de trabalho fora do domicílio, seja mais comum entre as mães dos alunos do que na população feminina em geral (ver Tabela 2.6)

Em um município caracterizado por grande presença de imigrantes, a distribuição de profissões raramente é independente da origem e do perfil étnico-racial dos indivíduos. Sírio-libaneses, por exemplo, eram desproporcionalmente representados entre negociantes. Dos 35 sírio-libaneses homens casados no censo local, vinte e oito (80%) foram registrados como negociantes. A existência de correlação entre a origem étnico-racial com ocupações, observada para sírio-libaneses, nos leva a indagar se seria possível identificar uma relação semelhante para a população negra. Os dados mostram, contudo, que tal relação não existiu entre pais

Tabela 2.5: Profissão dos Pais dos estudantes e da população masculina, São Carlos, 1907
(em porcentagem da respectiva população)

População Masculina n=9117		Pais de alunos n=192	
Colono	50,2	Colono	26,0
Camarada	11,8	Lavrador	11,5
Lavrador	9,5	Negociante	9,4
Empregado	3,6	Camarada	5,2
Negociante	3,0	Pedreiro	4,7
Pedreiro	1,5	Empregado	4,2
Carroceiro	1,4	Carpinteiro	3,6
Administrador	1,3	Empregado Cia Paulista	3,1
n.d	1,3	Sapateiro	3,1
Empregado Cia Paulista	1,2	Carroceiro	3,1

Fontes: Livros de chamadas - Arquivo do Estado de São Paulo.
São Carlos, Recenseamento de 1907.

Notas: "n.d" corresponde a não disponível. Os casos indisponíveis ocorrem quando a informação não foi preenchida o que pode significar tanto ausência de informação como ela não ser aplicável para o indivíduo em questão. Por exemplo, o fato de uma mulher não ter o campo "profissão" não preenchido pode ter sido causado tanto pela falta da informação quanto pelo fato da entrevistada se dedicar exclusivamente à atividade doméstica. Para definição de "Empregado" ver Capítulo 3.

negros, pelo menos com os dados disponíveis. Isto é, a presença de negros nas escolas não parece ser resultado de negros ocuparem postos de trabalho associados com maior grau de alfabetização, como profissões urbanas e ligadas ao comércio. Para os 24 pais de estudantes registrados como pretos ou mulatos no censo para os quais se tem informação sobre profissão, a maioria exercia os dois ofícios mais comuns entre a população negra em geral na época, quais sejam, colono e camarada. Observe-se, porém, que essas duas ocupações entre pais dos alunos possuíam ocorrência menor que a frequência na população em geral. Enquanto cerca de 70% dos negros na população eram colonos (39%) ou camaradas (31,2%), entre os pais essas duas ocupações perfazem 54% do total. Por outro lado, a ocupação de lavradores é maior entre os pais negros dos alunos do que entre o conjunto da população. Negros registrados como lavradores constituíam

Tabela 2.6: Profissão das Mães dos estudantes e da população feminina, São Carlos, 1907
(em porcentagem da respectiva população)

População feminina n=7934		Mães de alunos n=184	
Colono	49,0	n.d	38,6
n.d	17,1	Colono	26,6
Lavrador	9,2	Lavrador	10,9
Camarada	4,0	Negociante	7,1
Cozinheiro	2,1	Camarada	2,7
Empregado	1,9	Comerciante	2,7
Negociante	1,8	Costureiro	2,7
Costureiro	1,8	Empreiteiro	1,1
Lavadeira	1,2	Sapateiro	1,1
Empreiteiro	0,8	Lavadeira	1,1

Fontes: Livros de chamadas - Arquivo do Estado de São Paulo.
São Carlos, Recenseamento de 1907.

Notas: Ver nota da Tabela 2.5.

apenas 5,8% no total da população e 12,5% entre os pais dos alunos (Ver Tabela C.1 em anexo).

O maior número relativo de pais de alunos na ocupação de lavrador (ver Tabela 2.5) foi uma característica observada tanto para a população branca quanto para a população negra. Se se entende educação como um bem normal, ou seja, um bem cujo consumo aumenta quando a renda aumenta, e se se assume que lavradores estavam em situação material melhor do que colonos e camaradas, então a maior participação relativa de lavradores entre os pais dos estudantes pode ser reflexo da melhor condição material desse grupo em relação ao resto da população. Um dado que reforça a conjectura sobre a condição material de lavradores é a correlação entre essa profissão e propriedade da terra. Cerca de 36% dos proprietários de terra registrados no censo eram lavradores, apesar desses últimos constituírem apenas 9,5% entre homens casados. Em termos absolutos, a predominância de lavradores entre proprietários de terra é ainda mais notável, dado que 2.418 indivíduos registrados nessa ocupação tinham acesso à terra, de acordo com o censo. Por sua vez, negociantes, a segunda profissão com mais pro-

prietários, contava apenas com 445 indivíduos registrados como proprietários. Se ainda faltassem elementos para considerar a ocupação de lavrador associada a uma maior distinção social, basta dizer que todos os adultos com sobrenome Camargo Penteado ou Arruda Botelho, famílias que dominavam a elite política local, foram registrados como lavradores no censo municipal.¹³

Outro fator a ser considerado para avaliar a influência dos pais sobre a probabilidade de matrícula escolar dos filhos é o nível educacional do pai e da mãe. A Tabela 2.7 apresenta esses dados para os pais dos alunos, comparados com o conjunto da população. Em geral, os pais de crianças matriculadas tendiam a ter maior nível de alfabetização que o conjunto da população. Contudo, o inverso ocorria com homens pretos. A alfabetização na população preta masculina como um todo era de 15,4%, enquanto entre os pais pretos de alunos é de apenas 10,5%. Esse resultado pode ser considerado surpreendente e parece indicar que, entre famílias pretas, a alfabetização da mãe é um determinante mais relevante para o acesso à escola do que entre famílias brancas e mulatas. Para as famílias brancas e mulatas, a tendência dos pais dos estudantes terem taxas maiores do que a média populacional se mantém tanto para a população masculina quanto para a população feminina. Assim, a taxa de alfabetização de pais brancos (55,4%) é maior do que a taxa da população masculina da mesma cor de pele (46,5%), o mesmo valendo para mães brancas (29,8% contra 28,9%). Entre mulatos, essas diferenças são maiores, já que mais de 2/3 dos pais mulatos são alfabetizados, contra apenas 28,2% na população total. Entre as mães mulatas, a diferença é de 14 pontos percentuais (28,2% contra 14,2%).

Nesse ponto, convém lembrar de algumas características do sistema educacional brasileiro no século XIX. O sistema educacional durante o Império tinha como uma de suas características o desequilíbrio no número de matrículas de meninos e meninas nas escolas primárias, com vantagem para os primeiros. Em 1839, por exemplo, das 77 escolas da província paulista, 68 eram de meninos.¹⁴ Também havia limitação de conteúdo na educação feminina. Tanto a Lei de 1827

13. Truzzi, *Café e indústria*, p.42.

14. Mauriceia Ananias, "A legislação da instrução pública primária na província de São Paulo: 1834-1868" (tese de doutorado, Universidade Estadual de Campinas, 2005), 68.

Tabela 2.7: Alfabetização entre pais de estudantes e na população, São Carlos, 1907

Cor/raça	População Masculina		Pais		População Feminina		Mães	
	n	(%)	n	(%)	n	(%)	n	(%)
Branco	7913	46,5	202	55,4	6889	28,9	205	29,8
Mulato	234	28,2	6	66,7	212	14,2	6	50,0
Preto	971	15,4	19	10,5	833	6,1	22	9,1

Fontes: Livros de chamadas - Arquivo do Estado de São Paulo. São Carlos, Recenseamento de 1907.

do governo central quanto a Lei 34 de 1846 da província paulista, dois institutos legislativos importantes na construção do sistema educacional no século XIX, excluía meninas do ensino mais avançado de matemática. A lei 34 definia, em seu primeiro capítulo, que “*A instrução primaria comprehende a leitura, escripta, theoria e practica da arithmetica até proporções inclusivè, as noções mais geraes de geometria pratica, grammatica da lingua nacional, e principios da moral christã, e da doutrina da religião do estado*”. Contudo, a leitura dos demais capítulos deixa claro que essa definição estava sujeita a alterações dependendo do sexo dos alunos, pois já no capítulo 2º vê-se que para alunas aplicavam-se as mesmas disciplinas listadas no capítulo primeiro com exclusão da “*geometria; e limitada a arithmetica á theoria e practica das quatro operações; e tambem das prendas que servem á economia domestica*”. Essa disposição normativa teve sua inspiração no Capítulo 12 da lei de 1827.

A dominância masculina nas escolas iria diminuir ao longo do século XIX. Renato Colistete mostra que meninas representavam apenas 11,1% do corpo discente masculino das escolas primárias paulistas entre 1836 e 1840, mas essa razão subiu para 64,6 entre 1876-1880.¹⁵ Essa desigualdade na oferta de educação teria efeitos de longa duração, já que a menor taxa de matrícula do ensino primário para meninas no período imperial é possivelmente uma das razões que explicam o *gap* educacional entre a população adulta feminina e masculina no início do século XX.

Com efeito, o censo municipal de São Carlos mostra que, para a população maior que 15 anos, quase 75% das mulheres eram analfabetas, enquanto esse per-

15. Colistete, *O atraso*, 109.

centual era de 58,7% para homens. Quando se comparam as taxas de alfabetização por cor, para ambos os sexos, encontra-se uma diferença ainda maior: 62% da população branca com mais de 15 anos era analfabeta, contra 87% de analfabetismo na população negra —claramente refletindo a herança da escravidão. A consequência desses números é que, considerando as desigualdades de gênero e cor, a mulher negra era a que se encontrava em pior situação já que 91% das mulheres negras adultas eram analfabetas. Esse resultado torna ainda mais relevantes e surpreendentes os indícios anteriores de que as mães mulatas e pretas podem ter sido mais influentes na educação das crianças do que os homens em suas famílias. Mesmo com a herança da escravidão, a altíssima taxa de analfabetismo em 1907, e com a precariedade das escolas públicas isoladas, mães mulatas e pretas deram uma contribuição relativamente grande à educação de seus filhos.

Em suma, profissão, patrimônio e alfabetização dos pais parecem ter tido os efeitos esperados sobre a probabilidade das crianças terem acesso à escola em São Carlos, no início do século XX. A análise anterior indica que lavradores proprietários de terra e negociantes eram os grupos que tinham mais chances de ter filhos matriculados nas escolas, em todos os grupos étnicos. A alfabetização dos pais também foi um fator relacionado a maiores taxas de matrícula. Este último efeito, contudo, parece ter sido maior para mulatos, enquanto que para pretos o papel relevante foi ocupado pelas mães, que parecem ter tido um papel determinante na educação de seus filhos.

Ainda é preciso, no entanto, analisar o efeito de cada um desses fatores condicionalmente aos demais. Em uma situação ideal, estimar um modelo econométrico que calcule a probabilidade de uma criança estar matriculada em uma escola, condicionalmente às variáveis da própria criança e de seus pais, requer que se saiba quais crianças do censo de 1907 estavam na escola e quais não estavam. Como visto, infelizmente os estudantes das escolas cujos dados estão disponíveis representam apenas uma fração do total de estudantes em São Carlos, em 1907.

Na prática, portanto, não há meios de saber diretamente quais crianças não incluídas na base de dados estavam, de fato, fora da escola. Por esse motivo, na próxima seção, serão utilizadas informações do censo municipal de 1907 para estimar quais crianças registradas neste levantamento populacional estavam fora

da escola e quais estavam na escola. Dessa forma, será possível em seguida elaborar um modelo econométrico e analisar como as variáveis consideradas nessa seção se correlacionam condicionalmente à probabilidade de uma criança estar na escola.



Figura 2.3: Alunos de escolas de Ibaté, São Carlos, 1917

Fonte: Almanach-Album de São Carlos: 1916-1917

2.3 Determinantes da matrícula escolar: uma análise quantitativa a partir do censo municipal

A análise da seção 2.2 assumiu que era uma boa aproximação estimar a taxa de matrícula de estudantes negros em São Carlos no início do século XX apenas com informações dos estudantes das cinco escolas cujos dados foi possível ter acesso nos arquivos. Além disso, para a análise dos dados da seção 2.2, assumiu-se adicionalmente que as médias das variáveis demográficas da população adulta de São Carlos consistiam em uma boa estimativa dos dados demográficos dos pais de crianças não-matriculadas em escolas. Nesta seção, apresenta-se uma abordagem diferente para responder às mesmas questões da seção 2.2, isto é, primeiro, quais eram as taxas de matrícula escolar por cor em São Carlos em 1907? Segundo, como o perfil demográfico dos pais impacta a probabilidade de matrícula escolar?

Com os dados do censo municipal, assume-se que a criança estava ou esteve recentemente matriculada em escola se ela tinha entre 7 e 15 anos de idade e se, adicionalmente, cumpria um dos seguintes requisitos:

1. O nome consta entre as 267 crianças identificadas nas escolas.
2. O campo “Profissão” está marcado como “Estudante” no censo municipal.
3. A criança está registrada como alfabetizada no censo.

Como há interseção entre os critérios, o número total de casos será inferior à soma do número absoluto de estudantes estimado por cada critério separadamente. De todo modo, é instrutivo contabilizar em quanto cada critério contribui separadamente. Das 267 crianças com nomes identificados nas escolas, 179 se encaixam no critério etário estabelecido (7 a 15 anos). Apenas 94 pessoas tiveram no censo o campo “Profissão” registrado como “Estudante”, sendo que uma parcela dessas pessoas eram matriculadas no ensino secundário. Com a faixa etária de 7 a 15 anos, esse número cai para 81. O último critério (alfabetização) foi o que mais contribuiu para a estimativa, pois segundo o censo de São Carlos, 1.803 crianças

com idade entre 7 e 15 anos eram alfabetizadas. Ao incluir crianças estrangeiras, que podem ter sido alfabetizadas no exterior antes da imigração, corre-se o risco de superestimar a taxa de matrícula escolar da época para crianças brancas. Esse viés, contudo, é provavelmente pequeno e atua contra a hipótese de que havia uma elevada proporção de negros matriculados em escolas primárias no pós-abolição em São Carlos.

Adotando os três critérios (nome na lista de chamada, idade, profissão e alfabetização) enumerados acima, estima-se que 1.924 crianças entre 7 e 15 anos estiveram em algum momento matriculadas em uma escola primária em São Carlos. Esse número é maior que o número real de 1.315 para 1909, segundo os dados oficiais de matrícula. Essa diferença é esperada, dado que o *Anuário do Ensino* representa o total de alunos em um momento no tempo, enquanto nossa estimativa usa parcialmente o número acumulado de alunos ao longo do período 1906-1909. Além disso, como visto no parágrafo anterior, há o potencial viés causado pela inclusão de crianças estrangeiras alfabetizadas.

As taxas de matrícula por cor calculadas a partir dessa nova estimativa são apresentadas na Tabela 2.8. Dessa vez, as estimativas são um pouco mais modestas para a população negra. Enquanto na primeira abordagem o percentual de negros e brancos nas escolas não diferia significativamente da participação desses grupos no resto da população, a nova estimativa aponta para uma pequena vantagem para a população branca, já que entre brancos a taxa de matrícula estimada seria de 21,1%, enquanto pretos e mulatos tinham taxas de matrícula de 10,5% e 15,5%, respectivamente. Os números da Tabela 2.8 implicam ainda que brancos eram 93% (1790/1924) das crianças nas escolas de São Carlos, número que supera um pouco o percentual de brancos entre 7 e 15 anos na cidade em 1907, que era de 87,8%. Este resultado era esperado, já que os resultados da seção 2.2 foram calculados apenas para as escolas em que há dados disponíveis. Todas essas escolas eram escolas isoladas (escolas públicas estaduais localizadas em bairros e distritos). Os resultados apresentados nesta seção referem-se a todas as escolas incluindo, portanto, o grupo escolar do município. Da mesma forma, é possível que a redução da proporção de negros seja resultado da inclusão de escolas particulares. Por fim, vale ressaltar que, independentemente de cor, por

essa estimativa mais de 80% (7707/9631) das crianças estariam fora da escola em São Carlos na época.

Tabela 2.8: Taxas de matrícula por cor, São Carlos, 1907

Cor/raça	Matrculados		Não-Matriculados		Total n_1+n_2
	n_1	(%)	n_2	(%)	
Branco	1790	21.1%	6675	78.9%	8465
Mulato	37	15.5%	201	84.5%	238
Preto	97	10.5%	831	89.5%	928
Total	1924		7707		9631

Fontes: Livros de chamadas - Arquivo do Estado de São Paulo. São Carlos, Recenseamento de 1907.

Notas: Colunas 3 e 5 em percentuais dos valores na linha. Por exemplo, chega-se à conclusão que 21,1% das crianças brancas estão matriculadas na escola computando 1790/8465.

2.3.1 Análise de regressão

Como mencionado anteriormente, os modelos econométricos são empregados para estimar a correlação das variáveis demográficas dos estudantes e de suas famílias sobre a probabilidade de matrícula na escola. As famílias dos alunos foram identificadas no censo municipal quando a criança foi registrada como dependente e havia um adulto classificado como chefe de família acima do registro dessa criança. Considerou-se “mãe” quando a chefe da família estava registrada como sendo do sexo feminino ou quando, adjacente ao chefe da família de sexo masculino, havia um indivíduo do sexo feminino registrado como esposa. É importante seguir esse critério, pois considerar todas as crianças registradas próximas a adultos no censo como filhos daqueles adultos levaria a considerar trabalhadores infantis do lar, como criados, como sendo filhos da família em questão. A decisão de seguir esse critério teve como consequência a eliminação de 733 casos do banco de dados.

Foram estimados três modelos, utilizando como variável dependente a matrícula escolar, que assume 1 se a criança está ou esteve matriculada em uma escola de acordo com a estimativa feita antes, e zero caso a criança não esteve matricu-

lada. Como pode ser depreendido da Tabela 2.8, a distribuição de matrícula é bastante assimétrica, já que das 9.631 crianças listadas na base de dados, estimou-se que apenas 1.924 estiveram em algum momento matriculadas em uma escola primária. A assimetria na distribuição da variável dependente persistiu mesmo após serem eliminados os casos para os quais não foi possível identificar os pais dos estudantes. Como variáveis explicativas da probabilidade de matrícula, foram incluídas as seguintes informações: negro (1 se o aluno era preto ou mulato, zero caso ele fosse branco), colono (1 se o pai era colono, zero caso ele tivesse qualquer uma das demais profissões), mãe brasileira (1 se a mãe era brasileira, zero caso ela tivesse outra nacionalidade), feminino (1 se a criança era do sexo feminino, zero se a criança fosse do sexo masculino), pai proprietário (1 se o pai era proprietário de terra, zero caso o pai não fosse proprietário).¹⁶ Além dessas variáveis que descrevem características qualitativas da população de forma binária (*dummies*), utiliza-se também como variável explicativa o número de membros da família.

A Tabela 2.9 apresenta os resultados dos três modelos. O primeiro modelo utiliza mínimos quadrados ordinários e, nesse caso de variável dependente binária, é comumente chamado de modelo de probabilidade linear (MPL). Como pode ser visto na coluna (1) da tabela, o valor da constante no modelo de probabilidade linear (0,316) significa que, para uma família com o mínimo de membros (isto é, 3) e cujo valor para as demais variáveis é zero —vale dizer, uma criança branca do sexo masculino, com o pai não-colono, mãe não-brasileira e analfabeta—, a probabilidade de matrícula escolar é de 31%. A variável que mais aumenta a probabilidade de matrícula da criança é a alfabetização da mãe (0.224). Ter a mãe alfabetizada aumentava a probabilidade da criança se matricular em uma escola em 22 pontos percentuais, resultado significativo mesmo ao nível de 1% de confiança.

Também estima-se o mesmo modelo adicionando a *dummy* de alfabetização do pai. Os resultados deste modelo são apresentados na Tabela C.4 no Apêndice C. Com a informação sobre alfabetização do pai, o coeficiente da alfabetização

16. Quando havia a mesma informação para o pai e a mãe da criança, optou-se por escolher uma delas, para não aumentar a variância dos estimadores, já que há alta correlação entre as variáveis de pai e mãe

Tabela 2.9: Resultados

	<i>Variável dependente:</i>		
	Matrícula		
	<i>MPL</i>	<i>logit</i>	<i>probit</i>
	(1)	(2)	(3)
Negro	-0,108*** (0,013)	-0,839*** (0,117)	-0,461*** (0,065)
Pai Colono	-0,250*** (0,011)	-2,241*** (0,087)	-1,197*** (0,046)
Mãe Brasileira	0,012 (0,011)	-0,180** (0,079)	-0,072 (0,046)
Feminino	-0,020*** (0,007)	-0,168*** (0,064)	-0,117*** (0,036)
Mãe alfabetizada	0,224*** (0,012)	1,445*** (0,072)	0,829*** (0,041)
Pai Proprietário	0,134*** (0,015)	0,630*** (0,073)	0,382*** (0,044)
N. de membros	-0,004*** (0,002)	-0,044*** (0,013)	-0,023*** (0,007)
Constante	0,316*** (0,016)	-0,651*** (0,105)	-0,413*** (0,061)
Observações	8.827	8.827	8.827
R ²	0,272		
R ² Ajustado	0,272		
log-verossimilhança		-3,156,400	-3,156,700
Crit. Inf. Akaike		6,328,799	6,329,399
Erro padrão do resíduo	0,340 (df = 8819)		
F (Wald)	369,88*** (df = 7; 8819)		

*p<0,1; **p<0,05; ***p<0,01

Fontes: Livros de chamadas - Arquivo do Estado de São Paulo.
São Carlos, Recenseamento de 1907.

Notas: Erros robustos para heteroscedasticidade no MPL.

da mãe mantém a significância estatística a 1%, mas seu efeito parcial cai para 17 pontos percentuais - ainda assim, maior do que o observado para o coeficiente dos pais que é de 12 pontos percentuais. A segunda variável com maior impacto positivo, sobre a probabilidade de matrícula escolar é propriedade da terra (0,134). O modelo de probabilidade linear prediz que ter pai proprietário de terra aumentava a chance de matrícula escolar em quase 14 pontos percentuais.

Por outro lado, outras variáveis têm o efeito de diminuir a probabilidade de matrícula escolar. Dentre estas, a primeira variável de maior impacto é a profissão do pai. De fato, ter pai colono diminuía a probabilidade de matrícula em 25 pontos percentuais em relação às demais profissões. Esse resultado é consistente com o que foi observado na seção 2.2. Como visto, as profissões de pais associadas com maiores taxas de matrícula escolar foram negociante e lavrador.

Em segundo lugar entre as variáveis que afetam negativamente a probabilidade de matrícula aparece a cor da pele da criança. Pelo modelo, se uma criança era preta ou mulata, sua probabilidade de matrícula é 10 pontos percentuais menor do que a de crianças brancas, mantendo constantes as demais variáveis. É possível, porém, que esse resultado seja diferente para crianças pretas e mulatas. A estimação da taxa de matrícula para esses dois subgrupos da população negra apresenta uma história inconsistente. Usando dados da escola e apenas a informação de cor do censo (primeira abordagem, seção 2.2), viu-se que a razão do percentual de matriculados e percentual na população é maior para pretos, ou seja, a taxa de matrícula para esse subgrupo era maior do que a de mulatos. Contudo, o resultado se inverte usando dados da segunda abordagem da seção 2.3. Assim, como a análise até aqui não permite verificar se há diferença entre esses dois grupos, opta-se por realizar uma regressão adicional, usando duas *dummies* para cor (preto e mulato). Os resultados são apresentados na Tabela C.5, disponível no Apêndice C. Os resultados mostram que pretos tinham uma chance 10 pontos percentuais menor de estar na escola em relação a crianças brancas, enquanto essa mesma estatística é de 11 pontos percentuais para mulatos. A diferença, porém, não é estatisticamente significativa. É possível que a diferença das taxas em favor dos estudantes mulatos que aparece na segunda abordagem seja, em parte, explicada pelo fato das mães mulatas terem maior taxa

de alfabetização, como visto na Tabela 2.7.

Como notado anteriormente, ao longo do século XIX houve redução do hiato de matrículas entre meninos e meninas. Essa redução se manteve após a Proclamação da República, com a razão entre alunas e alunos alcançando 94,5% em 1906-1910 na instrução primária paulista.¹⁷ No entanto, os resultados mostram que, após controlar para outras variáveis, meninas ainda tinham, em comparação aos meninos, uma chance 2 pontos percentuais mais baixa de ingressar em uma escola no período. O coeficiente dessa variável é significativo a 1%. A única variável a não se revelar estatisticamente significativa foi a *dummy* “mãe brasileira”.

O modelo de probabilidade linear tem a vantagem de ter interpretação e cálculo simples, mas sofre de duas limitações. A primeira é que, como não há nenhum tipo de censura na variável dependente —ou seja, não há imposição de um intervalo fechado ao qual a variável deve pertencer—, nada impede que, para uma observação particular, a predição do modelo seja negativa ou maior que 1. Além disso, como o modelo é linear nos parâmetros, o efeito parcial de uma variável é constante, independentemente do nível dessa variável.¹⁸ No caso aqui examinado, por exemplo, o modelo de probabilidade linear prediz que o efeito de um membro adicional na família sobre a probabilidade de matrícula será de 0,2 pontos percentuais, independentemente do número de membros da família.

Como forma de contornar essas limitações, são estimados dois modelos adicionais: *probit* e *logit*, com os respectivos resultados descritos na Tabela 2.9. Vale notar que, diferentemente do modelo de probabilidade linear, os coeficientes dos modelos *probit* e *logit* não podem ser interpretados diretamente. Para obter uma interpretação análoga ao modelo de probabilidade linear, é preciso multiplicar os coeficientes por um fator que é de aproximadamente de 0,18 para o *logit* e 0,3 para o *probit*.¹⁹ Aplicando essa regra simples, encontram-se os efeitos marginais para

17. Colistete, *O atraso*, 223.

18. Jeffrey M Wooldridge, *Introductory econometrics: A modern approach*, 5ª ed. (Cengage Learning, 2012), 583.

19. Na prática, o efeito parcial de qualquer variável x nesses modelos é dado por: $\frac{\partial P(y=1|x)}{\partial x} = g(X\beta)f'(x)$. Como $g(X\beta)$ é função das variáveis independentes, seu valor pode variar. No presente caso, adota-se o *average marginal effect* (AME), dado por $n^{-1} \sum_{i=1}^n g(X\beta)$

o *probit* e o *logit*, cujos resultados são apresentados na Tabela 2.10. A Tabela 2.10 mostra que a hierarquia dos efeitos parciais não mudou nas estimações de *logit* e *probit*. Por exemplo, os valores na Tabela 2.10 implicam que ter pai colono reduz a probabilidade de matrícula em 25 pontos percentuais no modelo *logit* e 24 no modelo *probit*. Assim como no modelo de probabilidade linear, nos dois modelos, a variável com maior efeito positivo é a alfabetização da mãe, enquanto a variável com efeito negativo mais forte é ter pai colono.

Tabela 2.10: Efeitos marginais - *logit* e *probit*

	logit	Erro-padroa	probit	Erro-padroa
Negro	-0.094***	0.013	-0.092***	0.014
Pai Colono	-0.252***	0.01	-0.238***	0.009
Mãe Brasileira	-0.02**	0.009	-0.014	0.009
Feminino	-0.019***	0.007	-0.023***	0.007
Mãe Alfabetizada	0.163***	0.008	0.165***	0.008
Pai proprietário	0.071***	0.008	0.076***	0.009
N. membros	-0.005***	0.002	-0.005***	0.002

Fontes: Livros de chamadas - Arquivo do Estado de São Paulo.
São Carlos, Recenseamento de 1907.

Notas: *p<0,1; **p<0,05; ***p<0,01

O sinal dos coeficientes e sua magnitude relativa são muito similares ao modelo de probabilidade linear. Por exemplo, a coluna do modelo de probabilidade linear na Tabela 2.9 indica que uma criança do sexo feminino tem probabilidade 2 pontos percentuais mais baixa de estar matriculada na escola; esse valor é quase idêntico aos encontrados para os modelos *logit* (1,9%) e no *probit* (2,3%) na Tabela 2.10. Em todos os casos, o coeficiente é estatisticamente significativo mesmo a 1%.

Contudo, há diferenças na magnitude de coeficientes individuais e na significância estatística de algumas variáveis, embora não substanciais. Por exemplo, o efeito da *dummy* "Negro", que foi de 10 pontos percentuais no modelo de probabilidade linear, tem um pequena queda para 9 pontos percentuais no *logit* e no *probit*. Ter mãe alfabetizada aumentava a probabilidade de matrícula em 16 pontos percentuais de acordo com os dois últimos modelos, contra 18 pontos percentuais no modelo de probabilidade linear. Por fim, com discrepância maior, ter pai proprietário aumentava a chance de matrícula escolar em cerca de 7 pontos

percentuais no *probit* e no *logit* e 22 pontos percentuais no modelo de probabilidade linear. Além disso, houve diferença de significância estatística em algumas variáveis: “mãe brasileira”, que não é significativa no modelo de probabilidade linear, é significativa a 5% no *logit*.

De maneira resumida, portanto, apresentou-se nesta seção uma abordagem variada para identificar quais crianças listadas no censo municipal de 1907 estavam na escola e quais não estavam. A partir dessa informação, foram estimados modelos econométricos com o objetivo de avaliar a correlação de variáveis socioeconômicas dos alunos e de suas famílias sobre a probabilidade de matrícula. Os resultados não mudam a essência das conclusões da seção 2.2, quando foram analisadas as correlações não condicionais. Os resultados sugerem, assim, que determinadas profissões exercidas pelo pai, como lavrador e negociante, a alfabetização da mãe e o patrimônio familiar eram fatores que aumentavam as chances de uma criança estar na escola em São Carlos, no segundo quinquênio da década de 1900.

2.4 Conclusões

Este capítulo apresentou estimativas da matrícula escolar de crianças negras para o município de São Carlos no período 1906-1909, bem como examinou em detalhe quais fatores socioeconômicos da família tornavam a matrícula escolar mais provável. Corroborando resultados que vêm sendo apresentados pela historiografia da educação negra brasileira, os resultados aqui obtidos apontam para uma presença relevante da população descendente de escravizados nas escolas primárias no pós-abolição.²⁰

Os dados apresentados sugerem que, em uma população em que negros perfaziam 11,5% da população em idade escolar, crianças negras respondiam por 12,2% das matrículas em escolas isoladas. Embora em posição similar à população branca nas escolas isoladas, é possível que as crianças negras estivessem em desvantagem no conjunto de matrículas da educação primária em São Carlos, quando o grupo escolar municipal e as escolas privadas são incluídas. Como visto na

20. Fonseca, “Pretos”; Da Silva, “A escola do professor Pretextato”.

seção 2.1, entre as escolas com livros de chamadas disponíveis não está incluído, por exemplo, o grupo escolar do município.

A seção 2.3 procurou contornar a limitação imposta pela observação parcial dos livros de chamada ao usar outros critérios para considerar uma criança matriculada em escola, como o indicador de alfabetização do censo. As novas estimativas mostraram que a presença de negros nas escolas como um todo era de fato menor que o percentual observado nas escolas isoladas. Com efeito, considerando todas as escolas, negros eram cerca de 7% do corpo discente das escolas (ver primeira coluna da Tabela 2.8).

O presente capítulo não se estendeu na análise das diferenças de matrícula escolar entre negros e imigrantes. De todo modo, considerando os casos em que o chefe da família está identificado, filhos de imigrantes eram mais de 87% das crianças brancas em escolas isoladas. Mesmo essa estimativa, que desconsidera escolas privadas e o grupo escolar do município, pode subestimar a presença imigrante em escolas primárias. Em 1911, a presença estrangeira em grupos escolares era bastante elevada em municípios paulistas. Em Campinas, por exemplo, o percentual de alunos com pais estrangeiros nos grupos escolares era de 62,5%. Um percentual modesto diante de Pindamonhangaba, que tinha mais de 99% de estudantes estrangeiros no seu grupo escolar.²¹

Os resultados do capítulo mostram que, mesmo havendo significativos obstáculos impostos à população negra no acesso à escola, as crianças negras conseguiam ocupar espaço nas instituições de ensino primário públicas. De particular interesse foi o papel exercido por mulheres negras na educação de seus filhos. Entre famílias pretas, a educação da mãe parece ter sido fator fundamental na determinação da matrícula, especialmente considerando os indicadores de alfabetização relativamente baixos de pais pretos. A análise de regressão reforçou esse achado ao mostrar que a variável com maior impacto positivo na probabilidade de uma criança se matricular em uma escola primária é a alfabetização da mãe. Com relação a esse aspecto, crianças negras parecem ter sido especialmente prejudicadas pelas instituições educacionais do século XIX que, como

21. Rosa Fátima De Souza, *Templos de civilização: a implantação da escola primária graduada no Estado de São Paulo, 1890-1910* (Unesp, 1998), 109.

visto, inicialmente preteriram mulheres do acesso ao ensino, além de impedirem a educação de escravizados, contribuindo decisivamente para que mulheres negras constituíssem o grupo populacional com as maiores taxas de analfabetismo em São Carlos em 1907. Em termos contrafactuais, uma maior educação das crianças negras de sexo feminino durante o Império poderia ter resultado em um hiato menor entre as taxas de matrícula escolar entre a população branca e negra no início da República.

De modo geral, as estimativas realizadas demonstram a existência de uma demanda substancial por educação por parte de famílias negras. Esse resultado sugere o erro incorrido nas conclusões por parte de interpretações clássicas sobre o tema que, entre outras coisas, afirmaram que nessas famílias os “pais ignoravam a existência e a utilidade das escolas”.²²

22. Fernandes, *A integração*, 169.

Capítulo 3

A população negra e o trabalho: o legado da escravidão e as ocupações

No capítulo anterior, foram analisados os dados de acesso à educação primária entre negros em São Carlos, na segunda metade da década de 1900. As estimativas mostraram que uma parte relevante da população negra em idade escolar —cerca de 12% das crianças matriculadas nas escolas isoladas—, conseguiu acesso ao ensino primário no município. O fato da matrícula no ensino primário das crianças negras não divergir significativamente dos níveis observados para a população branca pode indicar que as condições econômicas de pelo menos parte de ex-escravos e descendentes não era muito diferente das observadas para imigrantes e grupos de brasileiros brancos. Qual era a situação econômica das famílias negras que teria possibilitado o seu investimento em educação?

Neste capítulo, apresenta-se uma análise do mercado de trabalho em São Carlos em 1907, com o intuito de sugerir respostas para essa pergunta. O mercado de trabalho no Oeste paulista após Abolição foi marcado por uma grande variedade de contratos e condições de trabalho. Como visto no Capítulo 1, o crescimento da lavoura cafeeira, viabilizado por melhorias nas técnicas de cultivo, redução do custo de transporte e cultivo de terras mais férteis, ameaçava aumentar o custo do trabalho no contexto de crise do regime escravista. A imigração europeia subsidiada amenizou o problema de mão de obra para os fazendeiros, mas, como argumentado, o sucesso da empreitada dos cafeicultores paulistas

foi apenas parcial, de modo que tanto trabalhadores imigrantes como nacionais parecem ter obtido um ganho material nas duas décadas que se seguiram à Abolição. Por outro lado, o legado da escravidão não desapareceu em meio às transformações econômicas, sociais e demográficas ocorridas no Oeste paulista no período. Embora ex-escravos e descendentes passassem a usufruir de um “ideal de liberdade”, com a possibilidade de livre mobilidade, perspectiva de acesso à terra e institucionalização e reconhecimento social de laços familiares, é certo que aspectos estruturais da sociedade escravocrata, como o racismo e as hierarquias nas relações interpessoais ainda se faziam presentes.¹

A complexidade da relação entre mercado de trabalho e cor da pele no pós-abolição em São Carlos é confirmada pelos resultados deste capítulo. Por um lado, os resultados mostram que houve clara predominância de negros em ocupações associadas ao trabalho doméstico e a menor renda, realidade que se manteve mesmo para as gerações mais jovens. Por outro lado, as estimativas aqui apresentadas reforçam os resultados já obtidos por Karl Monsma de que a posição de negros na estrutura ocupacional não era muito distinta daquela observada para imigrantes. Imigrantes também estavam presentes em ocupações domésticas, enquanto negros também se faziam presentes entre colonos e dominavam em ocupações artesanais. Na conclusão do capítulo, argumenta-se que essa aparente contradição é explicada por fatores estruturais e conjunturais e é coerente com o argumento principal desta tese de que a prosperidade do café no Oeste paulista teve benefícios temporários para a população negra.

3.1 Mercado de trabalho e as ocupações em São Carlos

Grande parte da literatura sobre o mercado de trabalho paulista no pós-abolição tem enfatizado a precariedade da situação ocupacional de ex-escravos e descendentes. Estudos feitos para a capital do estado comumente destacam os elevados percentuais de imigrantes na indústria e a relativa dificuldade enfrentada pelos

1. Mattos, *Das cores*, Cap.1-4; Monsma, *A reprodução*, 58-67.

negros para se inserirem no mercado formal.² No interior de São Paulo, o quadro não mudaria substancialmente. Na zona rural, o contrato de “colono”, principal forma de organização do trabalho na produção cafeeira, era dominante. Neste tipo de contrato, resultante da evolução do sistema de parceria, o trabalhador e sua família ficavam responsáveis pelo tratamento de determinado número de cafezais.³ Dependendo do tamanho da família, da qualidade da terra e do preço do café, os colonos poderiam auferir bons ganhos em sua atividade. Em geral, porém, os trabalhos clássicos e outros mais recentes sobre a agricultura paulista retrataram o acesso aos contratos de colonato como quase exclusivamente limitado a trabalhadores imigrantes. Ex-escravos e descendentes, por sua vez, estariam limitados a funções domésticas, de agregados e camaradas, marcadas pela maior dependência de grandes proprietários e limitadas possibilidades de se converter em ganho material relevante para as gerações futuras.⁴

Essas disparidades ocupacionais em relação à população branca e imigrante também é geralmente explicada pela discriminação contra o trabalhador nacional e o trabalhador negro, em particular. Os autores que viram na discriminação a principal causa das desigualdades observadas no mercado de trabalho apoiam-se no fato de que havia uma crença generalizada de que o trabalhador europeu era superior ao trabalhador nacional.⁵ De fato, estudos anteriores apresentaram evidências abundantes de que a elite agrária da época considerava o trabalhador nacional “racialmente inferior”, “excessivamente moroso” e “vadio”.⁶

Essas conclusões e resultados gerais, contudo, devem ser qualificados à luz de contribuições recentes e as apresentadas nesta tese. Em primeiro lugar, como

2. Andrews, por exemplo, cita o fato de que na capital paulista, em 1893, os percentuais de estrangeiros empregados no comércio, fábricas, transporte e artesanato eram de 72%, 79%, 81% e 86%, respectivamente. Andrews, *Negros e Brancos*, 111. Ver também Jacino, “O negro”, 127-178

3. Maria Teresa Schorer Petrone, “Imigração assalariada”, *História geral da civilização brasileira* 2 (1967): 274; Souza, “From Bonded Laborers to Educated Citizens? Immigration, Labor Markets, and Human Capital in São Paulo, Brazil (1820-2010)”, Capítulo 2.

4. Andrews, *Negros e Brancos*, 110-111; Dean, *Rio Claro*, 165; Holloway, *Imigrantes*, 259.

5. Nathaniel H Leff, “Economic development in Brazil, 1822-1913”, em *How Latin America Fell Behind: Essays on the Economic Histories of Brazil and Mexico, 1800-1914* (Stanford University Press, 1998), 39.

6. Celia Maria Marinho De Azevedo, *Onda negra, medo branco: o negro no imaginário das elites—século XIX*, vol. 6 (Annablume, 1987); Lúcio Kowarick, *Trabalho e vadiagem: a origem do trabalho livre no Brasil* (Editora Brasiliense, 1987).

o período posterior à Abolição é marcado pela ausência de fontes censitárias com informações sobre cor da pele, a pesquisa sobre o mercado de trabalho da época tem sido feita a partir de fontes com recortes geográficos e demográficos muito estritos. Por exemplo, um importante estudo sobre a capital paulista no período 1912-1920 utiliza informações de boletins de ocorrência para obter dados sobre a ocupação da população negra no período.⁷ Porém, uma vez que pessoas negras sofriam desproporcionalmente com a repressão policial, os boletins de ocorrência a qualquer tempo resultariam em sobre-representação de negros em qualquer profissão. A maior participação relativa de negros nos boletins não invalida em si os achados desse tipo de evidência, visto que a ausência de negros em diversos ofícios em uma fonte documental onde negros são potencialmente sobre-representados é de fato um indicador de segregação ocupacional. No entanto, a presença de auto-seleção nas fontes pesquisadas limita a inferência acerca de ocupações nas quais negros aparecem sobre-representados. A auto-seleção também consiste em um problema em trabalhos que se limitam a estudar um pequeno número de empresas.⁸

Outro problema encontrado na historiografia que se relaciona às fontes utilizadas é a generalização para todo o estado de São Paulo de resultados encontrados para a capital paulista. Florestan Fernandes e George Andrews, por exemplo, ao identificarem elevado nível de segregação ocupacional na capital, afirmam que essa segregação era replicada no campo, com trabalhadores nacionais confinados às regiões menos prósperas da cafeicultura, como o Vale do Paraíba, enquanto imigrantes dominavam as terras férteis do Oeste paulista.⁹ Na verdade, esses trabalhos não apresentam dados sobre essas regiões rurais a que se referem, limitando-se a generalizar evidências da região da capital de São Paulo, uma área urbana e bastante atípica em relação a outras partes do estado em termos de mercado de trabalho.

O recente trabalho de Karl Monsma ofereceu uma importante contribuição para superar as limitações da historiografia sobre o pós-abolição no conjunto de

7. Jacino, "O negro", 127-178.

8. Andrews, *Negros e Brancos*, Cap. 4; Ball, "Inequality in São Paulo's".

9. Andrews, *Negros e Brancos*, 110-111; Fernandes, *A integração*, 20.

São Paulo, ao examinar os dados censitários de São Carlos para 1907.¹⁰ O principal achado de Monsma em relação ao mercado de trabalho é que, assim como ocorria com os imigrantes, a ocupação mais comum para trabalhadores negros era a de colono.¹¹ Nesta seção, procura-se estender a análise de Monsma sobre a estrutura ocupacional de São Carlos em 1907. Em linhas gerais, a análise corrobora os principais resultados obtidos pelo autor, mas dois novos aspectos são estudados. Em primeiro lugar, os dados de ocupação do censo municipal de São Carlos são utilizados para avaliar criticamente a hipótese de Florestan Fernandes e George Andrews de que, no pós-abolição, ex-escravos e descendentes diminuíram a oferta de trabalho em ocupações vistas como típicas da escravidão.¹² Em segundo lugar, elabora-se um índice de ocupações como forma de mensurar a representação de negros em determinada ocupação em relação à representação de negros na população geral. O uso desse índice permitirá definir uma classificação ocupacional particularmente útil para uma avaliação mais geral sobre a posição de ex-escravos e descendentes no mercado de trabalho de um importante município cafeeiro de São Paulo.

Na próxima seção, detalha-se a construção do índice de equilíbrio racial e os principais resultados obtidos com a adoção dessa medida são analisados à luz da historiografia sobre o pós-abolição.

3.1.1 Classificação e resultados

O indicador que será utilizado nesta seção baseia-se no índice desenvolvido por Roger Ransom e Richard Sutch, em seu estudo sobre a economia do Sul dos Estados Unidos no pós-emancipação.¹³ O propósito do índice, que será chamado aqui de Índice de Equilíbrio Racial (IER), é mensurar o grau de sub e sobre-representação de determinado grupo demográfico em uma ocupação no mercado de trabalho de São Carlos com as informações disponíveis do censo municipal de

10. Monsma, *A reprodução*.

11. *Ibid.*, 94.

12. Andrews, *Negros e Brancos*, 138-139; Fernandes, *A integração*, 13-14.

13. Roger L Ransom e Richard Sutch, *One kind of freedom: The economic consequences of emancipation* (Cambridge University Press, 2001), 220-231.

1907.¹⁴

O índice de equilíbrio racial tem seu valor definido em um intervalo de -1 a 1. Valores próximos a -1 indicam sub-representação de negros, enquanto valores próximos de 1 sugerem sobre-representação desse grupo demográfico na ocupação. Ainda seguindo Ransom e Sucht, foram definidas 5 categorias de ocupações de acordo com o valor do índice de equilíbrio racial:

1. Brancos excluídos ($IER \geq 0,8$): ocupações em que a participação relativa de brancos é muito próxima ou igual a zero.
2. Dominância negra ($0,2 \leq IER < 0,8$): ocupações em que há sobre-representação de negros, mas não exclui brancos.
3. ocupações abertas ($-0,2 \leq IER < 0,2$): ocupações em que não há dominância de nenhum grupo em particular.
4. Dominância branca ($-0,8 \leq IER < -0,2$): ocupações em que há sobre-representação de brancos, mas não exclui negros.
5. Negros excluídos ($IER < -0,8$): ocupações em que a participação relativa de negros é muito próxima ou igual a zero.

Por consistência e para viabilizar comparações com o caso do Sul dos Estados Unidos, adotam-se aqui os mesmos intervalos propostos por Ransom e Sucht. Diferentemente desses autores, porém, considera-se na análise abaixo apenas as ocupações com cinco ou mais trabalhadores.¹⁵ Em que pese a arbitrariedade envolvida em qualquer classificação, argumenta-se que a utilidade da classificação aqui proposta —para comparar e classificar ocupações de forma consistente—supera os custos impostos pela subjetividade dos critérios adotados.

Os dados dos primeiros três grupos para a população masculina são apresentados na Tabela 3.1. Analisando primeiro o grupo (Brancos Excluídos), pode-se

14. Os detalhes técnicos do cálculo do índice de equilíbrio racial encontram-se no Apêndice B.

15. Esse critério busca eliminar ocupações que tinham pouca importância para a economia da época ou que resultam de uma resposta algo idiossincrática à pergunta do recenseador sobre ocupação, como, por exemplo, “Matador de Formigas” e “Chefe”, ambas com apenas 1 trabalhador registrado em todo censo.

notar que, entre as ocupações em que brancos são relativamente excluídos, há predominância de ocupações ligadas à atividade doméstica, como cozinheiro, criado e caseiro. A ocupação de criado, em particular, representa a maior proporção de homens negros (54%), alcançando um valor mais de três vezes superior à proporção de negros em todas as ocupações (13,5%), o que levou o índice de equilíbrio racial para seu valor máximo.

Tabela 3.1: IER das ocupações dos grupos 1 a 3, População negra masculina, São Carlos, 1907

Grupos	Ocupações	Total	Proporção de negros	IER
Branco excluídos				
	Cozinheiro	21	0,48	1.0
	Criado	13	0,54	1.0
	Doceiro	5	0,4	0.93
	Caseiro	5	0,4	0.93
	Portador	11	0,36	0.86
	Camarada	1156	0,35	0.83
Dominância negra				
	Agregado	6	0,33	0,79
	Carroceiro	158	0,27	0,61
	Empregado	345	0,22	0,43
	Maquinista	50	0,22	0,42
	Pintor	19	0,21	0,38
	Cocheiro	49	0,2	0,35
	Soldado	25	0,2	0,33
	Carpinteiro	89	0,19	0,29
	Pedreiro	150	0,19	0,27
Ocupações abertas				
	Empreiteiro	84	0,13	-0,02
	Operário	46	0,13	-0,03
	Mecânico	9	0,11	-0,14
	Serrador	9	0,11	-0,14
	Seleiro	28	0,11	-0,17
	Empregado Cia Paulista	135	0,1	-0,19
	Colono	4950	0,1	-0,2

Fonte: São Carlos, Recenseamento de 1907.

Notas: Dados referentes à população masculina com idade igual ou superior a 18 anos. Todos os nomes de ocupação foram padronizados para o substantivo masculino e ocupações equivalentes como “Carreiro” e “Carroceiro” foram agregadas. As denominações “Empregado”, “Operário” são bastante gerais de modo que não fica claro a que se referem e, possivelmente, reuniram grupos heterogêneos de trabalhadores. Optou-se, mesmo assim, por apresentar os nomes das ocupações conforme aparecem no censo.

Os dados da Tabela 3.1, que se referem à população masculina adulta (maiores de 18 anos), omitem alguns aspectos relevantes sobre a ocupação de criado no que diz respeito à composição de sexo, idade, cor da pele e nacionalidade. De acordo com o censo municipal, 229 pessoas exerciam a ocupação de criado em São Carlos em 1907, em sua maioria mulheres (186). Havia adultos e até idosos nessa ocupação, como nove indivíduos com idade igual ou superior a 60 anos. Além disso, a grande maioria dos criados era composta por crianças e adolescentes, com 62,4% dos indivíduos registrados como criados no censo tendo idade inferior a 18 anos. Se esses indivíduos com idade inferior a 18 anos forem levados em conta, a sobre-representação de negros nessa ocupação se reduz, passando estes últimos a responder por 43,6% dos criados, em vez de 54% como consta na Tabela 3.1. Mesmo assim, a sobre-representação seria grande o suficiente para classificar a ocupação de criado como sendo exclusivamente negra, de acordo com o índice de equilíbrio racial.

Em relação à cor da pele, a sobre-representação de pretos entre criados é maior do que a sobre-representação de mulatos, como mostra a Figura 3.1, que apresenta os percentuais de negros e imigrantes em ocupações domésticas. De fato, quando se calcula o índice de equilíbrio racial separadamente para pretos e mulatos, considerando ambos os sexos, nota-se que criado é uma ocupação listada como exclusiva de preto (Índice de Equilíbrio Racial de 0,91), mas aberta entre mulatos (Índice de Equilíbrio Racial de -0,03). Essa diferença ocorre em função do baixo percentual de mulheres mulatas entre criadas (1%). Mulatos constituíam 2,4% da população ocupada e 2% entre criados, enquanto pretos somavam 10,9% da população e 36% entre criados.

Para se avaliar a dimensão de nacionalidade, é ilustrativo considerar duas definições. A primeira, adotada pelo censo municipal de 1907, é a de *jus soli*, em que a nacionalidade é determinada pelo local de nascimento. Na segunda definição, atribui-se a nacionalidade a um indivíduo de acordo com sua ascendência, caso chamado de *jus sanguinis*. Essa diferença é importante, pois considerar apenas a definição adotada pelo censo poderia levar a subestimar o número de imigrantes em atividades domésticas, principalmente aquelas com grande número de crianças e adolescentes, como a ocupação de criado.

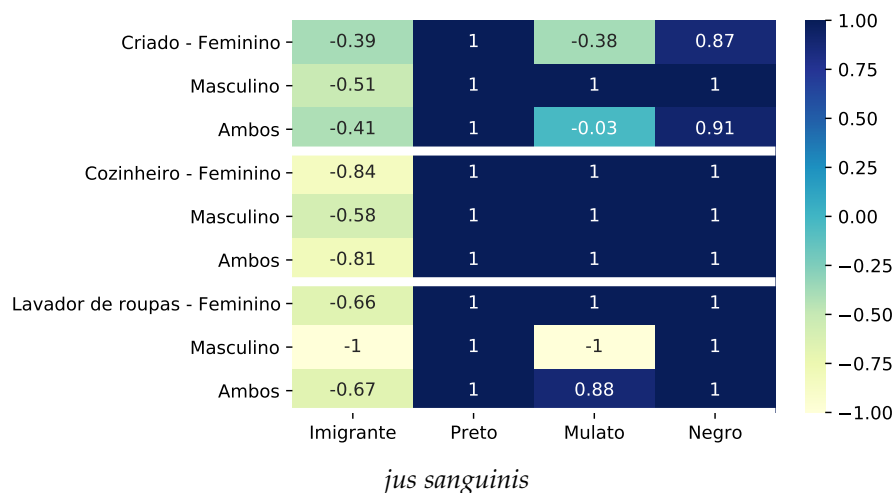
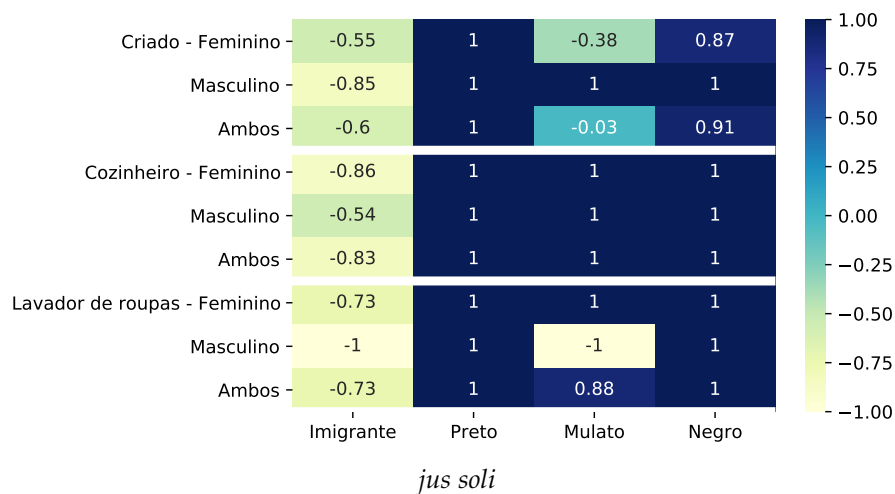


Figura 3.1: *IER* para imigrantes e negros em ocupações domésticas, São Carlos, 1907

Fonte: São Carlos, Recenseamento de 1907.

Notas: O *IER* varia de -1 a 1. Quanto mais perto de 1, maior é a sobre-representação na ocupação e mais escura é a célula. Por exemplo, mulheres imigrantes (*jus solis*) são efetivamente excluídas da ocupação de cozinheira ($IER=-0,86$), enquanto essa é uma ocupação classificada como exclusiva para mulheres pretas ($IER=1$). Como em toda a tese, “negro” é a agregação de “preto” e “mulato”. Ver texto para a classificação do *IER*.

Fica claro que imigrantes também estavam presentes na ocupação de criado, mas em proporção menor do que sua presença na população (ver comparação

entre a primeira e segunda Tabela na Figura 3.1). Considerando apenas nascidos no exterior (*jus soli*), homens estrangeiros estavam efetivamente excluídos da ocupação de criado (Índice de Equilíbrio Racial de -0,85). Já a participação de mulheres estrangeiras entre criados é maior, embora estas também estejam sub-representadas na categoria: elas correspondiam a 66,4% da população feminina ocupada e 38% entre criadas. É notável perceber a mudança no valor do índice de equilíbrio racial (de -0,55 para -0,39, no caso das mulheres) para a ocupação de criado quando se consideram filhos de imigrantes como estrangeiros (*jus sanguinis*). Esse dado indica que muitos imigrantes de primeira geração trabalhavam como criados em São Carlos, principalmente mulheres. Em particular, o painel b da Figura 3.1 mostra que mulheres imigrantes tinham representação entre criadas equivalente à representação observada para mulheres mulatas (índice de equilíbrio racial de -0,39 para mulatas e -0,38 para mulheres imigrantes). Quando se consideram as ocupações de cozinheiro e lavadeira, contudo, há pouca diferença pelo critério de nacionalidade.

Cozinheiro é a segunda ocupação com maior percentual de mulheres e homens negros, como mostram as Tabelas 3.1, discutida acima, e a Tabela 3.2, que apresenta os resultados separadamente para a população feminina. Havia 196 adultos trabalhando como cozinheiros em São Carlos em 1907, dos quais 175 eram mulheres, tornando essa ocupação tipicamente feminina. A dominância feminina, já observada para a ocupação de criado e agora para cozinheiros, aparece de forma ainda mais pronunciada na profissão de lavadeira, como pode ser visto na Tabela 3.2. Havia 121 adultos na ocupação de lavador de roupas, com mulheres respondendo pela quase totalidade desse número (119). Diferentemente do trabalho de criado, não há diferença de sobre-representação entre pretos e mulatos na função de cozinheiro. É importante notar ainda que, ao contrário do que ocorre com os criados, as ocupações de cozinheiro e lavadeira eram dominadas por adultos. A idade mediana de lavadeiras e cozinheiros era de 35 anos e 30 anos, respectivamente, enquanto que para criados a idade mediana era de 14 anos.

Uma outra ocupação que aparece como exclusiva de negros é de camarada. Karl Monsma mostra que, corroborando a literatura, os dados de São Carlos exibem uma grande presença de pretos e mulatos na ocupação de camarada. O autor

Tabela 3.2: IER das ocupações dos grupos 1 a 3, População negra feminina, São Carlos, 1907

Grupo	Ocupação	Total	Proporção de negras	IER
Branco excluídos				
	Cozinheira	175	0,68	1
	Lavadeira	119	0,45	1
	Doceira	8	0,50	1
	Seleira	5	0,40	0,96
	Empregada	165	0,36	0,88
	Camarada	338	0,36	0,88
	Criado	73	0,36	0,87
Ocupações abertas				
	Engomadora	18	0,28	0,65
	Pedreira	11	0,18	0,27
	Empreiteira	65	0,17	0,21
Dominância negra				
	Ferreira	8	0,12	-0,04
	Cocheira	9	0,11	-0,13

Fonte: São Carlos, Recenseamento de 1907.

Notas: Dados referentes à população feminina com idade igual ou superior a 18 anos. Para detalhes sobre os nomes das ocupações, ver nota da Tabela 3.1.

pondera, no entanto, que a maior presença de negros nessa ocupação é uma característica identificável entre homens casados. Monsma demonstra que, entre homens sem família, os percentuais de estrangeiros na função de camarada era próxima àquela observada para pretos e mulatos.¹⁶

Por qual razão então o índice de equilíbrio racial classifica a ocupação de camarada como exclusiva de negros? A resposta reside nos critérios etário e demográfico adotados. Os percentuais de imigrantes na ocupação de camarada só são comparáveis à população negra quando se consideram, como fez Monsma, portugueses e italianos do sexo masculino com idade entre 15 e 60 anos não associados a uma família. Ocorre que esta classe da população (homens não casados entre 15 e 60 anos) abrange menos de 10% do total de camaradas listados no censo. Quando se considera toda a população adulta, independentemente de situação matrimonial, constata-se que 36% dos camaradas eram negros, contrastando com a participação de negros na população adulta ocupada, que era

16. Monsma, *A reprodução*, 96.

apenas de 12,3%. Por essa razão, o índice de equilíbrio racial classifica camarada como sendo uma ocupação exclusivamente negra. Além disso, entre camaradas, não se observa diferença entre pretos e mulatos, como visto anteriormente para ocupações domésticas (ver Tabelas C.6 e C.8 no Apêndice C).

Um fato que contribui para que negros sejam desproporcionalmente representados entre camaradas são as diferenças nas razões entre homens e mulheres nesta ocupação no caso de imigrantes e negros. Enquanto entre imigrantes existiam quase 4 homens por mulher na ocupação de camarada, entre negros essa razão era de 3,4. Ou seja, proporcionalmente, havia mais mulheres negras na ocupação de camarada do que mulheres imigrantes. Dessa forma, qualquer medida de concentração ocupacional que se restrinja a analisar a população masculina irá subestimar a sobre-representação de negros entre camaradas. Coerente com esse fato, o índice de equilíbrio racial calculado separadamente para mulheres negras (0,88, Tabela 3.2) é maior do que o índice de equilíbrio racial calculado para homens negros (0,83, Tabela 3.1). Paralelamente, em larga medida, a menor razão de homens e mulheres entre camaradas negros é explicada pela baixa proporção de homens e mulheres entre camaradas mulatos. Entre mulatos, a razão sexual é de apenas 1,4 e a proporção de mulatos entre camaradas é maior do que a de pretos, visto que o índice de equilíbrio racial nesta ocupação para mulatos é de 1 e para pretos é de 0,89, ou seja, proporcionalmente existem mais mulatos entre camaradas que pretos.

Obviamente, a participação da população negra no mercado de trabalho em São Carlos, no início do século XX, não estava limitada às ocupações domésticas e de camarada. A Tabela 3.1 apresenta outras profissões dominadas por negros, e ocupações abertas, em que não há dominância clara de nenhum grupo demográfico. Entre as funções dominadas por negros, aparecem as ligadas à construção civil, como pintor, carpinteiro e pedreiro. O fato de essas ocupações exigirem certo nível de habilidades (*skills*) e serem dominadas por negros indica a presença de trabalhadores qualificados entre a população negra. Em sua pesquisa para a capital de São Paulo, Ramatis Jacino também encontrou um número expressivo (cerca de 12%) de negros entre carpinteiros.¹⁷

17. Jacino, "O negro", 153.

Com os dados do censo de São Carlos, é possível desagregar as ocupações por cor da pele. Nota-se, assim, que entre as três ocupações citadas anteriormente (pintor, carpinteiro e pedreiro), apenas uma não era dominada claramente por mulatos. Na função de carpinteiro, o índice de equilíbrio racial é de 1 para mulatos e apenas 0,02 para pretos, revelando uma sobre-representação muito maior de mulatos nesta ocupação do que de pretos. Nesta profissão, negros correspondiam a 22,1% do total, dos quais 11,5% eram mulatos, proporção esta significativa tendo em vista que a população de mulatos era de apenas 2,6% do total. A ocupação de pintor também tinha predominância de mulatos, com índice de equilíbrio racial de 0,91 para este grupo e 0,31 para pretos. Por sua vez, a ocupação de pedreiro apresentou maior equilíbrio, com o índice de equilíbrio racial de aproximadamente 0,3 para os dois grupos.

As ocupações classificadas pelo índice de equilíbrio racial como abertas incluem um grupo heterogêneo de atividades e são caracterizadas pela ausência de domínio de qualquer grupo demográfico em particular. Entre essas ocupações, aparece o colonato. Assim, corroborando a interpretação de Karl Monsma, o índice de equilíbrio racial indica que não havia predomínio de trabalhadores imigrantes e brasileiros brancos em relação a trabalhadores negros na atividade de colono.¹⁸

A Tabela 3.3 apresenta a lista de ocupações dominadas por brancos. Em outras palavras, nessas profissões a proporção de brancos é maior do que a representação desse grupo na população ocupada, mas não em uma proporção que se poderia classificá-las como exclusivas de brancos. De particular interesse, devido a seu peso demográfico, é a ocupação de lavrador. Como foi salientado no Capítulo 2, o termo lavrador no censo de São Carlos de 1907 parece denominar o trabalhador que cultivava a própria terra, fosse ela pequena roça, média ou grande lavoura. Coerente com essa hipótese é o fato de 70% da população adulta que se declarou como lavrador ter também declarado que era proprietária de terra, um percentual muito maior do que o valor observado para a população ocupada total (17,8%). Assim, o fato de negros serem relativamente pouco presentes entre lavradores é sintoma de um fenômeno maior, qual seja, a sub-representação de negros entre

18. Monsma, *A reprodução*, 96.

proprietários de terra. De acordo com o censo de São Carlos de 1907, 13,5% da população negra era proprietária de terra. Esse percentual era muito menor do que os 36% observados entre brasileiros brancos. O percentual de pretos com ocupação declarada e com propriedade de terra é de 9,8%, enquanto o mesmo percentual para mulatos é de 16,4%. Como visto no Capítulo 1, os dados sobre propriedade de terra no censo, porém, devem ser analisados com cautela, já que não está claro qual definição de propriedade foi utilizada.

Tabela 3.3: IER das ocupações dominadas por não-negros, População masculina negra, São Carlos, 1907

Ocupação	Total	Proporção de negros	IER
Trabalhador	60	0,10	-0,22
Lavrador	988	0,09	-0,25
Ferreiro	44	0,09	-0,28
Caixeiro	13	0,08	-0,37
Foguista	13	0,08	-0,37
Industrial	19	0,05	-0,55
Feitor	21	0,05	-0,59
Administrador	123	0,04	-0,65
Marcineiro	29	0,03	-0,70
Alfaiate	44	0,02	-0,80

Fonte: São Carlos, Recenseamento de 1907.

Notas: Dados referentes à população masculina com idade igual ou superior a 18 anos. Para detalhes sobre os nomes das ocupações, ver nota da Tabela 3.1.

Se a presença diminuta de negros em grupos ocupacionais como lavradores e industriais é explicada pelo baixo volume de recursos materiais em posse dessa população, seria a sub-representação de ex-escravos e descendentes em outras ocupações explicada pelas diferenças em habilidades e qualificações, isto é, em capital humano? Os resultados apresentados na Tabela 3.4, que lista as funções em que negros eram efetivamente excluídos, sugerem uma resposta para essa pergunta. Como pode ser observado, negros estavam excluídos de certas funções artesanais (sapateiro, padeiro e relojoeiro, por exemplo) e de todas as profissões que exigiam níveis maiores de alfabetização ou acumulação prévia de recursos. Não há negros, pretos ou mulatos, entre professores, guarda-livros, telegrafistas,

médicos, farmacêuticos, advogados ou dentistas. Entre negociantes, ocupação urbana que tende a exigir proficiência mínima de leitura e numeramento, além de recursos acumulados, o percentual de negros era de apenas 2%.

Os resultados analisados nesta seção permitiram avaliar um ângulo do legado da escravidão para a população negra. Os resultados indicam que os baixos níveis de educação e capital da população negra, resultado direto da experiência do cativo, limitaram as possibilidades profissionais de pretos e mulatos nas décadas que se seguiram à Abolição. Essa conclusão é reforçada pelo fato de que negros dominavam profissões artesanais e qualificadas, como maquinistas e carpinteiros (aba “Dominância negra”, Tabela 3.1) e competiam diretamente com imigrantes e brancos brasileiros em outras, como mecânicos e seleiros (aba “Ocupações abertas” Tabela 3.1). No entanto, a possibilidade de competição desaparecia quando se considera ocupações que demandavam qualquer forma de capital. Na verdade, as diferenças verificadas entre os dois subgrupos da população negra, pretos e mulatos, são evidências de que a realidade ocupacional de ex-escravos e descendentes no pós-abolição deveu-se em grande parte ao acesso praticamente nulo à educação, terra e capital durante a escravidão. Com efeito, as vantagens da população mulata em educação e patrimônio, vistas no Capítulo 1 (Tabelas 1.3 e 1.4), parecem se refletir aqui em melhores ocupações para esse grupo. Uma possível razão para essa tendência é o fato de que a ascendência mista da população mulata logrou diminuir a força do passivo do passado escravocrata.

Reconhecer a influência da educação, do capital e da terra sobre o mercado de trabalho no pós-abolição não implica em diminuir a importância da discriminação racial. No entanto, a efetividade da discriminação pode ter variado nas diferentes regiões do país, a depender das condições materiais enfrentadas por trabalhadores e empregadores. Em particular, é possível que em vista do quadro de escassez relativa do trabalho, a força da discriminação no mercado de trabalho no interior de São Paulo em meio ao rápido crescimento da economia cafeeira não tenha se reproduzido com a mesma intensidade que na capital do estado. Um indicativo de como a competição pode ter limitado a influência da discriminação é encontrada nos altos percentuais de negros entre colonos e outras ocupações artesanais.

Tabela 3.4: IER das ocupações em que negros são excluídos, População masculina negra, São Carlos, 1907

Ocupação	Total	Proporção de negros	IER
Negociante	297	0,02	-0,85
Sapateiro	75	0,00	-1,00
Jardineiro	5	0,00	-1,00
Verdureiro	5	0,00	-1,00
Dentista	5	0,00	-1,00
Açogueiro	5	0,00	-1,00
Relojoeiro	7	0,00	-1,00
Advogado	7	0,00	-1,00
Fazendeiro	8	0,00	-1,00
Folheiro	9	0,00	-1,00
Telegrafista	9	0,00	-1,00
Podador	9	0,00	-1,00
Mascate	10	0,00	-1,00
Médico	10	0,00	-1,00
Diretor	10	0,00	-1,00
Ajudante	11	0,00	-1,00
Chefe de Estacao	14	0,00	-1,00
Empregado Público	14	0,00	-1,00
Fogueteiro	18	0,00	-1,00
Carniceiro	18	0,00	-1,00
Artista	20	0,00	-1,00
Guarda Livros	21	0,00	-1,00
Cervejeiro	23	0,00	-1,00
Professor	24	0,00	-1,00
Farmacêutico	24	0,00	-1,00
Padeiro	24	0,00	-1,00
Hortelão	26	0,00	-1,00
Barbeiro	33	0,00	-1,00
Comerciante	69	0,00	-1,00
Estudante	5	0,00	-1,00

Fonte: São Carlos, Recenseamento de 1907.

Notas: Dados referentes à população masculina com idade igual ou superior a 18 anos. Para detalhes sobre os nomes das ocupações, ver nota da Tabela 3.1.

3.2 Conclusão

Neste capítulo analisou-se, por meio de dados sobre a estrutura ocupacional, a participação da população de ex-escravos e descendentes no mercado de trabalho. Viu-se que a relação entre cor da pele e ocupação em São Carlos, menos de 20 anos depois da abolição, era complexa. Por um lado, persistiam tendências observadas durante a escravidão, como a predominância de trabalhadores negros em tarefas domésticas. Em particular, mulheres pretas eram desproporcionalmente representadas entre lavadeiras, cozinheiros e criados. A dominância de trabalhadores negros em ocupações domésticas, bem como a permanência de trabalhadores negros jovens nessas ocupações, constituem fatos que contrariam afirmações feitas pela literatura em relação ao abandono por parte de ex-escravos de atividades similares à escravidão nos anos que se seguiram à Lei Áurea.¹⁹ É possível que tenha ocorrido uma redução da oferta global de trabalho por parte de ex-escravos no pós-abolição, como no Sul dos Estados Unidos depois da Guerra Civil, mas não há evidência de que essa oferta tenha se concentrado em um grupo particular de ocupações.

No caso americano, houve forte queda da oferta de trabalho da população negra depois da Guerra Civil.²⁰ Estima-se que o total de trabalhadores na agricultura tenha caído entre 17 e 24 pontos percentuais depois da emancipação e que o número de horas trabalhadas por dia tenha tido redução mínima de 9 pontos percentuais.²¹ É importante notar, contudo, que essa redução da oferta de trabalho por parte da população afro-americana no pós-emancipação não redundou em uma mudança significativa na distribuição ocupacional que caracterizava a escravidão.²² É provável ainda que diferenças de capital humano entre a população branca e a população de ex-escravos limitavam as possibilidades dos últimos de se engajarem em ocupações que exigiam certo nível de habilidade.²³ Não há razão para crer que o caso brasileiro seja diferente.

19. Andrews, *Negros e Brancos*, 138-139; Fernandes, *A integração*, 13-14.

20. Higgs, *Competition*, 40.

21. Ransom e Sutch, *One kind*, 45.

22. Higgs, *Competition*, 60.

23. Ransom e Sutch, *One kind*, 23-39.

Pode-se agora voltar à questão que abriu este capítulo. Afinal, quais condições econômicas teriam permitido a matrícula escolar de crianças negras em São Carlos na segunda metade da década de 1900? Para responder essa pergunta, é conveniente lembrar aqui a ocupação dos pais das crianças negras matriculadas nas escolas primárias. Com efeito, entre os pais das crianças negras, há predomínio de lavradores e colonos, já que 12 dos 24 pais de crianças negras dos quais se tem informações eram colonos (9) ou lavradores (3). Significativamente, apenas dois pais recebiam a denominação de empregado que, como visto, era uma ocupação dominada por negros, mas aqui aparece sub-representada entre pais de alunos. É possível, portanto, que o acesso ao colonato e à ocupação de lavrador tenham criado condições para que negros colocassem seus filhos nas escolas. A prosperidade cafeeira e a consequente concorrência por mão de obra podem ter criado as condições de melhores oportunidades de trabalho para a população negra, o que teria permitido o investimento por parte das famílias negras em capital humano, a despeito do legado escravocrata ainda se fazer presente na estrutura ocupacional. Teriam essas mesmas condições econômicas perdurado tempo suficiente para permitir que parte da renda do trabalho tenha se convertido na acumulação de capital? No Capítulo 4, avalia-se o acesso à terra por parte da população negra, com o objetivo de responder esta pergunta.

Capítulo 4

A população negra e a terra: aspectos quantitativos e qualitativos do acesso à propriedade rural

No Capítulo 1, argumentou-se que o crescimento da economia cafeeira no Oeste paulista, combinado à redução do fluxo migratório na década de 1900 em relação ao número de imigrantes da década anterior, teria levado a um quadro de relativa escassez de mão de obra na agricultura. A falta de braços e a crescente demanda por trabalho teriam viabilizado melhorias materiais para trabalhadores, que podem ter obtido ganhos monetários e melhores termos nos contratos nos primeiros anos do século XX. Essas condições aumentaram o poder de barganha e favoreceram tanto trabalhadores imigrantes quanto nacionais, inclusive os descendentes de escravizados.

Vários autores notaram essas condições do mercado de trabalho na agricultura paulista da época, embora geralmente com foco nos imigrantes.¹ Thomas

1. Carlos de Almeida Prado Bacellar, "O apogeu do café na Alta Mogiana", em *Na Estrada do Anhanguera: uma visão regional da história paulista* (Humanitas Publicações FFLCH/USP, 1999), 150; Marília Dalva Klaumann Cánovas, "A emigração espanhola e a trajetória do imigrante na cafeicultura paulista: o caso de Villa Novaes, 1880-1930." (diss. de mestrado, Universidade de São Paulo, 2001), 204; Herbert S Klein, *A imigração espanhola no Brasil* (Fapesp, 1994), 65-6; Herbert S Klein, "A integração dos imigrantes italianos no Brasil, na Argentina e Estados Unidos", *Novos Estudos CEBRAP* 25 (1989): 95-117.

Holloway, por exemplo, em seu estudo sobre os imigrantes europeus, registrou a gradual melhoria das condições contratuais conseguidas pelos trabalhadores imigrantes na agricultura paulista de 1860 em diante.² O sistema de colonato tornar-se-ia dominante no Oeste paulista a partir de 1880.³ Mesmo em períodos de crise, como a que se seguiu à queda dos preços do café na década de 1890, é possível perceber como a nova organização do trabalho deu flexibilidade aos trabalhadores, que podiam pelo menos migrar para outras regiões do estado ou deixar o país em direção à Argentina ou à sua terra natal, diferentemente do que acontecia com os antigos contratos de parceria de meados do século XIX.⁴

Embora as condições de trabalho e de vida da maioria dos imigrantes fossem certamente difíceis, é provável que a conjuntura da economia tenha favorecido pelo menos uma parcela das famílias imigrantes, permitindo que elas se tornassem pequenas proprietárias de terra. Comparando a Estatística Agrícola de 1905 com o censo de 1920, Thomas Holloway mostrou que a participação de estrangeiros entre proprietários de terra aumentou substancialmente, saindo de 12,8%, em 1905, para 25,2%, em 1920.⁵ Cabe destacar que esses dados possivelmente subestimam a ascensão econômica dos imigrantes, pois incluem descendentes diretos de imigrantes como brasileiros. Mais recentemente, Andre Lanza cruzou os dados da Estatística Agrícola de 1905 com os Registros de Matrícula de Imigrantes da Hospedaria de Imigrantes de São Paulo com o objetivo de traçar a mobilidade econômica dos colonos por meio da aquisição de propriedades rurais. O autor mostrou que, apesar dos percalços enfrentados e de variar conforme a origem do imigrante e o período de chegada em São Paulo, o acesso à terra pelos imigrantes vindos para o trabalho nas fazendas de café foi, de fato, uma realidade.⁶

2. Holloway, *Imigrantes*, 114-117. Ver também Da Costa, *Da senzala*.

3. Souza, "From Bonded Laborers to Educated Citizens? Immigration, Labor Markets, and Human Capital in São Paulo, Brazil (1820-2010)", 113-117.

4. Sobre as migrações e re-imigrações para a Argentina, ver André Luiz Lanza, "Imigrantes no Brasil e na Argentina: políticas de atração, fluxos, atividades e deslocamentos São Paulo e Buenos Aires, 1870-1930" (diss. de mestrado, Universidade de São Paulo, 2015) Sobre os contratos de trabalhos firmados pelos imigrantes, ver Pierre Denis, *O Brasil no século XX* (1909), 196-9, 202-3; Petrone, "Imigração", 118-119; Verena Stolcke, *Coffee Planters Workers And Wives: Class Conflict And Gender Relations On Sao Paulo Coffee Plantations* (Springer, 1988), 28-43

5. Holloway, *Imigrantes*, 230.

6. André Luiz Lanza, "De braços para a lavoura a proprietários rurais: Imigrantes e o acesso à

Por sua vez, conforme analisado nos Capítulos 2 e 3, a população negra também estava presente entre os colonos nas fazendas e nas escolas primárias, o que sugere que parte dos ex-escravos e de seus descendentes também ocupou espaços na economia cafeeira no Oeste paulista, possivelmente também beneficiando-se das melhores condições contratuais que favoreceram os imigrantes. Este capítulo aborda a questão da propriedade rural de ex-escravos e descendentes no contexto da expansão cafeeira no pós-abolição, a partir do exemplo do município de São Carlos no início do século XX. Na primeira seção, realiza-se um esforço para identificar proprietários de terra negros a partir do cruzamento dos dados do censo municipal de São Carlos de 1907 e da Estatística Agrícola de 1905. Nessa mesma seção, apresenta-se uma análise descritiva, comparando aspectos qualitativos e quantitativos do acesso à terra entre brasileiros brancos, negros e imigrantes. Na seção seguinte, utiliza-se um modelo de classificação de *machine learning* que usa a correlação entre cor da pele e nomes para estimar a presença de proprietários negros nos municípios vizinhos a São Carlos. A última seção conclui o capítulo, resumindo os principais achados e articulando suas implicações.

4.1 Proprietários negros em São Carlos

Em uma sociedade agrária, a propriedade da terra pode ser entendida como um indicador da ascensão econômica de um indivíduo ou grupo demográfico. Para a população de ex-escravos e seus descendentes, o sucesso da acumulação patrimonial ao longo do ciclo de vida é uma medida do quanto os obstáculos legais e práticos impostos pelas instituições do regime escravocrata limitaram as possibilidades econômicas da população negra. Dessa forma, analisar a relação entre cor da pele e acesso à terra no pós-abolição é um passo importante para entender a experiência econômica não apenas da população negra livre, mas também da população escrava que a precedeu.

Esta seção apresenta dados detalhados do acesso à terra em São Carlos no início do século XX. Os resultados são apresentados de modo desagregado para terra em São Paulo, 1886-1920” (tese de doutorado, Universidade de São Paulo, 2021), Capítulo 4.

cada um dos grupos demográficos analisados nesta tese —isto é, brasileiros brancos, imigrantes e negros. A separação em grupos foi possível mediante o cruzamento dos dados do censo municipal de 1907 e da Estatística Agrícola de 1905. Para esse cruzamento, foi utilizado um algoritmo de *matching* similar ao utilizado no Capítulo 2 para a identificação de alunos de escolas primárias, e cuja descrição detalhada se encontra no Apêndice A.

Segundo os dados de São Carlos na Estatística Agrícola de 1905, havia um total de 309 propriedades em nome de indivíduos, brasileiros (247) e estrangeiros (62). Com a implementação do algoritmo, foram identificados no censo municipal de 1907 os proprietários de 192 dessas propriedades, correspondentes, assim, a 62% do total. Como alguns proprietários possuíam mais de uma propriedade, o número final dos donos de terras é menor do que o total registrado. Originalmente, a Estatística Agrícola registrou 286 proprietários únicos, dos quais 168 (59%) tiveram a cor da pele identificada. Assim, desses 168 proprietários, 104 (61,9%) eram brasileiros brancos; 58 (34,5%) imigrantes; e 6 (3,6%) eram brasileiros negros. A Tabela 4.1, resultante do cruzamento dos dados da Estatística Agrícola com o censo de São Carlos, apresenta informações demográficas para cada um dos grupos de proprietários.

Antes de apresentar os resultados da análise dos grupos demográficos, convém notar que, diferentemente do que ocorre com brasileiros brancos e negros, os dados da Tabela 4.1 referem-se a todos os imigrantes listados na Estatística Agrícola de 1905. Essa plena correspondência no caso dos imigrantes acontece porque todos membros desse grupo foram considerados brancos, enquanto consideram-se brasileiros brancos e negros apenas aqueles identificados por meio do *matching*, que combinou informações do censo municipal com a Estatística Agrícola. Em outras palavras, os números de proprietários brasileiros, brancos e negros, representam o mínimo de cada grupo na Estatística Agrícola, enquanto o número de imigrantes proprietários representa o número exato de proprietários estrangeiros. Todos os 118 proprietários listados em São Carlos na Estatística Agrícola, cuja cor não foi possível determinar com o algoritmo e que, portanto, ficaram de fora da base analisada nesta seção, são brasileiros.

A participação dos grupos demográficos entre os proprietários de terra, apre-

Tabela 4.1: Dados demográficos de proprietários de terra, São Carlos, 1905/1907

	Brasileiros brancos	Imigrantes	Negros
Cor			
Branco	104	58	-
Mulato	-	-	3
Preto	-	-	3
Sexo			
Feminino	14	4	-
Masculino	90	54	6
Nacionalidade			
Brasileiro	104	-	6
Italiano	-	39	-
Português	-	13	-
Espanhol	-	4	-
Russo	-	1	-
outra	-	1	-
Profissão			
Administrador	1	1	-
Empregado Público	1	-	-
Fazendeiro	3	-	-
Lavrador	96	34	5
Mecânico	1	-	-
Negociante	1	-	1
Servente Justiça	1	-	-
Alfabetização			
Alfabetizado	88	21	2
Analfabeto	16	14	4

Fontes: São Carlos, Recenseamento de 1907.

São Paulo, Secretaria de Agricultura, Commercio e Obras Publicas do Estado de São Paulo. (1906-1907). Estatística agrícola e zootechnica do estado de São Paulo no anno agrícola de 1904-1905. São Paulo: Tip. Brazil.

Notas: Quando a informação demográfica estava disponível no Censo Municipal e na Estatística Agrícola optamos por usar a informação do Censo. Informações de profissão e leitura não estão disponíveis para uma parcela dos imigrantes (ver Apêndice A para detalhes.)

sentada na Tabela 4.1, sugere que brasileiros brancos, que representavam cerca de 62% do total, estavam desproporcionalmente representados entre proprietários quando se considera seu peso relativo na população, enquanto negros e imigran-

tes encontravam-se sub-representados. É possível precisar melhor qual o grau de sobre e sub-representação de cada um dos grupos demográficos analisados. Assim como no Capítulo 3, utiliza-se aqui o índice de equilíbrio racial para comparar as taxas de acesso à terra entre os três grupos demográficos estudados. Conforme visto antes, o índice de equilíbrio racial mede o grau de sobre e sub-representação de um grupo demográfico em uma ocupação, dada uma população de referência.

A versão padronizada do índice de equilíbrio racial varia de -1 a 1, com valores negativos indicando sub-representação e valores positivos indicando sobre-representação do grupo demográfico sob consideração. Assim, o índice indica que brasileiros brancos eram bastante sobre-representados entre proprietários de terra, pois o índice de equilíbrio racial para esse grupo, que é de 0,9, fica próximo ao valor máximo do índice. Brasileiros brancos representavam 19% da população masculina com mais de 20 anos, em São Carlos em 1907. Se esse percentual fosse reproduzido entre proprietários, seriam observados 33 proprietários de terra pertencentes ao grupo. O número observado, contudo, é de 104, o que é ainda mais expressivo quando se considera que essa figura, como notado anteriormente, representa o limite inferior do total de brasileiros brancos proprietários de terra em São Carlos, em 1905.

O índice de equilíbrio racial mostra ainda que o número de imigrantes e negros proprietários é menor do que o esperado, revelando desigualdade no acesso à terra para ambos os grupos. Imigrantes alcançam um índice de -0,62, enquanto negros chegam a -0,69. Em outras palavras, embora imigrantes também estivessem sub-representados entre proprietários de terra, sua sub-representação era relativamente menor do que a de negros. Com efeito, foram registrados 58 imigrantes e 6 negros proprietários de terra, quando o esperado (levando em conta suas participações na população adulta) seria observar 112 e 22, respectivamente. Ainda que os resultados mostrem uma pequena vantagem de imigrantes em relação aos negros no acesso à terra, essa evidência não deve ser exagerada. Em primeiro lugar, a diferença do índice de equilíbrio racial de imigrantes e negros é muito pequena e, assim, pode ser suscetível ao critério demográfico adotado para o grupo de referência (que, no caso, é a população masculina maior de 20 anos). Em segundo lugar, se existiam proprietários negros entre aqueles que não tiveram

a sua cor da pele identificada, então o índice de equilíbrio de negros seria necessariamente maior, caso em que o índice de negros se aproximaria ou superaria o valor observado para imigrantes. Por fim, como visto no Capítulo 1, os dados do censo de São Carlos indicam que a prevalência de proprietários era muito similar entre negros e imigrantes. Na população masculina com idade igual ou superior a 20 anos e autodeclarados proprietários, 13,5% eram negros e 14,2% imigrantes (ver Tabela 1.4).

A questão fundiária não se restringe ao acesso. A Estatística Agrícola de 1905 detalha informações sobre área total, uso e tipo da terra, lavouras cultivadas e valor da terra. Essa dimensão qualitativa da propriedade da terra deve também ser levada em conta na comparação entre brasileiros brancos, imigrantes e negros. A Tabela 4.2 apresenta as principais informações sobre as terras em propriedade de negros, considerando aqueles identificados simultaneamente no censo municipal de 1907 e na Estatística Agrícola de 1905.

Tabela 4.2: Dados de propriedades de negros, São Carlos, 1905/1907

Nome proprietário	Cor	Área (alq.)	Café (área)	Nacionais	Estrangeiros	Valor (mil-réis)
Arges Vinhas	Mulato	19	40,9	0	8	15.000
Arges Vinhas	Mulato	50	50,0	0	24	30.000
Bernardo Caetano	Preto	1	0,0	1	0	500
Daniel Antonio Gardino	Mulato	171	89,5	3	16	200.000
Francisco Antonio Borges	Mulato	90	50,0	5	25	150.000
Francisco Antonio Borges	Mulato	185	50,0	15	18	250.000
Jose Romao Dos Reis	Preto	236	0,0	4	0	20.000
Manoel Camargo	Preto	150	50,0	1	17	129.000

Fontes: São Carlos, Recenseamento de 1907.

São Paulo, Estatística agrícola e zootechnica do estado de São Paulo no anno agrícola de 1904-1905.

Notas: A coluna "Área" representa a área total em alqueires (1 alqueire paulista = 2,42 hectares ou 24.200 metros quadrados). A coluna "Café" representam o percentual da área total dedicada ao cultivo de café. As colunas "Nacionais" e "Estrangeiros" indicam o número de trabalhadores brasileiros e estrangeiros que trabalhavam na propriedade à época do levantamento.

A análise dos dados do tamanho das propriedades permite concluir que não existia nenhum latifundiário entre os proprietários negros. Usando a classificação

de Sérgio Milliet, as 8 propriedades de negros se dividem em pequenas (2), médias (2) e grandes (4).⁷ Por outro lado, o maior número de propriedades médias e grandes indica que as posses de negros eram expressivas. Analisando os mesmos dados, Karl Monsma chamou atenção para a diferença entre pretos e mulatos.⁸ No cálculo do autor foram identificados 3 proprietários pretos, nenhum deles empregando imigrantes ou plantando café, somando em conjunto, aproximadamente, 82 alqueires. Em contraste, o autor cita dois mulatos, Francisco Antônio Borges e Argeu Vinhas, que não só tinham terras maiores, como também empregavam estrangeiros nelas e cultivavam café.

Apesar dos dados apresentados nesta seção corroborarem os achados de Monsma, há duas diferenças importantes. Em primeiro lugar, embora a abordagem de cruzamento dos dados de Monsma e a adotada aqui tenham encontrado 3 proprietários pretos, há um resultado distinto e relevante em um dos indivíduos selecionados. O algoritmo de *matching* utilizado para identificar os proprietários negros não encontrou uma correspondência para Elesbão Galo, proprietário listado na Estatística Agrícola, no censo municipal de 1907. Karl Monsma parece ter incluído o indivíduo "Lisbon Joaquim Gallo", porém esse nome apresentou uma distância de *string* muito elevada para ser considerado um *match* pelo algoritmo utilizado nesta seção. Uma segunda diferença refere-se aos dados do proprietário Argeu Vinhas, mulato, que tinha, de acordo com a transcrição aqui realizada, não uma, mas duas propriedades listadas na Estatística Agrícola. Somadas, essas propriedades alcançavam uma área total de 69 alqueires e empregavam 32 trabalhadores estrangeiros.

Mesmo levando em conta esses resultados distintos, as diferenças entre proprietários pretos e mulatos permanecem, com vantagem para os últimos. Em média, mulatos empregavam um número maior de trabalhadores, estrangeiros e nacionais, em suas terras. Além disso, o valor médio da propriedade de mu-

7. A classificação de Sérgio Milliet definiu pequenas propriedades como aquelas com até 25 alqueires, médias propriedades tendo entre 25 e 100 alqueires, e grandes propriedades possuindo entre 100 e 500 alqueires. Acima do último limite, considera-se latifúndio (Ver Milliet, *Roteiro*, 76 e Caio Prado Junior, "Distribuição da propriedade fundiária rural no Estado de São Paulo", *Revista Geografia* 1, n. 1 (1935)).

8. Monsma, *A reprodução*, 306-307.

latos foi de 129 mil-réis, contra apenas 49,8 mil-réis entre pretos. O único aspecto em que proprietários pretos superaram mulatos foi em área total. O maior proprietário entre os donos de terra negros foi José Romão dos Reis, registrado como preto no censo e dono de uma propriedade de 236 alqueires. Assim, pretos possuíam 129 alqueires de terra em média, enquanto mulatos detinham propriedades de cerca de 103 alqueires em média.

Como foi notado antes, em conjunto as propriedades em posse de negros e imigrantes comparavam-se desfavoravelmente com as de brasileiros brancos. Como mostra a Tabela 4.3, que apresenta estatísticas descritivas da área total das propriedades rurais para os três grupos demográficos sob análise, a área total média controlada por brasileiros brancos era de 157,6 alqueires, contra 112,8 alqueires de negros e 44,5 de imigrantes. Havia maior desigualdade na distribuição da área total entre brasileiros brancos ($Gini=0,64$) e imigrantes ($Gini=0,68$) causada, em parte, pela presença de latifundiários nesses dois grupos. Se por um lado não foi encontrado nenhum latifundiário entre negros ($Gini=0,40$), havia 9 proprietários entre brasileiros brancos com terras que superavam os 500 alqueires. Entre imigrantes, havia apenas 1 latifúndio, mas, a ampla predominância de pequenas propriedades nesse grupo fazia com que a desigualdade fosse elevada e o tamanho da propriedade média baixo.

A importância do tamanho da propriedade, contudo, não deve ser exagerada. Outros indicadores presentes na Estatística Agrícola que também se relacionam com o potencial de produção agrícola devem ser analisados. A Tabela 4.4 compara os valores médios do percentual da área dedicada à plantação de café, número de trabalhadores e valor das propriedades de brasileiros brancos, imigrantes e negros. É possível notar, novamente, vantagem para os brasileiros brancos nas três medidas. Assim, em média, brasileiros brancos tinha condições de dedicar cerca de 10 pontos percentuais a mais da área total para a produção de café que imigrantes e negros. Além disso, enquanto proprietários negros e imigrantes empregavam, respectivamente, cerca de 17 e 7 trabalhadores em média, brasileiros brancos empregavam 33.

O tipo de terra considerado de maior valor à época era a terra roxa. Cerca de 30% das propriedades rurais de brasileiros brancos em São Carlos possuíam esse

Tabela 4.3: Estatísticas Descritivas da área total de propriedades rurais de Brasileiros brancos, Imigrantes e Negros, São Carlos, 1905/1907

	Brasileiros Brancos	Imigrantes	Negros
Número de Propriedades	123	61	8
Média	157,6	44,5	112,8
Desvio-padrão	281,3	128,4	85,3
Mínimo	3	0,5	1
25%	30,8	6,5	42,2
50%	68,5	20	120
75%	149,5	37,5	174,5
Máximo	1808	1000	236
Gini	0.64	0.68	0.40
Latifundiários	9	1	0

Fontes: São Carlos, Recenseamento de 1907.

São Paulo, Estatística agrícola e zootécnica do estado de São Paulo no ano agrícola de 1904–1905.

Notas: Área total em alqueires (1 alqueire paulista = 2,42 hectares ou 24.200 metros quadrados).

Tabela 4.4: Indicadores relativos à propriedade rural de Brasileiros brancos, Imigrantes e Negros, São Carlos, 1905/1907

Grupo	Café (área)	Nacionais	Estrangeiros	Terra roxa	Valor (mil-réis)
Brasileiros brancos	39,8	7,3	26,4	29,7	108.199,2
Imigrantes	27,9	1,7	5,9	8,5	19.442,6
Negros	28,0	3,6	13,5	12,5	99.312,5

Fontes: São Carlos, Recenseamento de 1907.

São Paulo, Estatística agrícola e zootécnica do estado de São Paulo no ano agrícola de 1904–1905.

Notas: A coluna “Café” representam a média do percentual da área total dedicada ao cultivo de café. As colunas “Nacionais” e “Estrangeiros” indicam o número médio de trabalhadores brasileiros e estrangeiros que trabalhavam na propriedade. A coluna “Terra roxa” indica o percentual de propriedade rurais do grupo que tinha essa qualidade de terra.

tipo de terra, enquanto entre imigrantes esse percentual era de apenas 8,5%. Entre brasileiros negros, o percentual de propriedades com terra roxa era de 12,5% devido exclusivamente a uma propriedade, ou seja, apenas 1 das 8 propriedades tinha terra roxa. O tipo de terra mais comum entre proprietários imigrantes e

negros era a terra branca, considerada menos fértil. Das 6 propriedades em posse de negros, 4 tinham terra branca (67%). Entre imigrantes, a existência de terra branca era proporcionalmente menor: 35 das 61 (57%) propriedades desse grupo apresentavam esse tipo de terra. Apesar das similaridades entre as propriedades de imigrantes e negros, a pequena vantagem observada para os últimos em relação ao tamanho médio das propriedades, ajuda a explicar a diferença do valor médio da propriedade entre imigrantes e negros. Assim, o valor médio da propriedade de negros era de aproximadamente 100 mil-réis, enquanto o valor médio da propriedade sob posse de imigrantes ficava em torno de 20 mil-réis.

Pode-se argumentar, contudo, que a diferença em favor de proprietários negros em relação a proprietários imigrantes que aparece em medidas de tendência central é um efeito estatístico provocado pela combinação de existência de *outliers* e reduzido número de informações no grupo de proprietários negros. As tabelas C.10 e C.11, no apêndice C, reproduzem as tabelas 4.3 e 4.4 sem a presença do proprietário Francisco Antônio Borges que, em termos de valor e quantidade de terras, é o principal proprietário rural no grupo de negros. Os resultados mostram que, mesmo após a remoção do principal proprietário negro, as diferenças nas médias do tamanho e valor das propriedades rurais se mantêm em favor de proprietários negros. Dessa forma, as vantagens observadas nos indicadores das propriedades rurais de negros em relação a imigrantes não parecem decorrer da escolha de medidas de tendência central para sumarizar os dados fundiários sob análise.

É relevante neste momento comparar as condições de acesso à terra entre negros e imigrantes observadas nos dados de São Carlos com as da experiência dos Estados Unidos após a Guerra Civil e o fim da escravidão naquele país. Assim como o Brasil, a economia e a demografia dos Estados Unidos foram muito influenciadas pela escravidão e pela imigração. Em que pese a existência de diferenças substanciais entre ambos, é possível que a formação de uma classe de proprietários em cada uma dessas subpopulações tenha enfrentado obstáculos análogos àqueles percebidos pela população negra e imigrante brasileira. De fato, uma tendência observada para São Carlos, que se repete nos Estados Unidos no pós-abolição, é a permanência da posição dominante da população branca na

posse da terra.⁹

Um outro paralelo ocorre quando se comparam os indicadores de propriedade da população negra livre nos Estados Unidos com os imigrantes de primeira geração. Como visto acima, em São Carlos, o valor médio da propriedade de imigrantes era menor do que aquele observado para negros em 1905-1907. Tendência similar também pode ser observada no caso dos Estados Unidos. Na Lousiana, ainda antes da Guerra Civil (1850), um indivíduo negro livre possuía, em média, US\$ 1.479 em propriedade, o que era quase o dobro do valor médio da propriedade de cidadãos nascidos no estrangeiro.¹⁰ Não está claro o quanto essa tendência se manteve ao longo do tempo e entre os diferentes estados, mas o crescimento do patrimônio da população negra no pós-emancipação nos Estados Unidos torna improvável que, na ausência de importação de capital, imigrantes de primeira geração apresentassem indicadores de propriedade superiores aos de afro-americanos.¹¹

De uma maneira geral, portanto, a análise dos dados apresentados nesta seção permite chegar a algumas conclusões sobre a questão fundiária relacionada aos diferentes grupos populacionais em São Carlos no início do século XX. Parece clara a posição privilegiada em que brasileiros brancos se encontravam em relação a imigrantes e brasileiros negros no que diz respeito ao acesso à terra no município. Brasileiros brancos não só tinham posse das maiores propriedades rurais, como também eram os principais produtores de café, contando com as terras de maior qualidade e com condições de dedicar a maior parte delas ao cultivo da lavoura mais lucrativa na época.

Também foram analisadas diferenças entre pretos e mulatos. Os dados apresentados nos capítulos anteriores mostraram a preponderância de mulatos sobre pretos quando foram consideradas as dimensões que compõem o capital humano. No Capítulo 1, viu-se que, para a população maior de 15 anos, a taxa de alfabetização de mulatos era aproximadamente 11 pontos percentuais maior

9. Ransom e Sutch, *One kind*, 78-80.

10. Loren Schweninger, *Black property owners in the South, 1790-1915* (University of Illinois Press, 1990), 4.

11. Entre 1890 e 1910, o patrimônio da população negra americana cresceu a uma taxa maior do que aquela observada para a população branca. Ver *ibid.*, 183-184

do que a de pretos (21,1% e 10,4%, respectivamente). No Capítulo 3, observou-se que pretos predominavam entre trabalhadores domésticos, enquanto mulatos apareciam com maior frequência entre trabalhadores mais qualificados, como carpinteiros e pintores. Uma possibilidade, então, é que a ascendência mista de mulatos tenha amenizado o passivo do passado escravocrata, fazendo com que essa população tivesse indicadores de capital humano melhores do que os da população preta, o que explicaria também sua melhor posição no mercado de trabalho. Os achados desta seção reforçam o entendimento acerca da vantagem socioeconômica de mulatos em relação a pretos. Foi visto que, com exceção do tamanho da propriedade, todos os indicadores quantitativos e qualitativos do acesso à terra eram melhores para proprietários mulatos do que pretos. Os dados de propriedade rural indicam, assim, que a acumulação de patrimônio pessoal parece ter sido uma possibilidade mais viável para mulatos do que para pretos.

Um outro fato que pôde ser verificado a partir dos dados analisados foi a proximidade das condições materiais de imigrantes e negros. A proporção relativa desses dois grupos entre proprietários de terra é similar. Ambos os grupos aparecem entre proprietários em proporção inferior à sua respectiva participação na população. Além disso, indicadores qualitativos, como qualidade da terra e percentual da área total dedicada ao cultivo de café, sugerem que proprietários imigrantes e negros eram parecidos entre si e muito diferentes de proprietários brasileiros brancos. Diferenças importantes, contudo, foram observadas na área média da propriedade e no valor das propriedades dos dois grupos. Em ambos os casos, as diferenças eram a favor de proprietários negros.

4.2 Para além de São Carlos: proprietários negros em outros municípios paulistas

A Estatística Agrícola de 1905 contém informações sobre propriedades rurais para todo o estado de São Paulo. Apesar da riqueza de dados, no que diz respeito aos proprietários rurais o documento limita-se a registrar os seus nomes e nacionalidades. Em particular, não é possível saber diretamente qual a cor da pele dos

proprietários. Na seção anterior, procurou-se superar esse obstáculo com o cruzamento dos dados da Estatística Agrícola com o Censo Municipal de São Carlos de 1907. Nesta seção, realiza-se uma abordagem complementar com o fim de estimar a cor dos indivíduos listados na estatística rural.

A abordagem proposta consiste no uso de um algoritmo de *machine learning* que estima a cor de indivíduos a partir de seus nomes. Dessa forma, em contraste com a seção anterior, em que a análise das condições de acesso à terra de negros e brancos em São Carlos teve por base uma variável *observada* de cor da pele, nesta seção a propriedade da terra de brancos e negros é avaliada a partir de uma *estimativa* da informação da cor da pele feita a partir do nome. Entre as vantagens do uso do algoritmo de *machine learning* inclui-se a ampliação do escopo regional do presente estudo.

A ampliação espacial do estudo, contudo, é feita de maneira controlada. A estimativa da cor da pele a partir do nome de um indivíduo é possível apenas em função de fatores históricos que levaram à distinção e classificação de nomes de negros e brancos. Sendo um fenômeno histórico, a correlação entre cor da pele e nome está limitada temporal e espacialmente. Por essa razão, embora informações de propriedades rurais estejam disponíveis para todo o estado de São Paulo em 1905, as estimativas desta seção foram feitas levando em conta um limite geográfico que, no caso, consiste em fazer uso apenas dos dados referentes aos municípios vizinhos a São Carlos (ver Figura 4.1).

Pressupõe-se assim que os municípios limítrofes de São Carlos possuíam realidades históricas e sociais similares entre si, com conexões entre suas populações e grupos demográficos como de brasileiros e imigrantes. A evolução demográfica no pós-abolição e a estrutura fundiária parecem corroborar esse pressuposto. Foi visto no Capítulo 1 que São Carlos assistiu a um elevado crescimento demográfico entre 1890 e 1900, com o aumento dos movimentos migratórios de nacionais e estrangeiros. A Tabela 4.5, que compara o tamanho das populações dos municípios vizinhos de São Carlos nos censos demográficos nacionais de 1890 e 1900, indica que um processo similar teria ocorrido naqueles municípios.

Como os dados da Tabela 4.5 usam diretamente as informações dos censos, existem duas potenciais limitações. Em primeiro lugar, os valores do total da

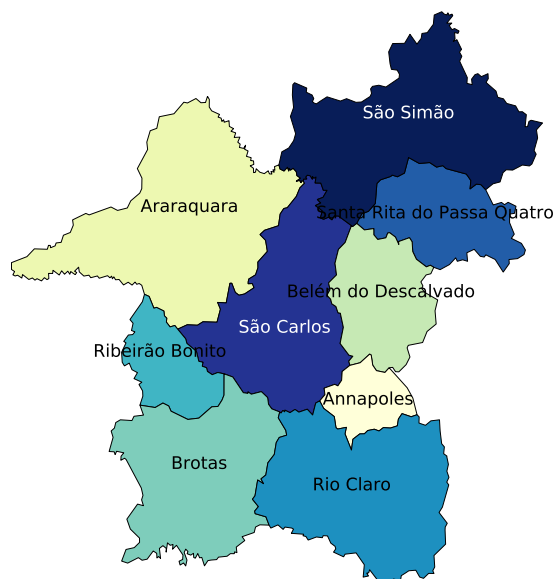


Figura 4.1: Municípios limítrofes de São Carlos, 1900

Fonte: Malha municipal 1900. IBGE.

população em cada município não controlam para mudanças territoriais ocorridas entre os dois censos. Em que pese o fato dos principais municípios da região, como São Carlos, Araraquara e São Simão não terem passado por desmembramento ou incorporação de distritos no período. Em segundo lugar, a baixa qualidade do censo de 1900 pode levar a uma estimativa equivocada do crescimento demográfico no período. Em particular, o crescimento demográfico de Araraquara e São Carlos parecem excessivamente elevados. No caso de São Carlos, o censo municipal de 1907 lista uma população de 38.641 pessoas, valor consistente com taxas de crescimento demográfico mais realistas que aquelas implicadas pelo total de 55.729 pessoas indicado pelo censo de 1900.

Embora a taxa de crescimento observada em São Carlos no período só tenha paralelo com Araraquara e São Simão, todos os municípios da região passaram por crescimento demográfico entre 1890 e 1900, sendo possível que, assim como São Carlos, a causa primária desse crescimento tenha sido a expansão cafeeira na região. Além disso, ainda que com intensidade variável em cada município,

as tendências demográficas observadas para São Carlos, destacadas no Capítulo 1, podem ter se repetido nos municípios vizinhos, tornando o perfil demográfico dos municípios vizinhos similar ao observado em São Carlos.

Tabela 4.5: Evolução demográfica, Região de São Carlos, 1890-1900

Município	1890	1900	Taxa de crescimento (a.a)
Anápolis (Analândia)	-	6.535	-
Araraquara	8.151	28.900	13.5
Belém do Descalvado	13.903	23.838	5.5
Brotas	9.507	10.029	0.5
Ribeirão Bonito	4.806	5.992	2.2
Rio Claro	24.584	31.891	2.6
Santa Rita do Passa Quatro	8.818	10.994	2.2
São Carlos	12.651	55.729	16.0
São Simão	3.477	24.850	21.7

Fonte: Censos nacionais de 1890 e 1900.

Além de possivelmente replicarem as tendências demográficas observadas em São Carlos, os municípios vizinhos possuíam estrutura fundiária similar, como pode ser visto na Tabela 4.6. Na maior parte dos municípios da região, observa-se a tendência de poucas propriedades controlarem a maior parte da área total, o que faz com que o índice de Gini da área das propriedades rurais alcance níveis elevados, com mínimo de 0,57 e máximo de 0,73. Além disso, assim como em São Carlos, os imigrantes tinham uma presença substancial entre os proprietários de terra nos municípios vizinhos, ainda que predominantemente entre pequenos proprietários. Em toda região, os estrangeiros totalizavam aproximadamente 40% dos pequenos proprietários e 15% dos latifundiários.

As semelhanças entre São Carlos e municípios vizinhos é importante para o exercício realizado a seguir. Algoritmos de *machine learning* utilizam informações conhecidas (observadas) de um grupo de treino para prever resultados de um grupo de teste. É fundamental que as amostras de treino e teste sejam parecidas, para que um algoritmo desenvolvido para uma amostra não produza resultados substancialmente diferentes para a outra amostra. Igualmente importante para a eficácia do algoritmo é que seja possível separar os grupos que se pretende

Tabela 4.6: Estatísticas fundiárias, Região de São Carlos, 1905

Município	Pequena	Média	Grande	Latifúndio	Gini	Estrangeiro
Anápolis	15,2	24,2	42,4	18,2	0,57	18,2
Araraquara	48,9	25,7	20,1	5,3	0,77	35,6
B. Descalvado	55,3	19,3	19,7	5,7	0,76	35,5
Brotas	53,3	33,2	10,7	2,8	0,69	33,6
Ribeirão Bonito	31,5	46,7	18,5	3,3	0,69	16,3
Rio Claro	40,2	36,6	19,8	3,3	0,67	37,6
Sta Rita P.Q	46,0	30,7	20,3	3,0	0,72	23,8
São Carlos	33,9	35,1	24,6	6,4	0,73	19,8
São Simão	51,5	25,1	21,0	2,4	0,73	13,1

Fonte: São Paulo, Estatística agrícola e zootécnica do estado de São Paulo no ano agrícola de 1904–1905.

Notas: Exceto Gini, todas as colunas estão em percentual. Para classificação, ver texto.

classificar com base nas informações disponíveis. Na próxima subseção, será visto que, por razões históricas, o nome de indivíduos é uma informação que permite distinguir brancos e negros no Brasil no início do século XX.

4.2.1 Nomes e cor da pele

A história de africanos e seus descendentes no Novo Mundo é marcada por padrões onomásticos decorrentes da própria experiência do cativo e de seu impacto sobre as atitudes das gerações de libertos que se seguiram ao fim do regime escravista. Nos Estados Unidos, por exemplo, vários estudos mostram a presença de correlação entre cor da pele e nomes e sobrenomes para diferentes épocas da história americana no *Postbellum*.¹² Analisando nomes de homens chefes de família, Lisa Cook e colegas identificaram pelo menos 21 nomes distintamente negros em todos os Estados Unidos, nas primeiras décadas do século XX.¹³ Embora não seja imediatamente evidente a causa da concentração de certos

12. Lisa D Cook, Trevon D Logan e John M Parman, “Distinctively black names in the American past”, *Explorations in Economic History* 53 (2014): 64–82; Roland G Fryer Jr e Steven D Levitt, “The causes and consequences of distinctively black names”, *The Quarterly Journal of Economics* 119, n. 3 (2004): 767–805.

13. Cook, Logan e Parman, “Distinctively black names in the American past”.

nomes entre afro-americanos, os autores sugerem que a adoção de nomes religiosos e honoríficos indica tanto a importância da religião na vida de ex-escravos e descendentes como a tendência de buscar ressignificação social a partir de nomes que refletiam poder e autonomia.¹⁴

No Brasil, também há evidência de que nomes devocionais são mais comuns na população negra. Um estudo, por exemplo, mostra que existe uma tendência de agrupamento de diferentes métricas do cromossomo Y que permitem caracterizar brasileiros brancos, negros sem nomes devocionais e negros com nomes devocionais como grupos independentes.¹⁵ Segundo o mesmo estudo, outros trabalhos já haviam mostrado que quanto maior a predominância de sobrenomes devocionais em um dado grupo de pessoas, mais escura era a pele de indivíduos deste grupo. A relação entre ascendência africana e sobrenomes religiosos também é atestada por estudos históricos.¹⁶

A evidência histórica de nomes tipicamente negros no Brasil, contudo, apresenta algumas diferenças em relação ao Sul escravista dos Estados Unidos. A adoção de nomes típicos por parte de ex-escravos e descendentes é, em parte, um possível resultado da busca de autonomia e ressignificação social no pós-abolição. No entanto, também é possível que esses mesmos incentivos operassem no sentido contrário, levando ex-escravos e descendentes a utilizar nomes mais comuns entre brancos. Um estudo feito para São Carlos no início do século XX indica que, entre alguns ex-escravos, foi comum o expediente de adotar sobrenomes de famílias tradicionais.¹⁷ A adoção de sobrenomes de famílias tradicionais teria o efeito de diminuir a correlação entre nome e cor da pele, dificultando análises desse tipo.

Mesmo assim, ao menos em São Carlos, a adoção de sobrenomes de famílias tradicionais por parte da população negra deu-se de forma apenas parcial, com os ex-escravos adotando apenas um dos sobrenomes. Como já salientado nesta

14. Nomes honoríficos encontrados pelos autores incluíam Master, King e Prince.

15. Ana AL Barbosa et al., "Family names and the length of the Y chromosome in Brazilian blacks", *Brazilian Journal of Genetics* 20 (1997).

16. José Tavares-Neto e Eliane S Azevêdo, "Racial origin and historical aspects of family names in Bahia, Brazil", *Human Biology*, 1977, 287–299.

17. Rogerio da Palma e Oswaldo Mário Serra Truzzi, "Renomear para Recomeçar: Lógicas Onomásticas no Pós-abolição", *Dados* 61, n. 2 (2018): 318.

tese, todos os indivíduos registrados no censo municipal de 1907 com sobrenomes das duas principais famílias de São Carlos, Arruda Botelho e Camargo Penteados, eram brancos. Assim, é possível que nomes como “Arruda” e “Camargo” tenham se tornado mais negros no pós-abolição, mas os nomes compostos permaneceram distintamente brancos.

De qualquer forma, com os dados do censo de São Carlos, é possível testar empiricamente se havia nomes distintamente negros, se sobrenomes devocionais eram mais comuns entre negros, e se sobrenomes parciais de famílias tradicionais do município eram mais negros do que os sobrenomes compostos. Para esse exercício, foi utilizado um índice desenvolvido por Lisa Cook em uma série de trabalhos sobre padrões onomásticos na população negra americana.¹⁸ Esse índice então foi implementado com os dados do censo municipal de São Carlos de 1907. O cálculo do índice é feito em dois estágios. No primeiro, são analisados os nomes mais comuns entre pessoas negras, mantendo-se apenas aqueles cuja frequência é maior do que a frequência mediana da distribuição de todos os nomes de negros. No segundo estágio, calcula-se a probabilidade de um indivíduo ser negro, dado que possua um dos nomes encontrados no primeiro estágio. A implementação do índice usando esses dois estágios é feita para os primeiros e os últimos nomes. Os resultados são apresentados na Tabela 4.7, que exibe os 10 primeiros e os 10 últimos nomes mais comuns entre negros de São Carlos em 1907.

Observando a Tabela 4.7, fica clara a existência de nomes tipicamente negros. Mais da metade das pessoas com nomes como Benedicto(a) (51%) e sobrenomes como Conceição (53%) são negros, mesmo com os negros constituindo apenas 12% da população total de São Carlos em 1907. Um grau menor, mas ainda elevado de concentração de negros, também aparece nas versões masculinas e femininas do nome Sebastião (36% e 34%, respectivamente) e de sobrenomes como Vicente (34,4%) e Santos (32,7%). Outros nomes devocionais que não aparecem na tabela, como Jesus e Cruz, também são desproporcionalmente comuns entre negros. Entre pessoas com nomes Jesus e Cruz, pessoas negras somavam 37% e

18. Cook, Logan e Parman, “Distinctively black names in the American past”; Lisa D Cook, “Inventing social capital: Evidence from African American inventors, 1843–1930”, *Explorations in Economic History* 48, n. 4 (2011): 507–518; Lisa D Cook, *African american inventors data set*, 2004.

Tabela 4.7: Nomes e sobrenomes mais comuns de negros, percentuais e números absolutos, São Carlos, 1907

	Brasileiros Brancos	Imigrantes	Negros	Total Negros
Primeiro nome	(%)	(%)	(%)	
Maria	50,1	36,6	13,3	485
Benedicto	44,0	4,9	51,1	166
Benedicta	44,5	3,9	51,6	131
Sebastiao	56,9	6,4	36,6	108
Manoel	36,5	46,0	17,4	105
Sebastiana	63,6	2,0	34,3	68
Joaquim	65,5	18,1	16,4	65
Francisca	46,6	35,3	18,1	63
Joanna	44,4	37,8	17,8	51
Carlos	58,6	26,5	14,9	51
Último nome				
Conceição	37,9	9,4	52,8	220
Santos	44,0	23,2	32,7	148
Francisco	25,5	27,0	47,4	65
Antonio	39,8	22,0	38,1	45
Maria	42,3	22,8	35,0	43
Vicente	43,8	21,9	34,4	22
Augusto	33,8	35,4	30,8	20
Benedicto	9,5	0,0	90,5	19
Paula	56,4	19,2	24,4	19
Luiza	23,5	23,5	52,9	18
População	48,2	39,4	12,3	4757

Fonte: São Carlos, Recenseamento de 1907.

Notas: A coluna "Total Negros" representa o número absoluto de negros com o nome em questão. As demais colunas estão em percentuais.

33%, respectivamente.

Um outro dado relevante da Tabela 4.7 é a ausência de sobrenomes das famílias tradicionais. Teriam esses sobrenomes permanecido brancos mesmo com a adoção deles por parte de alguns ex-cativos? Na Tabela 4.8, apresentam-se os valores dos índices, segundo grupos demográficos, para os nomes de três famílias tradicionais de São Carlos: Arruda Botelho, Camargo Penteado e Abreu Sampaio.

Como é possível notar na tabela, o uso de nomes tradicionais entre negros

Tabela 4.8: Sobrenomes tradicionais por grupos demográficos, São Carlos, 1907

Sobrenome Tradicional	Brasileiros brancos	Imigrantes	Negros	Total Negros
Singular				
Abreu	77,3	0,0	22,7	5
Arruda	78,2	2,6	19,2	37
Botelho	75,0	13,2	11,8	9
Camargo	60,4	1,0	38,6	117
Penteado	83,0	0,0	17,0	17
Sampaio	86,6	4,2	9,2	13
Composto				
Abreu Sampaio	100,0	0,0	0,0	0
Arruda Botelho	100,0	0,0	0,0	0
Camargo Penteado	100,0	0,0	0,0	0
População	48,2	39,4	12,3	

Fonte: São Carlos, Recenseamento de 1907.

Notas: A coluna "Total Negros" representa o número absoluto de negros com o nome em questão. As demais colunas estão em percentuais.

limitava-se, de fato, a apenas um dos nomes. Os nomes "Arruda" e "Camargo", por exemplo, aparecem entre negros em proporção maior (19,2% e 38,6%, respectivamente) do que o percentual de negros (12,3%) na população. O caso do sobrenome "Camargo" é o mais extremo. A proporção de negros com esse nome é maior do que a que foi observada para alguns sobrenomes devocionais, como Santos e Jesus. Não obstante, esse efeito desaparece quando são considerados os sobrenomes compostos. Não havia nenhum negro com nomes compostos das três famílias tradicionais analisadas. Em conjunto, as famílias "Abreu Sampaio", "Arruda Botelho" e "Camargo Penteado" possuíam um total de 48 membros, todos classificados como brancos no censo municipal de 1907.

Em resumo, os dados apresentados anteriormente sugerem que nomes e sobrenomes estavam correlacionados à cor da pele no pós-abolição. Na próxima seção, parte-se desse fato para estimar um modelo de classificação que, a partir da informação do nome, seja capaz de prever a cor da pele dos proprietários rurais nos municípios ao redor de São Carlos.

4.2.2 Classificação de proprietários

É possível agora realizar um exercício de classificação para se predizer a cor da pele dos proprietários rurais dos municípios limítrofes de São Carlos que foram listados na Estatística Agrícola de 1905. A combinação de técnicas de reamostragem, necessárias devido ao desbalanceamento das classes, e algoritmos levou ao teste de 57 modelos. O modelo de classificação com o melhor desempenho foi o *Multi-layer Perceptron Classifier* (daqui em diante, MLP). O MLP é um algoritmo de classificação de redes neurais e, como tal, consiste em uma generalização não-linear de modelos lineares.¹⁹

A Tabela 4.9 sintetiza os resultados dos três melhores modelos, ao lado das técnicas de reamostragem e das métricas selecionadas de desempenho (Acurácia, Kappa e AUC).²⁰ Todos os modelos foram implementados utilizando o pacote scikit-learn em Python.²¹

Tabela 4.9: Resultados de modelos preditivos selecionados

Modelo	Re-amostragem	Acurácia	Kappa	AUC
Random Forest	SMOTE random under sampling	0.84	0.13	0.68
SVM	random undersampling	0.77	0.25	0.72
MLP	random oversampling	0.85	0.22	0.74

Notas: Métricas calculadas usando a amostra de teste. Kappa de Cohen. Predições do Random Forest e MPL geradas usando o *cut-off* de 0.75 (ver texto). SVM e MLP equivalem a Support-Vector Machine e Multi-layer Perceptron Classifier, respectivamente.

Os modelos de classificação utilizam características observadas a fim de predizer a qual classe pertence uma determinada observação cuja característica seja desconhecida. No presente caso, em que se pretende predizer a cor de um indivíduo, todas as variáveis preditivas são índices onomásticos, conforme apresen-

19. Jerome Friedman, Trevor Hastie, Robert Tibshirani et al., *The elements of statistical learning* (Springer series in statistics New York, 2001), 394-397.

20. Ver Max Kuhn e Kjell Johnson, *Applied predictive modeling* (Springer, 2013), Capítulo 16 e Jason Brownlee, *Imbalanced classification with Python: better metrics, balance skewed classes, cost-sensitive learning* (Machine Learning Mastery, 2020)

21. F. Pedregosa et al., "Scikit-learn: Machine Learning in Python", *Journal of Machine Learning Research* 12 (2011): 2825–2830.

tados na Tabela 4.10. O fato de serem usadas apenas variáveis elaboradas a partir dos nomes dos indivíduos explica-se tanto pela limitação dos dados quanto pela forma como os modelos de classificação funcionam. Algoritmos de *machine learning* têm, primeiro, seus parâmetros estimados em um grupo de treino e, posteriormente, são aplicados em um novo grupo de teste. No presente caso, conhece-se a variável de resposta na amostra de teste e, portanto, tem-se um caso de *supervised learning*.²² Esse fato é importante porque significa que os grupos de treino e aquele em que o modelo será implementado devem ter as mesmas variáveis, exceto pela variável de resposta —no caso, cor da pele. Por essa razão, é preciso limitar as variáveis preditivas àquelas que podem ser obtidas nos dois bancos de dados. Assim, ainda que o modelo tenha sido treinado com os dados do censo municipal de 1907, não é possível utilizar outras informações além do nome dos indivíduos, uma vez que variáveis como profissão e idade não estão disponíveis na Estatística Agrícola.

Tabela 4.10: Variáveis preditivas utilizadas nos modelos

Índice	Descrição
Primeiro nome	<i>dummy</i> se indivíduo tem primeiro nome negro
Último nome	<i>dummy</i> se indivíduo tem último nome negro
Cumprimento	Número de caracteres da string sem espaço
Dimensão	Número de substrings separadas por espaço
Razão consoante	Razão entre o número de consoantes e o comprimento
Tradicional	Número de sobrenomes tradicionais
Oligarquia	<i>dummy</i> se tem sobrenome composto de oligarquias
Devocional	Número de nomes devocionais

Notas: Ver texto para quais nomes foram considerados pertencentes à oligarquia e devocionais.

Os índices onomásticos apresentados na Tabela 4.10 são explicados na segunda coluna. Apenas os índices “oligarquia” e “devocional” parecem requerer explicação adicional. O índice “oligarquia” consiste em uma variável *dummy* que assume valor 1 caso o nome do indivíduo contenha um sobrenome composto de uma família tradicional paulista, e zero caso contrário. Os sobrenomes compostos

22. Sarah Guido e Andreas Müller, *Introduction to machine learning with python*, vol. 282 (O’Reilly Media, 2016), 26.

utilizados para a construção desse índice foram obtidos na *Genealogia Paulistana*.²³ Para a listagem de sobrenomes devocionais, utilizou-se como referência o trabalho de Lucas Costa Scottini.²⁴

Além de limitar as variáveis preditivas às variáveis que podem ser obtidas nos dois bancos de dados, também procurou-se definir uma amostra de treinamento do modelo que fosse similar, em termos demográficos, às características da base de dados de proprietários rurais. Assim, foram objeto de análise apenas homens maiores de 20 anos. Estima-se que apenas 5% dos proprietários de terra nos municípios limítrofes com São Carlos eram mulheres. Nessas circunstâncias, estimar um modelo em uma amostra de treino com aproximadamente 50% de indivíduos do sexo feminino poderia levar a um viés indesejado. Por esse motivo, as mulheres foram excluídas da base de proprietários rurais e o modelo foi implementado apenas com a população masculina. Outro critério demográfico adotado se distingue do realizado na seção anterior: os grupos de pretos e mulatos foram agregados em uma única classe de negros. Um dos motivos para esse procedimento é que os modelos de classificação testados mostraram limitada capacidade de distinguir indivíduos dos dois grupos.

As principais métricas de desempenho do modelo MLP são apresentadas na Tabela 4.11. O modelo alcançou acurácia global de 85%, implicando que para cada 100 proprietários rurais na amostra de treino, o modelo previu corretamente a cor da pele de 85 proprietários. Elevadas taxas de acurácia global, contudo, são esperadas em modelos treinados em amostras caracterizadas por desbalanceamento de classes. Do total de 2.176 indivíduos da amostra de treinamento, apenas 299 são negros. O inconveniente causado por esse desbalanceamento de classes fica claro quando se considera que um modelo que prevê que todos os indivíduos são brancos teria 86% de acurácia.²⁵

Para contornar diagnósticos excessivamente otimistas na presença de classes

23. Ver Luiz Gonzaga da Silva Leme, *Genealogia paulistana* (Duprat & comp., 1904). Para a pesquisa, foi consultada a versão online do trabalho de Silva Leme que, entre outros recursos, lista alfabeticamente os nomes das famílias tradicionais paulistas.

24. Lucas Costa Scottini, "O que o Nome nos ensina? Padrões sociais e raciais de nomes e sobrenomes e performance escolar em São Paulo" (diss. de mestrado, Universidade de São Paulo, 2011), 48.

25. Kuhn e Johnson, *Applied predictive modeling*, Capítulo 16.

Tabela 4.11: Métricas de desempenho da classificação com *Multi-layer Perceptron Classifier*

Métrica	Valor
Acurácia	85,0
Acurácia Balanceada	58,9
Kappa (Cohen)	0,24
AUC (ROC)	0,74
Precisão	43,0
Sensibilidade	24,0
F1-score	31,0
Frequência predita de negros	Frequência observada de negros
168	299

Notas: Métricas calculadas usando a amostra de teste. Precisão e F1-score para classe principal (negros).

desbalanceadas, outras métricas foram implementadas para avaliar os resultados dos modelos. Duas medidas apresentadas na Tabela 4.11 são de particular interesse nesse contexto. A primeira é o Kappa de Cohen. Essa estatística mede o grau de concordância das predições do modelo com as predições esperadas, com base nos totais marginais da matriz de confusão (Figura 4.2). A matriz de confusão é obtida a partir da tabulação cruzada das classes previstas e observadas. No presente caso, em que há duas classes, a matriz de confusão é de ordem 2x2, sendo composta pelos casos Verdadeiro-Positivo (TP), Falso-Negativo (FN), Falso-Positivo (FP) e Verdadeiro-Negativo (TN).²⁶ A estatística de Kappa varia de -1 a 1, com valores negativos significando ausência de concordância do modelo utilizado com as predições esperadas e valores positivos indicando concordância.²⁷ A implementação do modelo MPL resultou em um Kappa de 0,24, o que indica moderada concordância com as predições esperadas (Tabela 4.11).

Uma segunda métrica importante no contexto de classes desbalanceadas é a “acurácia balanceada”. Essa medida é simplesmente o valor médio das taxas de sensibilidade —proporção de indivíduos classificados como negros em relação

26. Usamos as siglas no original em inglês (TP, FN, FP, TN) porque são a forma como esses termos comumente aparecem na literatura

27. Kuhn e Johnson, *Applied predictive modeling*, 255.

a todos os negros observados —, e de especificidade—proporção de indivíduos classificados como brancos em relação a todos os brancos observados.²⁸ Os valores das taxas de sensibilidade e especificidade podem ser recuperados a partir da matriz de confusão (Figura 4.2). Nos termos da matriz de confusão, a sensibilidade é o número de predições positivas corretas dividido pelo número total de positivos ($\frac{TP}{TP+FN}$), enquanto a especificidade é calculada como o número de predições negativas corretas dividido pelo número total de negativas ($\frac{TN}{TN+FP}$).

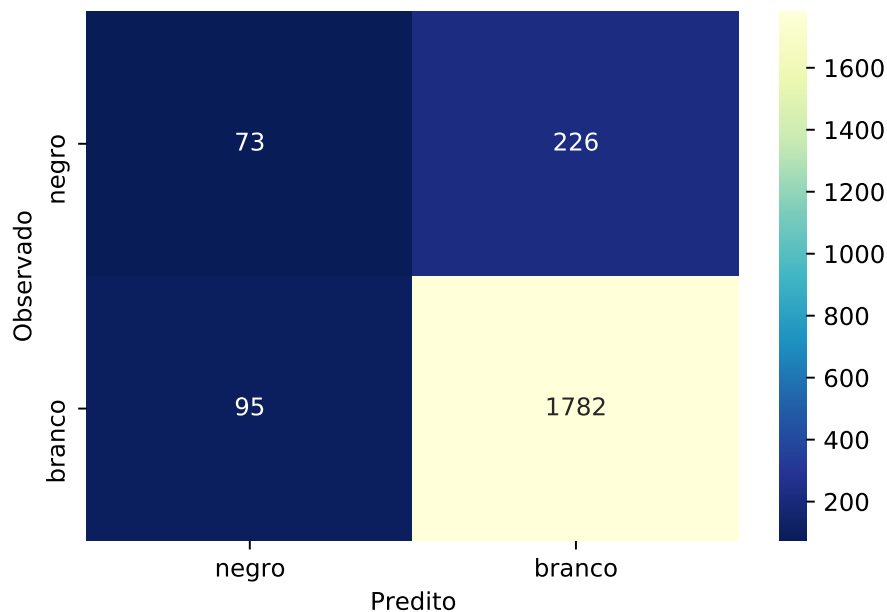


Figura 4.2: Matriz de Confusão

A partir da matriz de confusão, é possível notar que a especificidade do modelo é de 0,95, enquanto a sensibilidade é de 0,24, o que explica a acurácia balanceada de 59,7%. Esta combinação de valores de sensibilidade e especificidade é um resultado direto do corte definido para se considerar um indivíduo negro. Assim como outros modelos de classificação probabilísticos no contexto de variáveis de resposta binária, o MLP gera um vetor bidimensional de probabilidades que atribui, a cada observação, a probabilidade de pertencer a cada uma das classes.²⁹

28. Kuhn e Johnson, *Applied predictive modeling*, 256.

29. *Ibid.*, 247.

A decisão da classe prevista de cada observação é feita após estabelecer-se um limite a partir do qual a observação será considerada como pertencente à classe positiva ou negativa. Normalmente, usa-se 0.5 como o ponto de corte padrão, mas é possível utilizar outros limites, conforme o objetivo do estudo. No presente caso, considerou-se como negro apenas indivíduos com probabilidade de ser negro maior do que 75%.³⁰ Implementar esse critério representa um movimento ao longo da ROC (ilustrado na Figura 4.3) que, no presente caso, significa diminuir a sensibilidade do modelo.

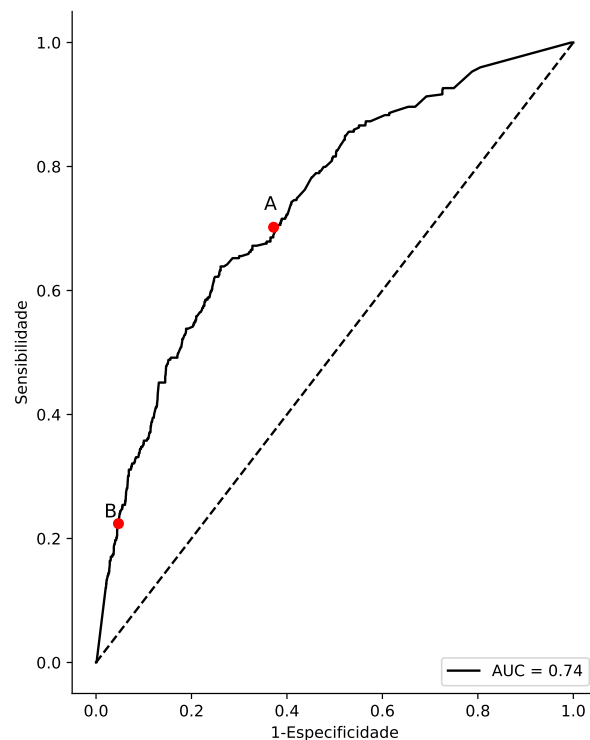


Figura 4.3: ROC - *Multi-layer Perceptron Classifier*

Notas: Curva de Característica de Operação do Receptor (ROC, na sigla em inglês). A Figura ilustra o *trade-off* entre sensibilidade e especificidade. O ponto A indica a combinação de sensibilidade e especificidade para um *cut-off* de 0.5, enquanto o ponto B indica a combinação associada ao *cut-off* de 0.75. Este último limite de aceitação foi o utilizado no modelo final.

A decisão de usar um limite de aceitação maior e, por consequência, subes-

30. Kuhn e Johnson, *Applied predictive modeling*, 424.

timar o número de proprietários negros foi tomada tendo em vista a hipótese de que o acesso à terra entre ex-escravos era raro. Atribuiu-se, assim, um peso maior ao erro de classificar proprietários brancos como negros do que ao erro de classificar proprietários negros como brancos. Dessa forma, tentou-se maximizar a probabilidade de um indivíduo ser negro, dado que ele foi classificado como negro (precisão). O custo consciente dessa opção foi ter um modelo que detecta negros com menor frequência e, por consequência, tem menor sensibilidade. O objetivo aqui, portanto, é mostrar que mesmo usando um modelo que subestima a presença de negros entre proprietários de terra, foi possível encontrar um número significativo de proprietários de terra que eram descendentes de escravizados.

É importante notar ainda que erros de classificação no algoritmo são esperados, sendo necessário apenas a identificação da direção do viés das estimativas, para que os resultados possam ser adequadamente interpretados. Desse modo, é importante destacar que, embora houvesse explícita correlação entre nomes e cor da pele no período do pós-abolição, é improvável que apenas essa informação seja suficiente para classificar um indivíduo como branco ou negro com precisão absoluta. A AUC do modelo, valor da área abaixo da Curva ROC que indica o grau de separação dos dois grupos analisados, apresentada na Figura 4.3, foi 0,74. Considerando que uma AUC=1 significa perfeita separação entre os dois grupos que se pretende classificar, e uma AUC=0.5 indica capacidade nula de distinção entre os grupos, então um valor de 0,74 demonstra que foi possível separar os dois grupos analisados apenas parcialmente, mas de maneira razoável. A taxa de precisão do grupo de negros ($\frac{TP}{TP+FP}$) foi de 0,43, o que é ilustrativo dessa limitação do modelo de classificação. Estima-se, assim, que a chance de ocorrência de negros no grupo de indivíduos classificados como negros é de aproximadamente 8/10 para 1. Esse valor é consideravelmente maior do que a chance de ocorrência na população, que é de 1/10 para 1, mas, como a razão é um pouco inferior a 1, significa que provavelmente um pouco mais da metade do grupo de indivíduos classificados como negros é composto de fato por brancos.³¹ Essa relação também

31. A chance de ocorrência de um evento é dada por

$$\alpha = \frac{P(A)}{P(A^c)}$$

pode ser observada na matriz de confusão.

Em contrapartida, como pode ser inferido da observação da figura 4.2, a precisão do modelo para classificar brasileiros brancos foi de 0,89 $\left(\frac{1782}{226+1782}\right)$, o que representa uma chance de ocorrência de 8 para 1 no grupo de brancos. Com efeito, é mais prudente interpretar os resultados desta seção como uma comparação entre um grupo de proprietários exclusivamente brancos com um grupo em que negros são aproximadamente a metade dos proprietários. Será visto a seguir que, mesmo considerando o sucesso parcial do algoritmo na separação dos grupos de interesse, é possível identificar padrões semelhantes àqueles observados para São Carlos.

4.2.3 Classificação nos municípios limítrofes de São Carlos

As predições do modelo de *machine learning* implementado permitem agora analisar as condições do acesso à terra entre brasileiros brancos, imigrantes e negros nos municípios limítrofes de São Carlos. Os municípios são: Araraquara, São Simão, Santa Rita do Passa Quatro, Ribeirão Bonito, Brotas, Rio Claro, Anápoles e Belém do Descalvado. A apresentação é análoga à que foi feita na seção 4.1, quando foram analisados os dados de proprietários de terra apenas no município de São Carlos. Três diferenças importantes, contudo, aparecem agora em relação aos dados apresentados nesta seção em relação à seção 4.1. A primeira é que, ao longo de toda esta seção, a variável de cor da pele é uma predição, baseada no melhor algoritmo de classificação descrito nas seções anteriores. Em segundo lugar, desta vez a análise concentra-se nas distribuições das variáveis de interesse, em vez de casos individuais. Essa ênfase acontece porque há uma incerteza maior em relação à cor da pele dos indivíduos, posto que a acurácia do modelo sugere que existem negros classificados como brancos e brancos classificados como negros. Dessa forma, estatísticas descritivas, como média e desvio-padrão, refletem melhor a realidade do que informações individuais. Por fim, esta seção limita-se a analisar a dimensão do acesso à terra. A impossibilidade da análise baseada em indivíduos ou que são muito afetadas por *outliers* inviabiliza um exame mais

Ver Sheldon Ross, *A first course in probability* (Pearson, 2014), 72.

detalhado das diferenças qualitativas do acesso à terra entre brasileiros brancos, imigrantes e negros no município limítrofes, como foi feito na seção 4.1 para São Carlos.

A Tabela 4.12 resume os dados da distribuição da área total entre brasileiros brancos, imigrantes e negros, segundo o modelo de classificação utilizado. O número de observações aponta que 12% (200) das propriedades rurais nos municípios limítrofes com São Carlos pertenceriam a negros. Quando foram analisados os dados reais de São Carlos, verificou-se que 8 (4,1%) das 192 propriedades rurais levantadas estavam em posse de negros. Assim, mesmo usando um modelo de classificação que deliberadamente subestima o número de proprietários negros, encontrou-se um número expressivo de propriedades rurais em posse de negros. Quando se analisa a informação de indivíduos e não a de propriedades rurais, o número previsto de proprietários únicos entre negros foi de 198, representando 12,7% do total.

Tabela 4.12: Distribuição da área de propriedades segundo o melhor modelo de classificação,
Municípios limítrofes de São Carlos, 1905

	Brasileiros brancos	Imigrantes	Negros
Observações	936	529	200
Média	123,7	64,5	97,9
desvio-padrão	335,9	211,8	239,8
min	1	1,0	2,0
10%	8	4,4	5,0
25%	15	8	10,0
50%	40	20	22,5
75%	120	45	72,5
90%	300	120	234,7
Máximo	7000	3000	2500

Fontes: São Paulo, Estatística agrícola e zootechnica do estado de São Paulo no ano agrícola de 1904–1905.

Notas: Classificação de brancos e negros feita conforme algoritmo MLP.

Municípios: Araraquara, São Simão, Santa Rita do Passa Quatro, Ribeirão Bonito, Brotas, Rio Claro, Anapólis, Belém do Descalvado.

Outra informação importante que pode ser depreendida da Tabela 4.12 é que, até a mediana, a distribuição da área total de imigrantes e negros é parecida en-

tre si e muito diferente da distribuição da área total entre brasileiros brancos. No primeiro decil, por exemplo, brasileiros brancos têm 8 hectares de terras, enquanto negros e imigrantes têm, respectivamente, 5 e 4,4 hectares. Essa tendência também está de acordo com o que foi observado nos dados observados para São Carlos, podendo ser percebida pela análise das diferenças do percentis das distribuições de cada um dos grupos. A mediana (2º quartil, 50%), por exemplo, de brasileiros brancos foi de 40 alqueires de terra, enquanto para imigrantes e negros, a mediana foi de 20 e 22, respectivamente. Como já observado, a distribuição da área total nos municípios vizinhos a São Carlos é caracterizada por muitas observações na base da distribuição e poucas observações no topo. Assim, à medida que se caminha ao longo da distribuição, observam-se áreas totais cada vez maiores. Esse efeito é observado nas três distribuições aqui discutidas, mas é uma tendência mais expressiva para brasileiros brancos e negros. É possível notar esse fato observando a razão entre quartis e decis das distribuições. Como visto, a área total mediana de brasileiros brancos é o dobro da mediana de imigrantes e negros (40 contra 20 e 22,5, respectivamente). No terceiro quartil, porém, a razão entre brasileiros brancos (120) e negros (72,5) se mantém, mas aumenta para quase 3 na comparação entre brasileiros brancos e imigrantes (45).

Desta forma, como visto acima, a distribuição da área de propriedades de imigrantes e negros possui percentis similares até a mediana. Por outro lado, para observações acima da mediana, a distribuição da área total de brasileiros negros começa a convergir para a distribuição de brasileiros brancos (por exemplo, no percentil 75%, negros tem 72,5 hectares, valor mais próximo ao observado para brasileiros brancos (120) que os 45 hectares observados para imigrantes). A Figura 4.4 mostra uma representação gráfica desse fato. É possível notar que, inicialmente, a distribuição acumulada de imigrantes e negros é bastante similar. Ou seja, para quaisquer valores de área entre zero e 20 alqueires, a probabilidade acumulada de se observar uma propriedade rural de determinado tamanho é aproximadamente a mesma para imigrantes e negros. Contudo, a partir da mediana, a distribuição acumulada da área total de propriedades rurais sob a posse de negros se desloca para a direita e se aproxima do perfil observado para brasileiros brancos.

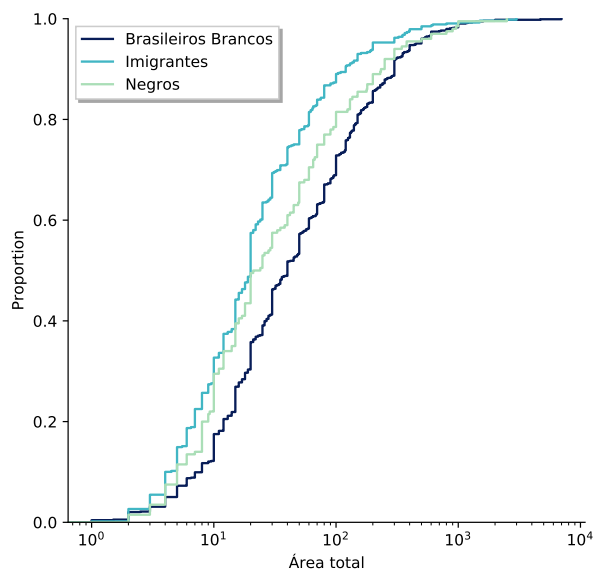


Figura 4.4: Distribuição acumulada da área total entre Brasileiros brancos, Imigrantes e Negros, São Carlos, 1907

Fontes: São Paulo, Secretaria de Agricultura, Commercio e Obras Publicas do Estado de São Paulo. (1906-1907). Estatística agrícola e zootécnica do estado de São Paulo no anno agrícola de 1904-1905. São Paulo: Tip. Brazil.

Nota: Eixo x em escala logarítmica. Classificação de brancos e negros feita conforme algoritmo MLP.

Por qual razão a distribuição da área total de propriedades de negros se desloca para a direita depois da mediana, ou seja, por que o tamanho das áreas de negros e brasileiros brancos se tornam progressivamente similares? Uma possibilidade é que a desigualdade fundiária entre brasileiros negros fosse de fato mais elevada quando comparada com imigrantes. Os dados indicam que a razão dos percentis 90/10 da área total entre negros era de 46 (234,7/5), enquanto esse mesmo valor é de 27 (120/4,4) para imigrantes. O motivo mais provável, porém, é que alguns brancos classificados como negros pelo algoritmo tenham propriedades rurais relativamente maiores, o que faz com que o procedimento desta seção superestime as medidas da distribuição da área total entre negros.

Pode-se ilustrar as consequências da classificação de proprietários brancos como negros, analisando-se casos individuais. Um exemplo é o proprietário “João Borba”, do município de Araraquara. O algoritmo de classificação atribuiu uma

probabilidade de aproximadamente 80% de João Borba ser negro, baseado em seu nome. João Borba tem dois latifúndios registrados na Estatística Agrícola, um de 600 alqueires e outro de 1000 alqueires. Se esse proprietário fosse na verdade branco, a média da área de propriedades de negros cairia de 98 para 90 e a razão dos decis 90/10, saindo de 46, passaria a ser de 41. Esse e outros casos tornam possível que a distribuição de propriedades rurais de negros tivesse média e desigualdade inferiores ao que se observa aqui.

Os resultados do modelo aplicado para a região em torno de São Carlos reforçam o notado a partir dos dados reais de São Carlos. Brasileiros brancos parecem ter formado, de fato, um grupo à parte, sendo proprietários de terras maiores do que as médias estimadas tanto para imigrantes quanto para negros nos municípios limítrofes. No entanto, algumas diferenças também foram observadas entre os resultados desta e da seção 4.1. Em particular, observou-se a presença de latifundiários entre os proprietários negros, o que tornou a distribuição fundiária dentro desse grupo altamente desigual. Pelos motivos apontados, esse resultado deve ser considerado com cautela. A classificação de alguns proprietários brancos como negros pode ter distorcido a distribuição de terras entre negros, fazendo o quadro geral desse grupo parecer mais favorável do que de fato era.

4.3 Conclusão

A acumulação de capital ao longo do ciclo de vida de um indivíduo é a medida mais clara de seu sucesso no mercado de trabalho e em outras dimensões de sua vida econômica. O mesmo é verdade para grupos demográficos. Em uma sociedade essencialmente rural, a acumulação de capital reflete-se em larga medida na propriedade da terra. Neste capítulo, procurou-se avaliar o sucesso econômico absoluto e relativo de brasileiros brancos, imigrantes e negros do Oeste paulista no pós-abolição. Os resultados mostraram que, apesar de brasileiros negros serem naturais da terra há várias gerações, a taxa de acesso à terra desse grupo não diferia significativamente daquelas de imigrantes recém-chegados. A incorporação de ex-escravos, inteiramente destituídos de propriedade pela abolição sem reparações em 1888, deve ter reforçado esse perfil de

desigualdade da população negra frente à população branca brasileira.

Uma diferença importante entre negros e imigrantes, no entanto, aparece no valor das propriedades. A análise dos dados de São Carlos mostra que os indicadores de qualidade da terra e tamanho médio da propriedade eram maiores para propriedades rurais de negros em relação a imigrantes. Consequentemente, o valor médio das propriedades de negros era maior do que o de proprietários imigrantes.

Um resultado claro da análise deste capítulo é que brasileiros brancos dominavam a propriedade rural da parte analisada do Oeste paulista à época. Esse grupo controlava não apenas a maior área das propriedades disponíveis, como também tinha acesso às melhores terras, o que permitia, entre outras coisas, dedicar parcela relevante da terra para o café, lavoura que permanecia a mais lucrativa na primeira década do século XX.

Dessa forma, os resultados deste capítulo demonstram que no Oeste paulista, duas décadas após a Abolição, a situação material de ex-escravos e descendentes não era muito diferente daquela prevalecente entre imigrantes europeus. Junto com as evidências apresentadas nos Capítulos 2 e 3, que indicam, respectivamente, que o acesso à educação e ao mercado de trabalho para as duas populações se dava em condições parecidas, os resultados deste capítulo sugerem que a origem das disparidades econômicas entre descendentes de europeus e de escravizados deve ser encontrada em um momento posterior ao que foi analisado, ou seja, após a década de 1900.

Conclusão

Em São Carlos da primeira década do século XX, a situação socioeconômica de ex-escravos e descendentes não era muito diferente daquela vivida por imigrantes recém-chegados no que diz respeito ao acesso à educação primária, oportunidades no mercado de trabalho e acesso à terra. Uma maneira de interpretar esse resultado é constatar que muitas afirmações a respeito da população negra brasileira que viveu o pós-abolição baseiam-se em evidências limitadas e esparsas. Os resultados do Capítulo 2 desta tese, por exemplo, mostram que mesmo tendo as taxas mais altas de analfabetismo na população e sofrendo com violência e racismo, negros procuraram matricular seus filhos nas escolas primárias. Essa é uma atitude que não se ajusta ao descrito por toda uma literatura que argumentou que ex-escravizados não valorizariam a educação e a busca de aperfeiçoamento. A “anomia”, tese já bastante criticada pela literatura, surge aqui como inconsistente e equivocada frente às evidências históricas.³²

O presente trabalho mostra que não só negros enfrentaram as adversidades existentes para matricular seus filhos nas escolas primárias, como também alcançaram melhores posições em um mercado de trabalho ainda profundamente marcado pelo legado da escravidão. Como demonstrado no Capítulo 3, mulheres negras ainda eram desproporcionalmente representadas entre cozinheiras e lavadeiras, e a ocupação de criado era muito comum entre crianças negras. No entanto, ex-escravos e descendentes ainda estavam relativamente bem representados em ocupações médias e qualificadas, como visto para as ocupações de colono, maquinista e carpinteiro. Mais que isso, partindo de condições patrimoniais pro-

32. Petrônio Domingues, *Protagonismo negro em São Paulo: história e historiografia* (Edições Sesc SP, 2019), Capítulo 1; Andrews, *Negros e Brancos*, 119-120.

vavelmente precárias, negros tornaram-se um grupo relevante de proprietários rurais e produtores de café, como demonstrado no Capítulo 4.

Uma outra maneira de interpretar os resultados da tese, complementar à primeira, é que o fato das condições socioeconômicas de brasileiros negros não diferirem significativamente de imigrantes, mas serem diametralmente opostas às aquelas observadas para brasileiros brancos, indica que brasileiros negros eram tão estranhos ao próprio país quanto estrangeiros recém-chegados. Como se ex-escravos e descendentes fossem estrangeiros na própria terra, como se o país que habitavam tivesse sido construído por eles, mas não para eles. Novamente, o exemplo da educação primária ilumina o argumento: o fato de mulheres negras serem o subgrupo demográfico com os menores níveis de educação foi uma causa direta das políticas educacionais do século XIX e, em particular, da disposição do Estado brasileiro, em suas diversas instâncias, de dificultar ou impedir o acesso de escravizados à educação. Como visto no Capítulo 2, o baixo nível educacional de mulheres negras parece ter sido o principal obstáculo à matrícula escolar de crianças negras em escolas primárias.

A analogia com a situação de estrangeiros pode se estender para a questão fundiária. A condição jurídica de populações escravizadas impossibilitou a propriedade de terra e, mesmo em situações excepcionais em que a propriedade de terra tenha se verificado entre negros livres na ordem escravocrata, a fragilidade da situação material e legal certamente dificultou a acumulação e manutenção do patrimônio familiar ao longo do tempo. Assim, mesmo que seja possível encontrar proprietários negros no pós-abolição, como demonstrado nos dados apresentados no Capítulo 4, esses proprietários provavelmente encontravam-se na condição de pioneiros em seu âmbito familiar. Situação não muito diferente dos imigrantes pobres, sem recursos e pecúlio acumulados, que haviam acabado de chegar em São Paulo.

Há, porém, limites para o paralelo entre a situação de imigrantes e negros no pós-abolição. O mais evidente deles é o racismo. Como demonstrado por Karl Monsma em seu estudo sobre São Carlos, a população negra sofria desproporcionalmente com violência.³³ Para o autor, a violência exercia a função social de

33. Monsma, *A reprodução*, Cap. 5.

delimitar espaços em que negros podiam e não podiam participar da sociedade. Dessa forma, a violência não teve apenas um efeito deletério imediato, refletido nas maiores taxas de mortalidade e no sofrimento físico e emocional de crianças e adultos negros, mas também produziria consequências que se propagaram nas gerações futuras ao construir e sedimentar lugares sociais de incluídos e excluídos na sociedade.

Um possível exemplo do mecanismo sugerido por Karl Monsma ocorre na educação. Relembrando sua experiência escolar, o cantor e compositor brasileiro Gilberto Gil relata um episódio em que, ao tentar fazer uma pergunta em sala, o professor retrucou “Cala boca, negro boçal!”.³⁴ O episódio ocorreu na década de 1950, mas não há razão para crer que casos como esse não fossem ainda mais frequentes e violentos 20 anos após a Abolição. Teria a violência no âmbito escolar dificultado acumulação de capital humano por parte da população de ex-escravos e descendentes? Como esta tese apresenta apenas dados da matrícula escolar e não examina indicadores de permanência, como frequência escolar e taxas de reprovação, não há meios de responder a essa pergunta no presente estágio da pesquisa.

Apesar de descrever um aspecto fundamental da situação de ex-escravizados e seus filhos, a tese que vê na violência e no racismo nas relações interpessoais as principais causas da desigualdade entre negros e descendentes de imigrantes tem limitações. Como já comentado na introdução desta tese, nos Estados Unidos, a população negra conseguiu diminuir a diferença em educação, renda e riqueza em relação à população branca nas décadas que se seguiram à Emancipação.³⁵ Não há dúvida sobre o volume de violência a que a população negra foi submetida no período. Entre 1882-1915, cerca de 3 mil negros foram mortos em linchamentos no Sul dos Estados Unidos.³⁶ Após a Guerra Civil, tentativas de construir escolas para negros frequentemente esbarravam em dura oposição de cidadãos

34. Simbolicamente, “boçal” é um termo que tem origem na escravidão e era usado para se referir a escravos não nascidos no Brasil. Para o relato do episódio, ver Gilberto Gil e Regina Zappa, *Gilberto bem perto* (Harper Collins Brasil, 2013), 58

35. Higgs, “Accumulation of Property”; Margo, “Comment and Further Evidence”; Higgs, *Competition*, 120,146.

36. Higgs, *Competition*, 90.

brancos que, em casos extremos, queimavam as escolas.³⁷ Embora a evolução econômica da população afro-americana pudesse ter sido maior caso a violência e hostilidade de cidadãos brancos estivessem ausentes, o fato de uma evolução favorável ter sido observada mesmo em um cenário de extrema violência indica que há limites no poder da coerção física de impedir avanços socioeconômicos de um subgrupo demográfico.

Esta tese argumenta que há razões para crer que outros fatores, além do racismo direto e da violência, tiveram efeito sobre as possibilidades econômicas de populações outrora escravizadas. Em particular, argumentou-se que, embora participassem da economia em condições similares, havia diferenças estruturais entre a população de imigrantes e a população de ex-escravos e descendentes. Em primeiro lugar, a população de imigrantes tinha taxas de crescimento demográfico maiores, o que fazia com que a taxa de crescimento da renda nas famílias de imigrantes fosse maior que a taxa de crescimento da renda de brasileiros negros, mesmo se não houvesse diferenças de produtividade do trabalho nos dois grupos. No longo prazo, essa diferença demográfica pode ter produzido efeitos dramáticos nas possibilidades de acúmulo de capital.

Outra dimensão que a população imigrante diferia de forma substancial da população negra brasileira é a que se refere à educação. Como visto no Capítulo 1, a taxa de alfabetização entre imigrantes era mais do que o dobro das taxas observadas para a população negra. Como argumentado no Capítulo 2, a alfabetização dos pais era um fator determinante da matrícula escolar de crianças. É provável que essas diferenças no nível de educação não tenham produzido efeitos imediatos durante o período de dominância da economia cafeeira, em função da conhecida baixa complementaridade entre capital humano e trabalho agrícola. Contudo, a urbanização, industrialização e crescente complexidade da economia nas décadas seguintes podem ter construído um perfil de demanda por trabalho que elevou o prêmio da educação, contribuindo para aumentar as desigualdades entre negros e descendentes de imigrantes.

É, portanto, evidente que a herança da escravidão deve também ser medida pelo impacto que essa instituição teve sobre as possibilidades de acumulação

37. Ransom e Sutch, *One kind*, 27.

de capital humano e físico das populações escravizadas. Não fosse a escravidão, a população negra brasileira teria melhores níveis educacionais, mais patrimônio material e melhores ocupações no mercado de trabalho no início do século XX, contando, assim, com condições melhores de competição econômica com a população branca. Além da comparação com brasileiros brancos e imigrantes, uma evidência adicional em favor da relação entre escravidão e indicadores socioeconômicos no pós-abolição é oferecida pela comparação entre pretos e mulatos. Como demonstrado ao longo da tese, os mulatos que, com sua ascendência mista parecem ter reduzido o peso do passivo escravocrata, tinham melhores níveis educacionais, melhores ocupações e, em geral, melhores condições de acesso à terra.

Mais de 100 anos depois dos fatos e estatísticas narrados nesta tese, indicadores socioeconômicos da população negra continuam em patamares mais baixos que aqueles observados para a população branca.³⁸ Em larga medida, essa situação deve-se à falha do Estado brasileiro em diminuir, ao longo do século XX, as desigualdades educacionais e de capital construídas ao longo do período escravocrata.

38. Instituto Brasileiro de Geografia e Estatística, “Desigualdades sociais por cor ou raça no Brasil”, *Estudos e Pesquisas-Informação Demográfica e Socioeconômica* 41 (2019).

Referências

Fontes Primárias

Livros de chamadas, 1906-1909, Livros: E-3910, E-3912, E-3914, E-3916, E-3674, São Paulo, Arquivo Público do Estado de São Paulo.

Piza, Antonio de Toledo. *Relatorio do anno de 1900 apresentado em 13 de janeiro de 1902 ao cidadão Dr. Bento Pereira Bueno, Secretario de Estado dos Negocios do Interior e da Justiça*. São Paulo: Typographia do Diário Official 1903

São Carlos, Recenseamento da população de S. Carlos do Pinhal no anno de 1907, de accordo com a lei n. 126 votada pela Camara Municipal em sessão de 4 de fevereiro de 1907 (São Carlos: Manuscrito, 1907). 8 v.

São Paulo, Secretaria de Agricultura, Commercio e Obras Publicas do Estado de São Paulo. (1906-1907). *Estatistica agricola e zootechnica do estado de São Paulo no anno agricola de 1904-1905*. São Paulo: Tip. Brazil.

Bibliografia

- Abramitzky, Ran, Leah Platt Boustan, Katherine Eriksson, James J Feigenbaum e Santiago Pérez. *Automated linking of historical data*. Relatório técnico. National Bureau of Economic Research, 2019.
- Ananias, Mauriceia. “A legislação da instrução pública primária na província de São Paulo: 1834-1868”. Tese de doutorado, Universidade Estadual de Campinas, 2005.
- Andrews, George Reid. *Afro-Latin America: Black Lives, 1600-2000*. Harvard University Press, 2016.
- . *Negros e brancos em São Paulo, 1888-1988*. Editora da Universidade do Sagrado Coração, 1998 [1991].
- Bacellar, Carlos de Almeida Prado. “O apogeu do café na Alta Mogiana”. Em *Na Estrada do Anhanguera: uma visão regional da história paulista*. Humanitas Publicações FFLCH/USP, 1999.
- Ball, Molly Catherine. “Inequality in São Paulo’s Old Republic: A Wage Perspective, 1891-1930”. Tese de doutorado, University of California, Los Angeles, 2013.
- Barbosa, Ana AL, Iglénir J Cavalli, Kiyoko Abé, Maria G Santos e Eliane S Azevêdo. “Family names and the length of the Y chromosome in Brazilian blacks”. *Brazilian Journal of Genetics* 20 (1997).

- Barros, Surya Aaronovich Pombo de. “Negrinhos que por ahi andão: a escolarização da população negra em São Paulo (1870-1920)”, 2005.
- Brownlee, Jason. *Imbalanced classification with Python: better metrics, balance skewed classes, cost-sensitive learning*. Machine Learning Mastery, 2020.
- Camargo, Jose Francisco de. *Crescimento da população no Estado de São Paulo e seus aspectos econômicos*. Vol. 1. São Paulo: Instituto de Pesquisas Econômicas, 1981.
- Cánovas, Marília Dalva Klaumann. “A emigração espanhola e a trajetória do imigrante na cafeicultura paulista: o caso de Villa Novaes, 1880-1930.” Diss. de mestrado, Universidade de São Paulo, 2001.
- Chalhoub, Sidney. *Visões da liberdade: uma história das últimas décadas da escravidão na corte*. Editora Companhia das Letras, 2011.
- Colistete, Renato Perim. *O atraso em meio à riqueza: uma história econômica da educação primária em São Paulo, 1835 a 1920*. São Paulo, 2016.
- Cook, Lisa D. *African american inventors data set*, 2004.
- . “Inventing social capital: Evidence from African American inventors, 1843–1930”. *Explorations in Economic History* 48, n. 4 (2011): 507–518.
- Cook, Lisa D, Trevon D Logan e John M Parman. “Distinctively black names in the American past”. *Explorations in Economic History* 53 (2014): 64–82.
- Da Costa, Emília Viotti. *Da senzala à colônia*. 5ª ed. Unesp, 2010 [1997].
- Da Silva, Adriana Maria Paulo. “A escola do professor Pretextato dos Passos e Silva: questões a respeito das práticas de escolarização no mundo escravista”. *A história da educação dos negros no Brasil*. Niterói: EdUFF, 2016, 141–162.
- De Azevedo, Celia Maria Marinho. *Onda negra, medo branco: o negro no imaginário das elites—século XIX*. Vol. 6. Annablume, 1987.
- De Souza, Rosa Fátima. *Templos de civilização: a implantação da escola primária graduada no Estado de São Paulo, 1890-1910*. Unesp, 1998.

- Dean, Warren. *Rio Claro: um sistema brasileiro de grande lavoura, 1820-1920*. Rio de Janeiro: Paz e Terra, 1977.
- Denis, Pierre. *o Brazil no seculo XX*. 1909.
- Domingues, Petrônio. *Protagonismo negro em São Paulo: história e historiografia*. Edições Sesc SP, 2019.
- Fernandes, Florestan. *A integração do negro na sociedade de classes*. Vol. 1. São Paulo: Dominus editora, 1965.
- Fonseca, Marcus Vinicius. "Pretos, pardos, crioulos e cabras nas escolas mineiras do século XIX." Tese de doutorado, Universidade de São Paulo, 2007.
- Friedman, Jerome, Trevor Hastie, Robert Tibshirani et al. *The elements of statistical learning*. Springer series in statistics New York, 2001.
- Fryer Jr, Roland G, e Steven D Levitt. "The causes and consequences of distinctively black names". *The Quarterly Journal of Economics* 119, n. 3 (2004): 767–805.
- Gaffre, Louis Albert. *Visions du Brésil*. São Paulo: F. Alves & Cia., 1912.
- Geografia e Estatística, Instituto Brasileiro de. "Desigualdades sociais por cor ou raça no Brasil". *Estudos e Pesquisas-Informação Demográfica e Socioeconômica* 41 (2019).
- Gibson, Campbell, e Kay Jung. *Historical census statistics on population totals by race, 1790 to 1990, and by Hispanic origin, 1790 to 1990, for the United States, regions, divisions, and states*. US Census Bureau Washington, DC, 2002.
- Gil, Gilberto, e Regina Zappa. *Gilberto bem perto*. Harper Collins Brasil, 2013.
- Guido, Sarah, e Andreas Müller. *Introduction to machine learning with python*. Vol. 282. O'Reilly Media, 2016.
- Higgs, Robert. "Accumulation of property by Southern blacks before World War I". *The American Economic Review*, 1982, 725–737.

- Higgs, Robert. *Competition and coercion: Blacks in the American economy 1865-1914*. Cambridge University Press, 1977.
- Holanda, Sérgio Buarque de. "Prefácio". Em *Memórias de um colono no Brasil*. (Autor: T. Davatz). São Paulo: Livraria Martins, 1941.
- Holloway, Thomas H. *Imigrantes para o café: café e sociedade em São Paulo, 1886-1934*. Rio de Janeiro: Paz e Terra, 1984.
- Jacino, Ramatis. "O negro no mercado de trabalho em São Paulo pós-abolição-1912/1920". Tese de doutorado, Universidade de São Paulo, 2012.
- Klein, Herbert S. *A imigração espanhola no Brasil*. Fapesp, 1994.
- . "A integração dos imigrantes italianos no Brasil, na Argentina e Estados Unidos". *Novos Estudos CEBRAP* 25 (1989): 95-117.
- Kowarick, Lúcio. *Trabalho e vadiagem: a origem do trabalho livre no Brasil*. Editora Brasiliense, 1987.
- Kuhn, Max, e Kjell Johnson. *Applied predictive modeling*. Springer, 2013.
- Lago, Luiz Aranha Corrêa do. *Da escravidão ao trabalho livre: Brasil, 1550-1900*. Editora Companhia das Letras, 2014 [1978].
- Lamounier, Maria Lúcia. "Between slavery and free labour: experiments with free labour and patterns of slave emancipation in Brazil and Cuba c. 1830-1888. 1993". Tese de doutorado, London School of Economics e Political Science, 1993.
- Lanza, André Luiz. "De braços para a lavoura a proprietários rurais: Imigrantes e o acesso à terra em São Paulo, 1886-1920". Tese de doutorado, Universidade de São Paulo, 2021.
- . "Imigrantes no Brasil e na Argentina: políticas de atração, fluxos, atividades e deslocamentos São Paulo e Buenos Aires, 1870-1930)". Diss. de mestrado, Universidade de São Paulo, 2015.

- Leff, Nathaniel H. "Economic development in Brazil, 1822-1913". Em *How Latin America Fell Behind: Essays on the Economic Histories of Brazil and Mexico, 1800-1914*. Stanford University Press, 1998.
- Levenshtein, Vladimir I. "Binary codes capable of correcting deletions, insertions, and reversals". 10, n. 8 (1966): 707-710.
- Lowrie, Samuel H. "The Negro Element in the Population of Sao Paulo, a Southernly State of Brazil". *Phylon (1940-1956)* 3, n. 4 (1942): 398-416.
- Lundquist, Jennifer Hickes, Douglas L Anderton e David Yaukey. *Demography: the study of human population*. Waveland Press, 2014.
- Lutz, Mark. *Learning python: Powerful object-oriented programming*. "O'Reilly Media, Inc.", 2013.
- Margo, Robert A. "Accumulation of Property by Southern Blacks Before World War I: Comment and Further Evidence". *American Economic Review*, 1984, 768-776.
- . *Obama, Katrina, and the persistence of racial inequality*. Relatório técnico. National Bureau of Economic Research, 2016.
- Mattos, Hebe. *Das cores do silêncio: os significados da liberdade no sudeste escravista*. 3ª ed. Editora Unicamp, 2013 [1995].
- Melo, Hildete Pereira de, João Lizardo de Araújo e Teresa Cristina de Novaes Marques. "Raça e nacionalidade no mercado de trabalho carioca na Primeira República: o caso da cervejaria Brahma". *Revista Brasileira de Economia* 57, n. 3 (2003): 535-569.
- Milliet, Sérgio. *Roteiro do café e outros ensaios: contribuição para o estudo da história econômica e social do Brasil*. 3ª ed. São Paulo: Departamento de Cultura, 1941.
- Monbeig, Pierre. *Pioneiros e fazendeiros de São Paulo*. São Paulo: HUCITEC, 1984.

- Monsma, Karl. *A reprodução do racismo: fazendeiros, negros e imigrantes no oeste paulista, 1880-1914*. São Carlos: Editora da Universidade Federal de São Carlos, 2016.
- . “Conflito simbólico e violência interétnica: europeus e negros no oeste paulista, 1888-1914”. *História em revista* 10, n. 10 (2004).
- Palma, Rogerio da, e Oswaldo Mário Serra Truzzi. “Renomear para Recomeçar: Lógicas Onomásticas no Pós-abolição”. *Dados* 61, n. 2 (2018): 311–340.
- Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel et al. “Scikit-learn: Machine Learning in Python”. *Journal of Machine Learning Research* 12 (2011): 2825–2830.
- Petrone, Maria Teresa Schorer. “Imigração assalariada”. *História geral da civilização brasileira* 2 (1967): 274.
- Petrone, Maria Tereza Schorer. “Imigração”. Em *O Brasil republicano: sociedade e instituições (1889-1930)*, vol. 3. Bertrand Brasil, 1977.
- Prado Junior, Caio. “Distribuição da propriedade fundiária rural no Estado de São Paulo”. *Revista Geografia* 1, n. 1 (1935).
- Prado Júnior, Caio. *História econômica do Brasil*. São Paulo: Brasiliense, 2006 [1945].
- Ransom, Roger L, e Richard Sutch. *One kind of freedom: The economic consequences of emancipation*. Cambridge University Press, 2001.
- Ross, Sheldon. *A first course in probability*. Pearson, 2014.
- Sanderson, Michael. *Education, economic change and society in England 1780-1870*. Cambridge University Press, 1995.
- Santos Gomes, Flávio dos. *Negros e política (1888-1937)*. Zahar, 2005.
- Schwarcz, Lilia Moritz. *Lima Barreto-triste visionário*. Editora Companhia das Letras, 2017.

- Schweninger, Loren. *Black property owners in the South, 1790-1915*. University of Illinois Press, 1990.
- Scottini, Lucas Costa. “O que o Nome nos ensina? Padrões sociais e raciais de nomes e sobrenomes e performance escolar em São Paulo”. Diss. de mestrado, Universidade de São Paulo, 2011.
- Sedgewick, Robert, e Kevin Wayne. *Computer science: An interdisciplinary approach*. Addison-Wesley Professional, 2016.
- Silva Leme, Luiz Gonzaga da. *Genealogia paulistana*. Duprat & comp., 1904.
- Souza, Bruno Gabriel Witzel de. “From Bonded Laborers to Educated Citizens? Immigration, Labor Markets, and Human Capital in São Paulo, Brazil (1820-2010)”. Tese de doutorado, Georg-August-Universität Göttingen, 2019.
- Stolcke, Verena. *Coffee Planters Workers And Wives: Class Conflict And Gender Relations On Sao Paulo Coffee Plantations*. Springer, 1988.
- Summerhill, William R. *Order against progress*. Stanford University Press, 2003.
- Tavares-Neto, José, e Eliane S Azevêdo. “Racial origin and historical aspects of family names in Bahia, Brazil”. *Human Biology*, 1977, 287–299.
- Truzzi, Oswaldo Mário Serra. *Café e indústria: São Carlos: 1850-1950*. São Carlos: Imprensa Oficial, 2007.
- . “Padrões de nupcialidade na economia cafeeira de São Paulo (1860-1930)”. *Revista Brasileira de Estudos de População* 29, n. 1 (2012): 169–189.
- Truzzi, Oswaldo Mário Serra, e Maria Silvia Beozzo Bassanezi. “População, grupos étnico-raciais e economia cafeeira: São Carlos, 1907”. *Revista Brasileira de Estudos de População* 26, n. 2 (2009): 197–218.
- Wooldridge, Jeffrey M. *Introductory econometrics: A modern approach*. 5^a ed. Cengage Learning, 2012.

Wright, Gavin. *Old South, new South: Revolutions in the southern economy since the Civil War*. Basic Books, 1986.

Apêndice A

Algoritmo de *matching* de nomes

Neste Apêndice detalha-se os algoritmos de *matching* usados para cruzar os dados dos livros de chamadas e da Estatística Agrícola com o censo municipal de São Carlos de 1907. Para fins de simplificação, chama-se o primeiro algoritmo, que cruza os nomes das listas de chamadas com o censo, de *matching dos estudantes*, enquanto o segundo algoritmo, que cruza os nomes dos proprietários da Estatística Agrícola com os nomes do censo, será chamado *matching de proprietários*.

Ambos os algoritmos estão divididos em dois estágios. Como o primeiro estágio é comum ao *matching* de estudantes e de proprietários, dedica-se a próxima seção à explicação dessa etapa. As seções A.2 e A.3 apresentam, respectivamente, os algoritmos de *matching* do segundo estágio de estudantes e proprietários.

A.1 Primeiro estágio

Para a melhor compreensão do algoritmo, assume-se que as bases de dados em que desejamos realizar o *matching* são chamadas de A_1 e A_2 , com A_1 sendo a base de estudantes ou proprietários, enquanto A_2 será sempre o censo. Todo o algoritmo está baseado em funções de distância de *string*. Por essa razão, apresenta-se primeiro cada uma das funções utilizadas e, posteriormente, explica-se os detalhes operacionais e os critérios adotados.

Como visto no texto principal, uma *string* é qualquer sequência ordenada de

caracteres que representa uma combinação de letras, números ou símbolos.¹ Assim, cada nome de uma pessoa em um registro histórico é uma *string* que pode ser representada por uma sequência de *bytes* por um computador. Para considerar que dois registros correspondem à mesma pessoa, adota-se um conjunto de regras de decisão que têm como principal vetor de informação medidas de distância de *string*. Uma *medida de distância de string* é um índice que usa como *input* duas *strings* e tem como *output* um valor que pode ser crescente ou decrescente na distância das *strings*.

Uma medida de *string* bastante conhecida é a distância de Levenshtein que calcula o mínimo de operações necessárias para transformar uma *string* na outra. Formalmente, a distância de Levenshtein é definida como segue:

$$lev_{a,b}(i,j) = \begin{cases} \max(i,j) & \min(i,j) = 0, \\ \min \begin{cases} leva,b(i-1,j) + 1 \\ leva,b(i,j-1) + 1 \\ leva,b(i-1,j-1) + 1_{a_i \neq b_j} \end{cases} & \min(i,j) \neq 0 \end{cases} \quad (\text{A.1})$$

Onde a e b são as *strings* para as quais se pretende calcular a distância de Levenshtein e i e j são os comprimentos de cada uma das *strings*, respectivamente. Assim, se existem duas *strings*, a e b , para se calcular a distância de Levenshtein primeiro verifica-se se o comprimento de uma das *strings* é zero (i.e. $\min(i,j) = 0$), em caso afirmativo, a distância Levenshtein é simplesmente o valor da *string* de maior comprimento ($\max(i,j)$). Caso as duas *strings* tenham comprimento não nulo, computa-se, separadamente, o número de *remoções* necessárias em a (i.e. $lev_{a,b}(i-1,j)$) o número de *inserções* necessárias em a (i.e. $lev_{a,b}(i,j-1)$) e, em caso de a e b permanecerem diferentes, computa-se o número de substituições (i.e. $lev_{a,b}(i-1,j-1)$) para torná-las iguais. Após esses procedimentos soma-se o números de operações realizadas. O resultado dessa soma será a distância de Levenshtein entre a e b .

1. Mark Lutz, *Learning python: Powerful object-oriented programming* ("O'Reilly Media, Inc.", 2013).

Na prática, no algoritmo de *matching* utiliza-se a versão normalizada da distância de Levenshtein, dado por:

$$\frac{(|a| + |b|) - lev_{a,b}(i, j)}{|a| + |b|} \quad (A.2)$$

Além da distância de Levenshtein, o algoritmo de *matching* considera outras 3 medidas de distância de *string* como forma de dar robustez aos *matchings*. As outras 3 medidas utilizadas são *partial ratio*, *token sort ratio*, *token set ratio*. Todas essas medidas de distância de *string* são variações do índice de Levenshtein que buscam, através de operações suplementares, contornar limites associados à versão básica do índice.

A primeira medida, *partial ratio*, procura maximizar a probabilidade de dois registros serem idênticos baseados em *substrings*, o que ajuda a lidar com o problema de associado a *strings* com cumprimentos muito diferentes. A *token sort ratio*, simplesmente aplica a distância de Levenshtein normalizada (equação A.2) para as duas *strings* ordenadas alfabeticamente, sem distinção de letras maiúsculas e minúsculas. Esta medida minimiza o problema provocado por *strings* equivalentes, mas com *substrings* fora de ordem. Por fim, a *token set ratio* é calculada da mesma forma que a *token sort ratio*, mas além de ordenar a *string* antes de aplicar a equação A.2, nessa medida as *strings* são divididas em dois grupos. No primeiro grupo agrega-se os caracteres de interseção e no segundo os caracteres remanescentes. A *token set ratio* é crescente no percentual que a interseção representa do comprimento da *string* e no grau de semelhança dos conjuntos remanescentes. Esta última medida de string será especialmente importante no segundo estágio dos proprietários. Isso porque, para muitos casos, Estatística Agrícola apresenta o nome dos proprietários em ordem invertida, com o sobrenome vindo primeiro. Além disso, há muitas abreviações de sobrenome tanto na Estatística Agrícola como no censo. Como será visto adiante, o índice *token set ratio* nos ajuda a lidar com essas dificuldades.

Operacionalmente, a implementação do algoritmo consistiu na construção de um dicionário de dicionários em que as chaves do primeiro dicionário são os nomes na base de dados A_1 , enquanto as chaves do segundo dicionário são os nomes

das 4 medidas de distância de *string* utilizadas. Os valores do dicionário interno são dados por uma lista contendo 5 *tuples* binárias em que a primeira entrada é um *matching* em A_2 e a segunda entrada é o valor da medida de distância de *string*. Nesse contexto, um dicionário é uma forma particular de *symbol table* que é uma estrutura de dados caracterizada por pares associados de *chave-valor*.² Como qualquer *symbol table*, dicionários podem ser aninhados um número arbitrário de vezes, com possível repetição de valores, mas sem repetição de chaves dentro do mesmo dicionário.

Para dar maior concretude a essa descrição teórica, apresenta-se abaixo uma entrada do primeiro dicionário, correspondente ao aluno “Alberto Martins”:

```
1 {ratio: [(Alberto Martins, 100),
2   (Gilberto Martins, 90),
3   (Alvaro Martins, 83),
4   (Alberto Mattos, 83),
5   (Alberto Masi, 81)],
6  partial_ratio: [(Alberto Martins, 100),
7   (Gilberto Martins, 93),
8   (Pedro Martins, 85),
9   (Alberto Masi, 83),
10  (Rita Martins, 83)],
11 token_sort_ratio: [(Alberto Martins, 100),
12  (Gilberto Martins, 90),
13  (Maria Alberto, 86),
14  (Alvaro Martins, 83),
15  (Alberto Mattos, 83)],
16 token_set_ratio: [(Alberto Martins, 100),
17  (Gilberto Martins, 90),
18  (Maria Alberto, 86),
19  (Alvaro Martins, 83),
20  (Alberto Mattos, 83)]}
```

As chaves desse dicionário interno são os nomes das medidas de distância de *string*: *ratio*, *partial ratio*, *token sort ratio*, *token set ratio*. Para melhor compreensão do que significam os valores do dicionário interno, tome-se como exemplo

2. Robert Sedgewick e Kevin Wayne, *Computer science: An interdisciplinary approach* (Addison-Wesley Professional, 2016).

a chave *ratio*. O valor dessa chave são todos os caracteres dentro do símbolo de colchete, ou seja: [(Alberto Martins, 100), (Gilberto Martins, 90), (Alvaro Martins, 83), (Alberto Mattos, 83), (Alberto Masi, 81)]. A estrutura de dados cercada por colchetes é uma lista, enquanto a cercada por parenteses é uma *tuple*. Como dito acima, cada *tuple* tem um par de *matching* e valor da medida de distância de *string*. Nesse caso específico, o melhor *matching* é dado pelo nome “Alberto Martins” que corresponde a um *matching* perfeito (*ratio*=100) de “Alberto Martins”, nome na base de estudantes.

O primeiro estágio do algoritmo se encerra após a construção desse dicionário para todos os nomes da base A_1 .

A.2 Segundo estágio: *matching* dos estudantes

No caso dos estudantes, o segundo estágio inicia-se com a junção dos bancos A_1 e A_2 , utilizando o *matching* mais comum como variável chave ou o melhor *matching* pelo índice de Levenshtein, quando não houver mais comum. Após a junção dos bancos, cria-se uma classificação em que os *matchings* são separados em 3 grupos, de acordo com os parâmetros definidos na Tabela A.1.

Tabela A.1: Classificação dos *matchings*

Grupo	Definição	Decisão
Grupo 1: <i>matching</i> perfeitos	nomes idênticos, únicos e idade compatível	Aceita
Grupo 2: <i>matching</i> quase-perfeitos	nomes parecidos, únicos e idade compatível	Analisa caso a caso
Grupo 3: homônimo	melhor <i>matching</i> tem múltiplos casos no censo	Aceita após eliminar interseção

Nota: Só há eliminação de casos no Grupo 2. No total, foram validados 267 *matches*. A contribuição de cada grupo para o total foi a seguinte: Grupo 1: 62, Grupo 2: 184 e Grupo 3: 19.

A decisão é feita em dois estágios. No primeiro estágio os 3 primeiros grupos são classificados e aceita-se todos casos classificados no grupo 1. Os casos classificados nos grupos 2 e 3 são analisados caso a caso no segundo estágio onde

decide-se quais casos aceitar e quais casos rejeitar. Para aceitação ou rejeição dos *matchings* no grupo 2 foram avaliados pelo pesquisador cada par individualmente.

Reconhecidamente, ao relegar ao pesquisador a validação do melhor *matching* encontrado, introduzi-se um elemento de arbitrariedade ao algoritmo. Optou-se por essa abordagem, contudo, pois a definição de *score* de corte para a medida de distância de *strings* seria igualmente arbitrária e teria o custo adicional de aumentar o número de falsos positivos e falsos negativos. Um exemplo concreto disso pode ser dado pela análise de dois pares de nomes classificados no grupo 2 no primeiro estágio de implementação do algoritmo. O aluno "Roque Mestieri" da Escola Masculina de Ibaté tem como melhor *matching* no censo o indivíduo chamado "Roque Domingos Mestieri", a distância de Levenshtein entre esses dois nomes é de 9. Uma distância de Levenshtein menor é encontrada entre os nomes "Luiza Carrigelli", aluna da Escola Feminina de Santa Cruz, e "Luiz Farinelli" seu melhor *matching* no censo. É uma conjectura razoável supor que a probabilidade do primeiro par de informações se referirem a mesma pessoa é maior do que para a segunda pessoa. No entanto, o estabelecimento de uma linha de corte de, por exemplo, 5 pontos de distância Levenshtein nos levaria a aceitar um caso possivelmente falso e rejeitar um caso possivelmente verdadeiro.

Nota-se ainda dois fatos que devem ser considerados que têm o potencial de mitigar os efeitos do elemento de subjetividade do algoritmo para a validade da análise. O primeiro fato é o de que, independente da forma de seleção dos nomes dos grupo 2 e 3, sempre será possível testar a robustez dos resultados obtidos para diferentes índices e valores de distância de *strings*. Além disso, não é demais salientar que todos os nomes utilizados como *matchings* (aceitos ou não) pertencem ao seletivo grupo de *matchings* dos nomes com as cinco menores distâncias de *strings* para cada um dos 4 índices utilizados. Em que pese o fato de que o universo de comparação para cada nome era o total de pessoas listadas no censo com idade igual ou inferior a 20 anos, ou seja, mais de 22 mil nomes.

No grupo 3, a decisão de aceite ou rejeição foi baseada no número de entradas idênticas nas variáveis do censo. Por exemplo, se o nome de um aluno é Luiz Augusto e existem dois indivíduos registrados como Luiz Augusto no censo em idade escolar, então utiliza-se no banco final apenas as variáveis registradas

igualmente no censo para esses dois indivíduos. Assim, se os “Luiz Augustos” do censo são ambos registrados como proprietários, adiciona-se essa informação no banco final. Contudo, se um Luiz Augusto está registrado no censo como “Preto” e outro como “Branco” a informação de cor da pele entra como *missing* no banco final.

Como resultado da implementação do primeiro estágio e do segundo estágio, chegou-se a 62 indivíduos no Grupo 1, 310 no grupo 2 e 19 no grupo 3. A base final, após a eliminação de 124 *matches* rejeitados no grupo 2, ficou com 267 indivíduos.

A.3 Segundo estágio: *matching* dos proprietários

No caso dos proprietários, o primeiro estágio também consistiu na construção de dicionários de dicionários com a chave do primeiro dicionário sendo o nome do indivíduo na Estatística Agrícola e as chaves internas sendo compostas de índices de distância de *string*.

É ilustrativo apresentar o *matching* dos proprietários utilizando a nomenclatura do algoritmo ABE, já amplamente utilizado na literatura ³. A versão básica do ABE está dividida em três estágios, quais sejam, *i) cleaning ii) blocking* e *iii) merging*.⁴ No presente caso, a limpeza (*cleaning*) dos nomes foi feita em dois passos. Primeiro os pronomes de tratamento e títulos foram removidos (Dona, Coronel, Dr., etc) e eliminou-se abreviações quando possível (e.g., Alm.da passou a ser Almeida), além de eliminar pessoas jurídicas como Câmara de São Carlos e bancos. No segundo sub-estágio apenas efetuou-se a padronização das *strings*.

A etapa intitulada *blocking* consiste em restringir o conjunto de possíveis *matches*. Em outras palavras, nesta etapa elimina-se indivíduos que não podem formar um *matching* em função de características observadas. No presente caso, o *blocking* consistiu em procurar na Estatística Agrícola apenas por nomes de indivíduos do censo municipal maiores de 18 anos marcados como proprietários. Por fim, o

3. ABE é acrônimo para o nome dos autores Ran Abramitzky, Leah Boustan, Katherine Eriksson: Ran Abramitzky et al., *Automated linking of historical data*, relatório técnico (National Bureau of Economic Research, 2019)

4. Ibid., 8.

último estágio para união das informações das duas bases consiste em escolher uma medida de distância de *string* e determinar um limite a partir do qual o *matching* é aceitável. Em algoritmos do tipo ABE a medida de distância de *string* comumente utilizada é a distância de Jaro–Winkler e, uma vez definido o *cut-off* neste índice, a junção das bases é efetuada. Em função das especificidades do presente caso, o critério de aceitação será diferente.

Após a limpeza, bloqueio de nomes não elegíveis para o *matching* e construção do dicionário de dicionários de *matchings*, utilizou-se este último para construir uma base de dados. Esta base contém os 5 melhores *matchings* para cada proprietário usando a distância de Levenshtein padronizada e a *token set ratio*. Então, definiu-se a seguinte regra de pré-aceitação: o nome do censo será considerado um *matching* se o índice de Levenshtein for maior que 90 ou se um dos 5 melhores *matchings* pela *token set ratio* for selecionado pelo pesquisador na análise caso a caso. Chamamos esse passo de pré-aceitação porque a aceitação completa demanda a aplicação de outros critérios que serão detalhados adiante. A análise caso a caso aqui, em que pese o custo associado de impor uma arbitrariedade ao algoritmo, foi implementada pelas mesmas razões arroladas na seção anterior.

A seguir apresenta-se um exemplo de análise de caso que levou à seleção de um dos *matches* listados na chave de *token set ratio*. A Tabela A.2 mostra os cinco melhores *matchings* no censo do proprietário Manoel Bangoiro dos Santos listado na Estatística Agrícola. Os 5 melhores *matchings* são mostrados tanto para o índice básico de Levenshtein como para a *token set ratio*. Como o melhor *matching* de Levenshtein é inferior ou igual a 90, procurou-se entre os 5 melhores *matchings* da *token set ratio* e selecionamos de Manoel Bangoiro dos Santos como sendo Manoel dos Santos Barnagoiro no censo. Esse procedimento foi implementado para todos os casos que o índice de Levenshtein foi menor ou igual que 90. Casos em que nenhuma das opções de *matchings* da *token set ratio* era similar ao nome de interesse foram eliminados.

Além das medidas de *string* utilizam-se mais dois critérios para validar os *matchings* encontrados. O primeiro critério de validação consistiu no cruzamento das informações da variável de nacionalidade, disponível nas duas bases. Em muitos estudos que utilizam algoritmos do tipo ABE, variáveis comuns às ba-

Tabela A.2: Seleção do matching do proprietário “Manoel Bangoiro dos Santos” listado na Estatística Agrícola, São Carlos 1905/1907

5 melhores matchings (base levenshtein)		5 melhores matchings (token set ratio)	
Manoel dos Santos	79	Manoel dos Santos	100
Antonio dos Santos	77	Manoel dos Santos Barnagoiro	96
Manoel Antonio de Mattos	76	Manoel dos Santos Mattos	83
Jose A. dos Santos	70	João dos Santos	80
Daniel dos Santos	70	Manoel dos Santos Ferreira	79

ses que se pretende juntar são utilizadas na etapa de *blocking*. No presente caso, optou-se por uma abordagem mais flexível ao utilizar a variável de nacionalidade apenas como validação. Essa decisão foi tomada em função tanto da inconsistências das variáveis de nacionalidade nas duas bases, quanto em função da imprecisão de alguns registros de nacionalidades da Estatística Agrícola. Em relação ao primeiro ponto, basta notar que algumas nacionalidades estão presentes na Estatística Agrícola, mas não estão presentes no censo, e vice-versa. Exemplos incluem as nacionalidades “russo” e “turco”. A primeira só existe na Estatística Agrícola, enquanto a segunda só existe no censo. Um problema mais grave que a ausência de correspondência exata de nacionalidades de grupos minoritários é a imprecisão de alguns registros na Estatística Agrícola em relação a grupos nacionais majoritários. Um exemplo são tirolezes meridionais, falantes de italiano de origem austríaca que podem ter sido registrados como austríacos na Estatística Agrícola e italianos no censo. Parece ter sido esse o caso de Giacomo Brumati, registrado como austríaco na Estatística Agrícola e como italiano no censo.

Os problemas relatados acima com a variável de nacionalidade levou à adoção do seguinte critério: foram eliminados todos os casos em que a nacionalidade dos *matchings* eram diferentes e que, cumulativamente, tinham distância de Levenshtein maior que 4. Em outras palavras, se o nome dos *matchings* é virtualmente idêntico, a informação de nacionalidade não foi considerada.

O último critério de validação dos *matchings* foi a eliminação de casos em que o melhor *matching* no censo tinha múltiplos casos. Infelizmente, alguns nomes na

Estatística Agrícola eram demasiadamente comuns para que se pudesse determinar unicamente o registro correspondente no censo. A decisão de eliminar esses casos pode ter levado a uma subestimação do número de brasileiros proprietários de terra. Por fim, considera-se “branco” todos os indivíduos estrangeiros, incluindo os que não foram encontrados no censo. Para esses casos, com exceção da cor e sexo, as informações demográficas entraram no banco final como *missing*.

A Estatística Agrícola de São Carlos lista 311 propriedades, com a implementação do algoritmo apresentado neste Apêndice foi possível identificar no censo municipal de 1907 os proprietários de 192 dessas propriedades. Como alguns proprietários têm mais de uma propriedade, o número de proprietários identificados é necessariamente menor. Com efeito, nossa base final analisada no âmbito do Capítulo 4 contou com 176 proprietários.

Apêndice B

Índice de Equilíbrio Racial

Neste anexo apresenta-se os detalhes técnicos para o cômputo do índice de Equidade Racial, usado para o estudo das ocupações no Capítulo 3 e em outras partes do texto. Originalmente, o *IER* foi proposto no estudo de Ransom e Sucht para economia do Sul dos Estados Unidos no *post bellum*.¹ No estudo citado, o *IER* é chamado de *Racial Balance Index*, sendo “Índice de Equilíbrio Racial” apenas uma tradução livre. É importante notar que as instruções básicas para o cálculo do *IER* já se encontram na obra de Ransom e Sucht.² No entanto, a apresentação original do cômputo do índice pula algumas etapas e contém um erro, o que pode confundir o leitor. Neste anexo apresenta-se a rotina detalhada do cômputo do índice, bem como se corrige a imprecisão na explicação original. Para o cômputo do *IER*, assume-se que a probabilidade de x trabalhadores negros estarem ocupados em uma ocupação j , com n trabalhadores segue uma distribuição binomial:

$$P_j(X = x_j) = \binom{n_j}{x_j} p^{x_j} (1 - p)^{n_j - x_j} \quad (\text{B.1})$$

Em que p é uma constante que representa a proporção de trabalhadores negros na força de trabalho ocupada.

Usando propriedades da distribuição binomial é possível definir o nível de equilíbrio racial de uma determinada ocupação pela diferença da quantidade de

1. Ransom e Sutch, *One kind*.

2. *Ibid.*, 220-231.

negros observados na ocupação (x_j) e seu valor esperado ($n_j p$). Medindo essa distância no número de desvios-padrões temos:

$$v_j = \frac{x - n_j p}{\sqrt{n_j p (1 - p)}} \quad (\text{B.2})$$

O *IER* é apenas uma versão padronizada do índice v e, assim, tem seu valor variando de -1 a 1. Valores próximos a -1 indicam sub-representação de negros, enquanto valores próximos de 1 indicam sobre-representação desse grupo demográfico na ocupação j .

Para chegar a essa versão padronizada, contudo, são necessários mais alguns passos. Substituindo bn por x , temos:

$$\frac{n(b - p)}{\sqrt{np(1 - p)}} \quad (\text{B.3})$$

Elevando ao quadrado e aplicando a raiz quadrada:

$$\sqrt{\frac{n^2 (b - p)^2}{np(1 - p)}} \quad (\text{B.4})$$

$$\sqrt{\frac{n (b - p)^2}{p(1 - p)}} \quad (\text{B.5})$$

Como b é uma proporção, seu valor necessariamente varia de 0 a 1. Assim, é possível definir dois limites para v como função de b . Se $b = 0$, vale:

$$\sqrt{\frac{np^2}{p(1 - p)}} \quad (\text{B.6})$$

$$\sqrt{\frac{np}{(1 - p)}} \quad (\text{B.7})$$

Se $b = 1$:

$$\sqrt{\frac{n(1 - p)^2}{p(1 - p)}} \quad (\text{B.8})$$

$$\sqrt{\frac{n(1-p)}{p}} \quad (\text{B.9})$$

Então, uma versão padronizada do índice pode ser construída dividindo-se v por:

$$\frac{\sqrt{np^{1-b}(1-p)^b}}{\sqrt{p^b(1-p)^{1-b}}} \quad (\text{B.10})$$

Realizando a divisão:

$$\psi = \frac{\frac{n(b-p)}{\sqrt{np(1-p)}}}{\frac{\sqrt{np^{1-b}(1-p)^b}}{\sqrt{p^b(1-p)^{1-b}}}} \quad (\text{B.11})$$

$$\psi = \frac{n(b-p)}{\sqrt{np(1-p)}} \cdot \frac{\sqrt{p^b(1-p)^{1-b}}}{\sqrt{np^{1-b}(1-p)^b}} \quad (\text{B.12})$$

$$\psi = \frac{n(b-p)}{\sqrt{np(1-p)}} \cdot \frac{1}{\sqrt{np(1-p)}} \cdot \sqrt{\frac{p^b(1-p)^{1-b}}{p^{-b}(1-p)^{b-1}}} \quad (\text{B.13})$$

$$\psi = \frac{n(b-p)}{np(1-p)} \cdot \frac{p^{2b}}{(1-p)^{2(b+1)}} \quad (\text{B.14})$$

$$\psi \frac{b-p}{(1-p)^b p^{1-b}} \quad (\text{B.15})$$

O *IER* é simplesmente um rearranjo da expressão acima:

$$IER = \frac{b-p}{p \left(\frac{1-p}{p}\right)^b} \quad (\text{B.16})$$

Apêndice C

Tabelas Complementares

Tabela C.1: Profissão dos Pais negros dos estudantes e da população masculina negra, São Carlos, 1907
(em porcentagem da respectiva população)

População negra masculina n=1924		Pais negros n=24	
Colono	39,6	Colono	37,5
Camarada	22,4	Camarada	16,7
n.d	8,6	Lavrador	12,5
Lavrador	5,9	Carpinteiro	8,3
Empregado	4,4	Empregado	8,3
Carroceiro	1,6	Campeiro	4,2
Pedreiro	1,4	Emp Farmácia	4,2
Criado	1,0	Empreiteiro	4,2
Empreiteiro	0,6	Pintor	4,2
Cozinheiro	0,6	-	-

Fontes: Livros de chamadas - Arquivo do Estado de São Paulo.
São Carlos, Recenseamento de 1907.

Notas: Ver nota da Tabela 2.5.

Tabela C.2: Profissão dos Mães negras dos estudantes e da população feminina negra, São Carlos, 1907
(em porcentagem da respectiva população)

População negra feminina n=1778		Mães negras n=24	
Colono	33,7	Colono	33,3
n.d	16,1	n,d	20,8
Camarada	9,1	Camarada	12,5
Cozinheiro	7,0	Lavrador	12,5
Lavrador	5,2	Lavadeira	8,3
Criado	4,1	Campeiro	4,2
Empregado	3,6	Empregado	4,2
Lavadeira	3,0	Empreiteiro	4,2
Costureiro	0,6	-	-
Empreiteiro	0,4	-	-

Fontes: Livros de chamadas - Arquivo do Estado de São Paulo.
São Carlos, Recenseamento de 1907.

Notas: Ver nota da Tabela 2.5.

Tabela C.3: Distribuição etária entre estudantes, São Carlos, 1906-1909

Faixa etária	Total	(%)
(0, 5]	45	16,9
(5, 10]	146	54,7
(10, 15]	62	23,2
(15, 20]	14	5,2

Fontes: Livros de chamadas - Arquivo do Estado de São Paulo.
São Carlos, Recenseamento de 1907.

Notas: Idade em Fevereiro de 1907 e, portanto, não corresponde necessariamente à idade da criança na escola. Por exemplo, uma criança de 7 anos que entrou na base como resultado de seu nome estar em uma lista de chamada em uma escola em 1909, teria no censo 5 ou 4 anos dependendo da data de aniversário.

Tabela C.4: Resultados - adicionando variável de alfabetização do pai

	<i>Variável dependente:</i>		
	Matrícula		
	<i>MPL</i> (1)	<i>logit</i> (2)	<i>probit</i> (3)
Negro	-0,084*** (0,013)	-0,668*** (0,120)	-0,363*** (0,067)
Pai Colono	-0,236*** (0,011)	-2,188*** (0,089)	-1,165*** (0,047)
Mãe Brasileira	0,011 (0,011)	-0,232*** (0,081)	-0,098** (0,047)
Feminino	-0,020*** (0,007)	-0,176*** (0,065)	-0,120*** (0,036)
Mãe alfabetizada	0,167*** (0,012)	0,999*** (0,078)	0,581*** (0,044)
Pai alfabetizado	0,120*** (0,009)	1,079*** (0,074)	0,581*** (0,041)
Pai proprietário	0,130*** (0,015)	0,624*** (0,074)	0,372*** (0,044)
N. de membros	-0,005*** (0,002)	-0,054*** (0,013)	-0,029*** (0,008)
Constante	0,274*** (0,016)	-1,058*** (0,112)	-0,623*** (0,064)
Observações	8,827	8,827	8,827
R ²	0,289		
R ² Ajustado	0,289		
log-verossimilhança		-3,046,580	-3,055,019
Crit. Inf. Akaike		6,111,160	6,128,037
Erro padrão do resíduo	0,336 (df = 8818)		
F (Wald)	353.52*** (df = 8; 8818)		

*p<0,1; **p<0,05; ***p<0,01

Fontes: Livros de chamadas - Arquivo do Estado de São Paulo.
São Carlos, Recenseamento de 1907.

Nota: Erros robustos para heteroscedasticidade no MPL.

Tabela C.5: Resultados - usando *dummies* para preto e mulato

	<i>Variável dependente:</i>		
	Matrícula		
	MQO (1)	<i>logit</i> (2)	<i>probit</i> (3)
Preto	-0,105*** (0,014)	-0,861*** (0,131)	-0,465*** (0,073)
Mulato	-0,118*** (0,024)	-0,771*** (0,215)	-0,447*** (0,121)
Pai Colono	-0,250*** (0,011)	-2,241*** (0,087)	-1,197*** (0,046)
Mãe Brasileira	0,012 (0,011)	-0,181** (0,079)	-0,072 (0,046)
Feminino	-0,020*** (0,007)	-0,168*** (0,064)	-0,117*** (0,036)
Mãe alfabetizada	0,224*** (0,012)	1,446*** (0,072)	0,829*** (0,041)
Pai proprietário	0,134*** (0,015)	0,630*** (0,073)	0,382*** (0,044)
N. de membros	-0,004*** (0,002)	-0,044*** (0,013)	-0,023*** (0,007)
Constante	0,316*** (0,016)	-0,650*** (0,105)	-0,413*** (0,061)
Observações	8,827	8,827	8,827
R ²	0,272		
R ² Ajustado	0,272		
log-verossimilhança		-3,156,329	-3,156,690
Crit. Inf. Akaike		6,330,659	6,331,381
Erro padrão do resíduo	0,340 (df = 8818)		
F (Wald)	323,68*** (df = 8; 8818)		

*p<0,1; **p<0,05; ***p<0,01

Fontes: Livros de chamadas - Arquivo do Estado de São Paulo.
São Carlos, Recenseamento de 1907.

Nota: Erros robustos para heteroscedasticidade no MPL.

Tabela C.6: IER das ocupações dos grupos 1 a 3, População mulata masculina, São Carlos, 1907

Grupos	Ocupação	Total	Proporção de mulatos	IER
Branco excluídos				
	Camarada	1156	0,06	1,00
	Carpinteiro	89	0,08	1,00
	Maquinista	50	0,08	1,00
	Operario	46	0,07	1,00
	Seleiro	28	0,07	1,00
	Cozinheiro	21	0,14	1,00
	Criado	13	0,08	1,00
	Portador	11	0,27	1,00
	Mecânico	9	0,11	1,00
	Agregado	6	0,17	1,00
	Caseiro	5	0,20	1,00
	Pintor	19	0,05	0,91
Dominância negra				
	Feitor	21	0,05	0,76
	Cocheiro	49	0,04	0,54
	Empregado	345	0,04	0,54
	Soldado	25	0,04	0,52
	Empreiteiro	84	0,04	0,38
	Pedreiro	150	0,03	0,29
Ocupações Abertas				
	Empregado Cia Paulista	135	0,03	0,17
	Lavrador	988	0,02	-0,03
	Alfaiate	44	0,02	-0,08
	Ferreiro	44	0,02	-0,08

Fonte: São Carlos, Recenseamento de 1907.

Notas: Ver notas das Tabelas 3.1 e 3.2.

Tabela C.7: IER das ocupações dos grupos 1 a 3, População mulata feminina, São Carlos, 1907

	Ocupação	Total	Proporção de mulatas	IER
Branco excluídos				
	Camarada	338	0,08	1,00
	Cozinheiro	175	0,08	1,00
	Empregado	165	0,09	1,00
	Lavador	119	0,05	1,00
	Pedreiro	11	0,09	1,00
	Operario	10	0,10	1,00
	Ferreiro	8	0,12	1,00
	Empreiteiro	65	0,05	0,86
Ocupações Abertas				
	Lavrador	812	0,02	-0,03

Fonte: São Carlos, Recenseamento de 1907.

Notas: Ver notas das Tabelas 3.1 e 3.2.

Tabela C.8: IER das ocupações dos grupos 1 a 3, População preta masculina, São Carlos, 1907

	Ocupação	Total	Proporção de pretos	IER
Branco excluídos				
	Cozinheiro	21	0,33	1,00
	Criado	13	0,46	1,00
	Doceiro	5	0,40	1,00
	Camarada	1156	0,29	0,89
	Carroceiro	158	0,27	0,81
Dominância negra				
	Caseiro	5	0,20	0,54
	Empregado	345	0,18	0,45
	Agregado	6	0,17	0,36
	Cocheiro	49	0,16	0,34
	Soldado	25	0,16	0,32
	Pintor	19	0,16	0,31
	Pedreiro	150	0,15	0,28
	Maquinista	50	0,14	0,20
Ocupações abertas				
	Carpinteiro	89	0,11	0,02
	Serrador	9	0,11	0,01
	Trabalhador	60	0,10	-0,07
	Empreiteiro	84	0,10	-0,11
	Portador	11	0,09	-0,14
	Colono	4950	0,09	-0,17

Fonte: São Carlos, Recenseamento de 1907.

Notas: Ver notas das Tabelas 3.1 e 3.2.

Tabela C.9: IER das ocupações dos grupos 1 a 3, População preta feminina, São Carlos, 1907

Ocupação	Total	Proporção de pretas	IER
Branco excluídos			
Cozinheiro	175	0,60	1,00
Lavador	119	0,39	1,00
Criado	73	0,34	1,00
Doceiro	8	0,50	1,00
Seleiro	5	0,40	1,00
Camarada	338	0,28	0,86
Engomador	18	0,28	0,86
Empregado	165	0,27	0,84
Ocupações abertas			
Empreiteiro	65	0,12	0,10
Cocheiro	9	0,11	0,01
Pedreiro	11	0,09	-0,14

Fonte: São Carlos, Recenseamento de 1907.

Notas: Ver notas das Tabelas 3.1 e 3.2.

Tabela C.10: Indicadores relativos à propriedade rural de Brasileiros brancos, Imigrantes e Negros (sem outlier), São Carlos, 1905/1907

Grupo	Café (área)	Nacionais	Estrangeiros	Terra roxa	Valor (mil-réis)
Brasileiros brancos	39,8	7,3	26,4	29,7	108.199,2
Imigrantes	27,9	1,7	5,9	8,5	19.442,6
Negros	25,1	1,5	10,8	16,6	65750,0

Fontes: São Carlos, Recenseamento de 1907.

São Paulo, Estatística agrícola e zootécnica do estado de São Paulo no ano agrícola de 1904–1905.

Notas: A coluna “Café” representam a média do percentual da área total dedicada ao cultivo de café. As colunas “Nacionais” e “Estrangeiros” indicam o número médio de trabalhadores brasileiros e estrangeiros que trabalhavam na propriedade. A coluna “Terra roxa” indica o percentual de propriedade rurais do grupo que tinha essa qualidade de terra.

Tabela C.11: Estatísticas Descritivas da área total de propriedades rurais de Brasileiros brancos, Imigrantes e Negros (sem *outlier*), São Carlos, 1905/1907

	Brasileiros Brancos	Imigrantes	Negros
Número de Propriedades	123	61	6
Média	157,6	44,5	104,5
Desvio-padrão	281,3	128,4	94,6
Mínimo	3	0,5	1
25%	30,8	6,5	26,8
50%	68,5	20	100
75%	149,5	37,5	165,8
Máximo	1808	1000	236
Gini	0.64	0.68	0.46
Latifundiários	9	1	0

Fontes: São Carlos, Recenseamento de 1907.

São Paulo, Estatística agrícola e zootécnica do estado de São Paulo no ano agrícola de 1904–1905.

Notas: Área total em alqueires (1 alqueire paulista = 2,42 hectares ou 24.200 metros quadrados).

Apêndice D

Figuras Complementares

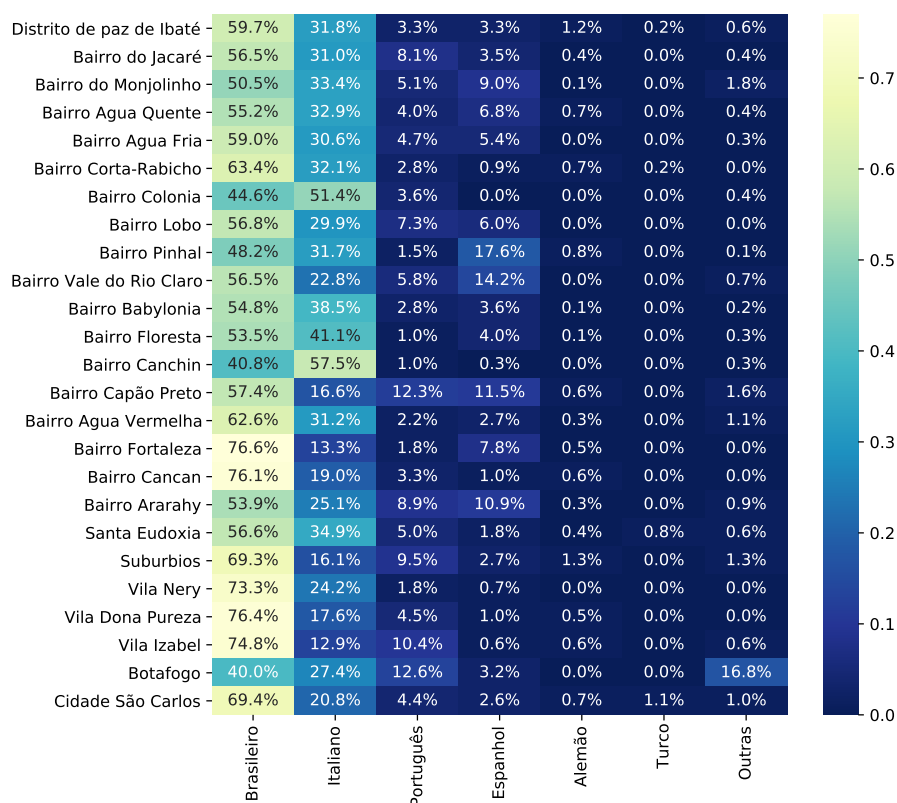


Figura D.1: Distribuição de nacionalidade por bairro, São Carlos, 1907

Fonte: São Carlos, Recenseamento de 1907.