

APPENDIX TO

**Black Americans' Landholdings and Economic Mobility after Emancipation:  
Evidence from the Census of Agriculture and Linked Records**

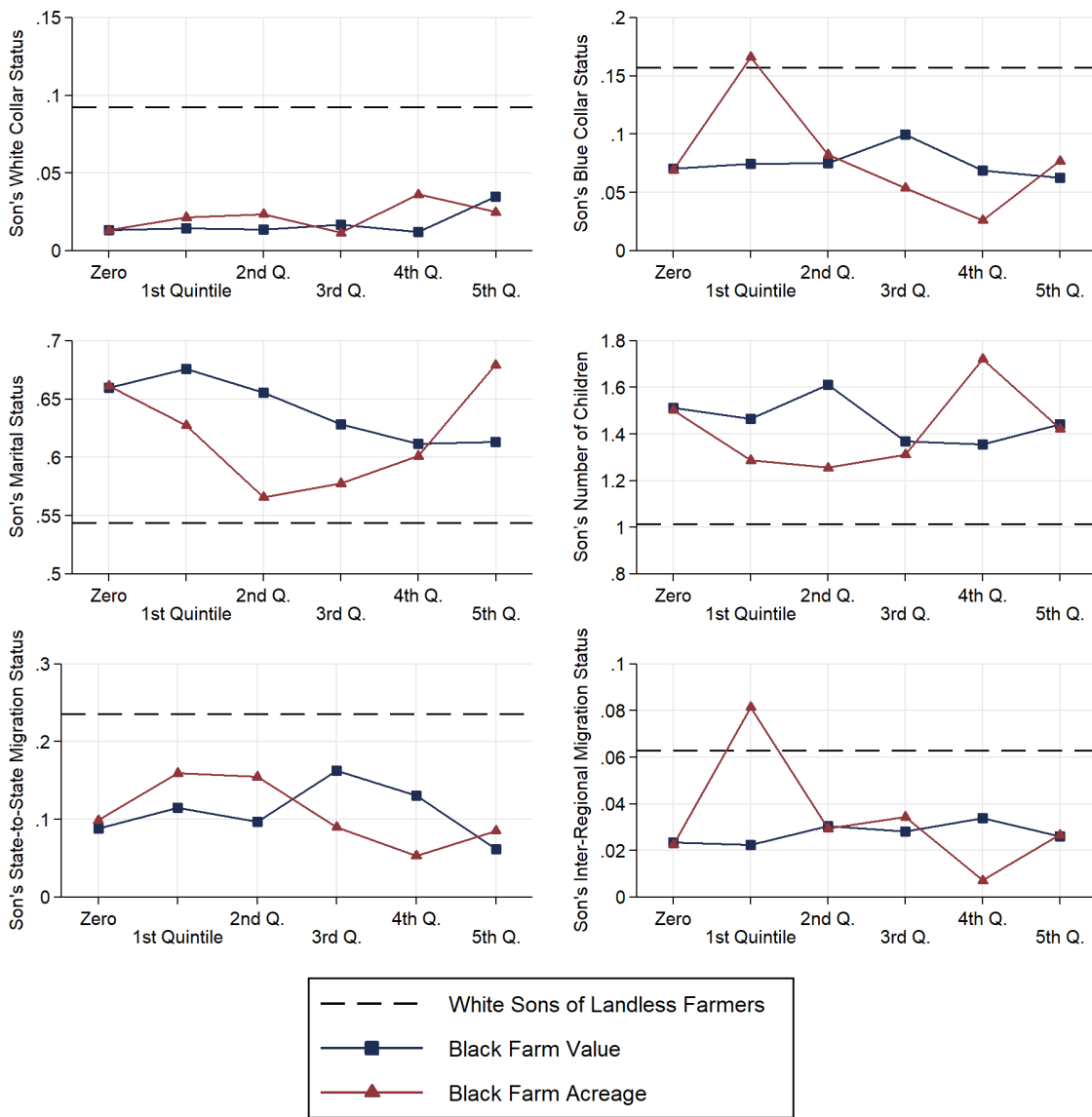
APPENDIX I. APPENDIX TABLES AND FIGURES

APPENDIX II. DATA APPENDIX

- a. LINKED SAMPLE
- b. IMPUTING MISSING OCCUPATIONS FOR SONS OBSERVED IN 1900
- c. 1880 AGRICULTURE CENSUS LINK
- d. ASSIGNMENT OF INCOME SCORES AND INCOME SCORE RANKS

APPENDIX I. TABLES AND FIGURES

**FIGURE A1 – SONS’ 1900 ADDITIONAL OUTCOMES BY FATHERS’ 1880 LANDOWNING STATUS**

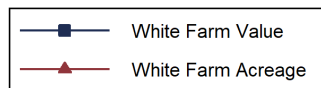
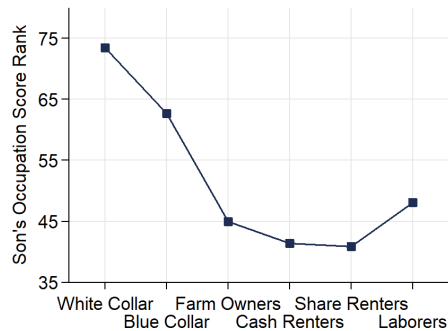
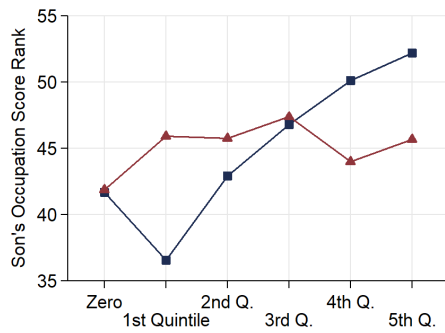
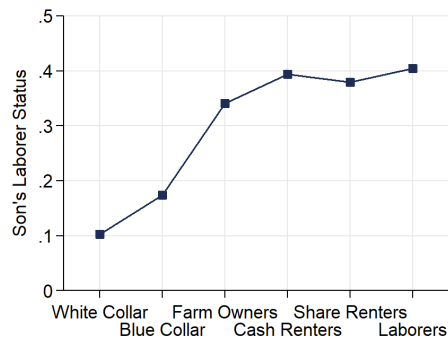
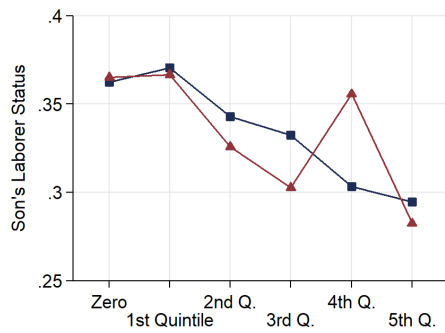
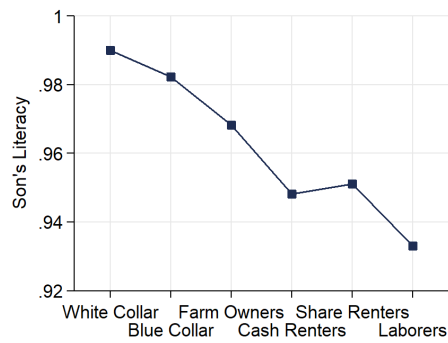
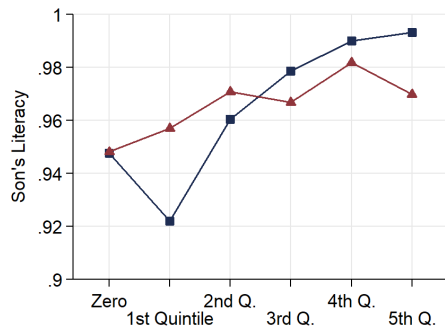
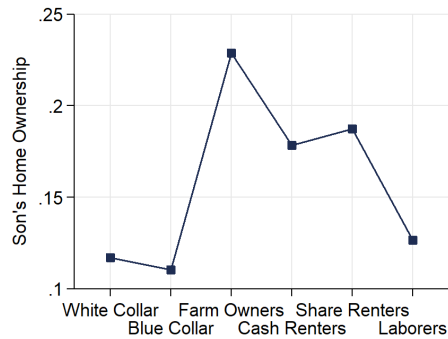
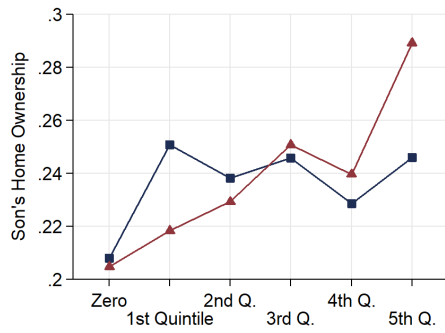


Notes: See notes to Figure 2 in the main text.

**FIGURE A2 – SONS’ 1900 OUTCOMES BY FATHERS’ 1880  
LANDOWNING OR OCCUPATION STATUS – WHITE SONS ONLY**

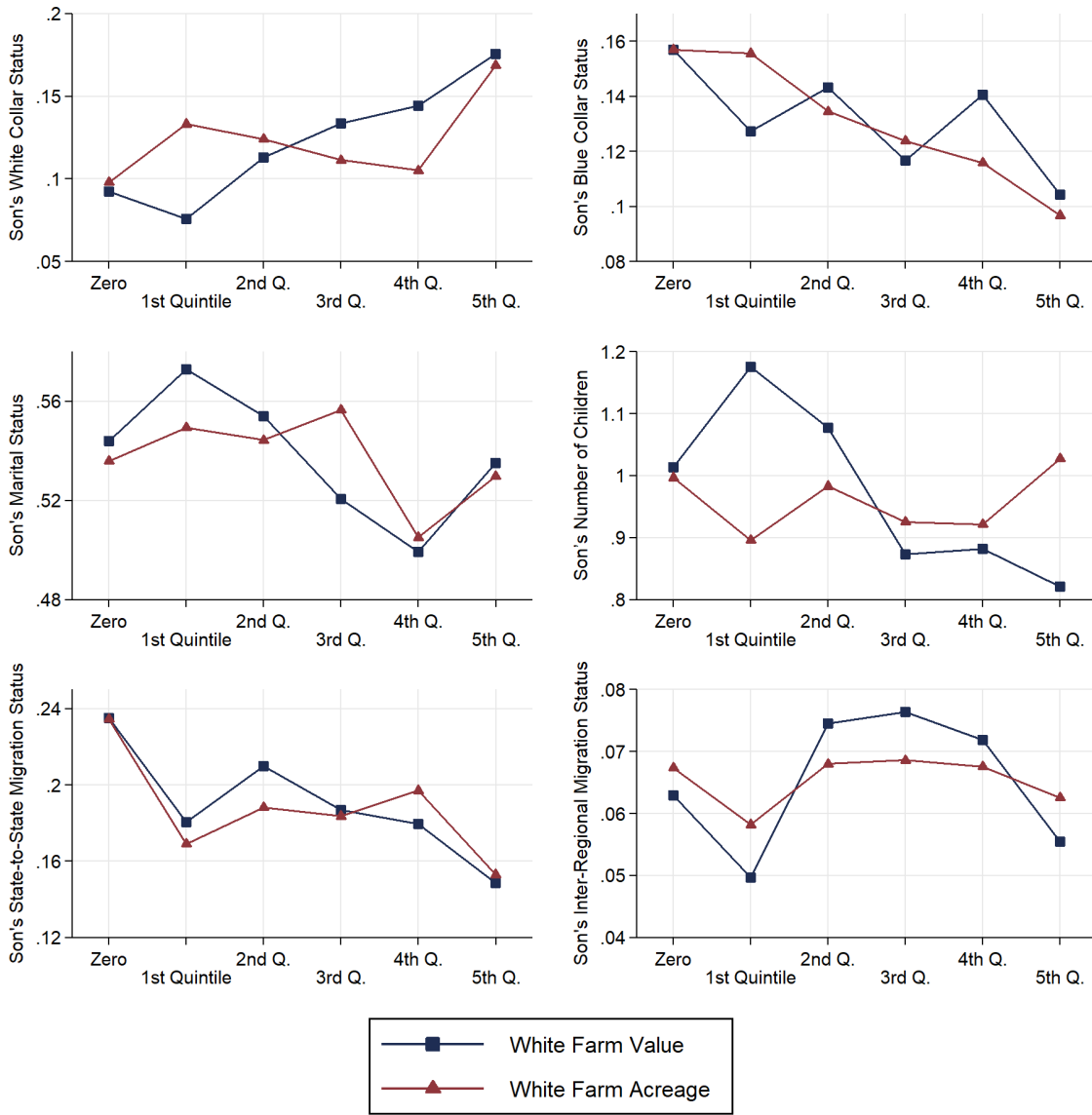
**QUINTILES OF FARM OWNERSHIP**

**OCCUPATION CATEGORIES**



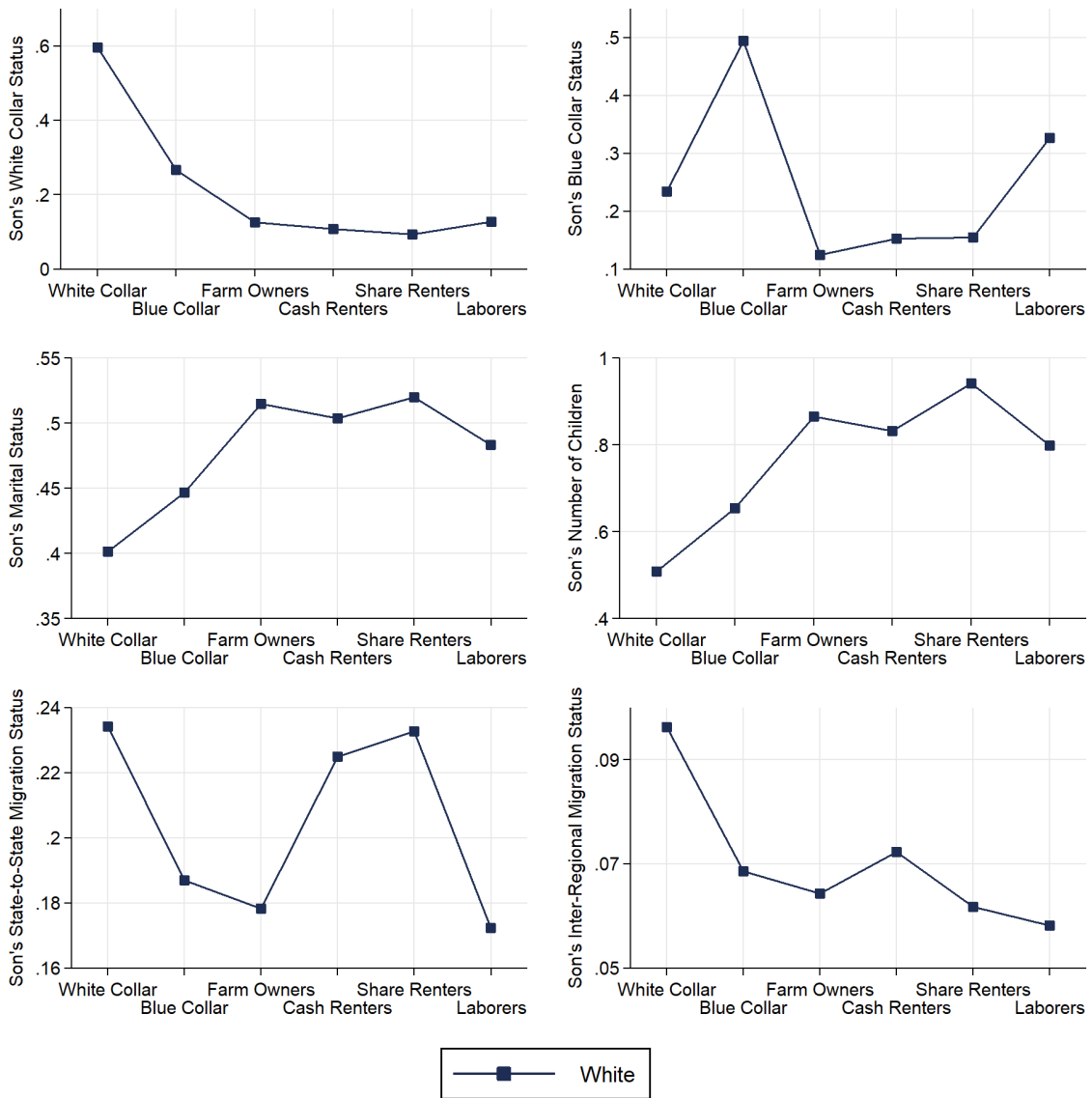
Notes: See notes to Figure 2 in the main text.

**FIGURE A3 – SONS’ 1900 ADDITIONAL OUTCOMES  
BY FATHERS’ 1880 LANDOWNING STATUS – WHITE SONS ONLY**



Notes: See notes to Figure 2 in the main text.

**FIGURE A4 – SONS’ 1900 ADDITIONAL OUTCOMES BY FATHERS’ 1880 OCCUPATION STATUS – WHITES ONLY**



Notes: See notes to Figure 2 in the main text.

**TABLE A1: RETURN TO QUINTILES OF FATHERS' FARM VALUE  
(RELATIVE TO NON-OWNERS)**

	BLACK		WHITE	
	U.S.	South Only	U.S.	South Only
<b>SONS' OCCUPATION INCOME SCORE RANK</b>				
1st Quintile	-0.1853 (0.4720)	-0.1028 (0.4705)	-2.6167** (1.2338)	-2.5062 (1.7425)
2nd Quintile	0.1055 (0.5893)	0.1439 (0.5903)	0.5407 (1.2551)	0.8710 (1.9192)
3rd Quintile	-0.0834 (0.4773)	-0.1259 (0.4771)	2.9432** (1.2591)	3.7294* (2.2499)
4th Quintile	0.2496 (0.4940)	0.3807 (0.4940)	4.8980*** (1.2649)	7.3141** (3.0835)
5th Quintile	0.6576 (0.5520)	0.4820 (0.5604)	5.8146*** (1.2915)	8.0435** (3.7938)
<b>SONS' HOME OWNERSHIP (BINARY)</b>				
1st Quintile	0.0201 (0.0180)	0.0206 (0.0180)	0.0225 (0.0194)	0.0410 (0.0280)
2nd Quintile	0.0379* (0.0207)	0.0385* (0.0208)	0.0143 (0.0199)	0.0310 (0.0308)
3rd Quintile	0.0290 (0.0201)	0.0288 (0.0203)	0.0327 (0.0203)	0.0502 (0.0373)
4th Quintile	0.0788*** (0.0232)	0.0836*** (0.0236)	0.0257 (0.0201)	0.0165 (0.0444)
5th Quintile	0.0882*** (0.0241)	0.0849*** (0.0248)	0.0537** (0.0209)	0.0352 (0.0531)
<b>SONS' LITERACY (BINARY)</b>				
1st Quintile	-0.0194 (0.0284)	-0.0195 (0.0285)	-0.0113 (0.0125)	0.0018 (0.0221)
2nd Quintile	0.0512 (0.0317)	0.0516 (0.0318)	0.0148 (0.0109)	0.0459** (0.0213)
3rd Quintile	0.0710** (0.0297)	0.0732** (0.0300)	0.0256*** (0.0098)	0.0744*** (0.0211)
4th Quintile	0.0948*** (0.0280)	0.0963*** (0.0284)	0.0316*** (0.0088)	0.0853*** (0.0221)
5th Quintile	0.0770*** (0.0295)	0.0780** (0.0305)	0.0314*** (0.0086)	0.0769*** (0.0237)
<b>SONS' LABORER STATUS (BINARY)</b>				
1st Quintile	0.0593** (0.0290)	0.0573** (0.0291)	0.0234 (0.0234)	0.0353 (0.0322)
2nd Quintile	-0.0150 (0.0306)	-0.0154 (0.0307)	-0.0139 (0.0230)	0.0075 (0.0338)
3rd Quintile	-0.0548* (0.0299)	-0.0550* (0.0302)	-0.0423* (0.0230)	-0.0206 (0.0373)
4th Quintile	0.0074 (0.0294)	0.0034 (0.0298)	-0.0888*** (0.0225)	-0.0171 (0.0461)
5th Quintile	-0.0677** (0.0302)	-0.0652** (0.0310)	-0.0982*** (0.0230)	-0.0464 (0.0582)
<b>N</b>	<b>3,182</b>	<b>3,109</b>	<b>4,934</b>	<b>1,639</b>

*Notes:* Table contains regression coefficient for quintile dummies, representing the quintile of fathers' 1880 farm value, and are interpreted relative to a farm value of \$0. Sample includes only farmers; renters for share are assigned a farm value of \$0. Renters for cash are assigned a farm value equal to the value of their livestock. Farm owners are assigned the combined value of the property and livestock. Regressions also include fixed effects for sons' age in 1900, fathers' age in 1880, fathers' urban status, and fathers' 1880 state fixed effects.

**TABLE A2: RETURN TO QUINTILES OF FATHERS' FARM VALUE  
(RELATIVE TO NON-OWNERS), OTHER OUTCOMES**

	BLACK		WHITE	
	U.S.	South Only	U.S.	South Only
<b>SONS' STATE-TO-STATE MIGRATION</b>				
1st Quintile	0.0157 (0.0185)	0.0155 (0.0186)	-0.0343* (0.0204)	-0.0684** (0.0274)
2nd Quintile	-0.0097 (0.0188)	-0.0099 (0.0189)	-0.0207 (0.0205)	-0.0567* (0.0292)
3rd Quintile	0.0516** (0.0225)	0.0481** (0.0226)	-0.0494** (0.0204)	-0.0564* (0.0330)
4th Quintile	0.0381* (0.0202)	0.0331 (0.0203)	-0.0490** (0.0208)	0.0133 (0.0435)
5th Quintile	-0.0184 (0.0162)	-0.0175 (0.0162)	-0.0803*** (0.0205)	-0.0282 (0.0502)
<b>SONS' INTER-REGIONAL MIGRATION</b>				
1st Quintile	-0.0074 (0.0091)	-0.0075 (0.0091)	0.0065 (0.0119)	-0.0228* (0.0123)
2nd Quintile	-0.0014 (0.0100)	-0.0017 (0.0100)	0.0101 (0.0126)	-0.0056 (0.0154)
3rd Quintile	-0.0028 (0.0096)	-0.0040 (0.0094)	0.0097 (0.0128)	-0.0033 (0.0182)
4th Quintile	0.0062 (0.0109)	0.0057 (0.0108)	0.0073 (0.0130)	0.0131 (0.0242)
5th Quintile	-0.0039 (0.0098)	-0.0006 (0.0094)	-0.0171 (0.0127)	0.0108 (0.0330)
<b>SONS' WHITE COLLAR STATUS</b>				
1st Quintile	0.0012 (0.0070)	0.0011 (0.0070)	-0.0262* (0.0144)	-0.0119 (0.0192)
2nd Quintile	0.0023 (0.0081)	0.0022 (0.0081)	0.0144 (0.0155)	0.0421* (0.0229)
3rd Quintile	0.0007 (0.0082)	0.0006 (0.0084)	0.0384** (0.0162)	0.0646** (0.0280)
4th Quintile	-0.0017 (0.0069)	-0.0021 (0.0071)	0.0479*** (0.0169)	0.1345*** (0.0412)
5th Quintile	0.0203* (0.0119)	0.0213* (0.0124)	0.0799*** (0.0180)	0.1247** (0.0524)
<b>SONS' BLUE COLLAR STATUS</b>				
1st Quintile	0.0000 (0.0156)	0.0002 (0.0156)	-0.0153 (0.0178)	-0.0294 (0.0233)
2nd Quintile	-0.0029 (0.0169)	-0.0029 (0.0169)	-0.0150 (0.0180)	-0.0352 (0.0245)
3rd Quintile	0.0221 (0.0193)	0.0214 (0.0193)	-0.0511*** (0.0175)	-0.0600** (0.0263)
4th Quintile	-0.0063 (0.0164)	-0.0029 (0.0166)	-0.0341* (0.0186)	-0.0564* (0.0315)
5th Quintile	-0.0115 (0.0155)	-0.0133 (0.0155)	-0.0780*** (0.0182)	-0.0226 (0.0454)

<b>SONS' MARITAL STATUS</b>				
1st Quintile	0.0003 (0.0247)	0.0032 (0.0247)	0.0183 (0.0223)	0.0160 (0.0319)
2nd Quintile	-0.0156 (0.0263)	-0.0200 (0.0263)	0.0024 (0.0225)	-0.0240 (0.0349)
3rd Quintile	-0.0228 (0.0251)	-0.0246 (0.0254)	-0.0207 (0.0228)	-0.1021** (0.0403)
4th Quintile	-0.0493* (0.0261)	-0.0519** (0.0263)	-0.0286 (0.0225)	-0.0726 (0.0490)
5th Quintile	-0.0371 (0.0272)	-0.0324 (0.0278)	0.0045 (0.0228)	-0.1806*** (0.0558)
<b>SONS' NUMBER OF CHILDREN</b>				
1st Quintile	-0.0922 (0.0960)	-0.0822 (0.0962)	0.0657 (0.0620)	0.1149 (0.0907)
2nd Quintile	0.0698 (0.1118)	0.0721 (0.1123)	0.0706 (0.0624)	0.0974 (0.1023)
3rd Quintile	-0.1108 (0.1089)	-0.1045 (0.1100)	-0.0747 (0.0622)	-0.2040* (0.1179)
4th Quintile	-0.1549 (0.1035)	-0.1649 (0.1046)	-0.0217 (0.0623)	-0.0744 (0.1674)
5th Quintile	0.0267 (0.1222)	0.0413 (0.1261)	-0.0634 (0.0625)	-0.2314 (0.1975)
<b>N</b>	<b>3,182</b>	<b>3,109</b>	<b>4,934</b>	<b>1,639</b>

Notes: See notes to Table A1.



**TABLE A3: RETURN TO FATHERS' OCCUPATION CATEGORIES  
(RELATIVE TO LABORERS)**

	BLACK		WHITE	
	U.S.	South Only	U.S.	South Only
<b>SONS' OCCUPATION INCOME SCORE RANK</b>				
Return to White Collar Status	4.3883*** (0.6412)	3.9178*** (0.6750)	21.7760*** (0.1558)	29.2695*** (0.4159)
Return to Blue Collar Status	3.3726*** (0.2684)	3.3008*** (0.3200)	9.9445*** (0.1389)	12.7533*** (0.3786)
Return to Farm Ownership	-0.2565 (0.3270)	-0.2972 (0.3317)	2.5032*** (0.4424)	3.3950*** (0.7995)
Return to Cash Renting	-1.2147*** (0.2843)	-1.2246*** (0.2839)	0.2223 (1.5535)	2.9528 (2.3344)
Return to Share Renting	-0.7134** (0.2798)	-0.7609*** (0.2795)	-0.4923 (0.9528)	2.4331 (1.5049)
<b>SONS' HOME OWNERSHIP (BINARY)</b>				
Return to White Collar Status	0.0343** (0.0158)	0.0412** (0.0204)	0.0015 (0.0020)	-0.0047 (0.0050)
Return to Blue Collar Status	0.0182*** (0.0066)	0.0256*** (0.0088)	0.0010 (0.0018)	0.0002 (0.0046)
Return to Farm Ownership	0.0774*** (0.0141)	0.0802*** (0.0146)	0.0745*** (0.0068)	0.0736*** (0.0120)
Return to Cash Renting	0.0063 (0.0108)	0.0085 (0.0108)	0.0286 (0.0216)	0.0014 (0.0337)
Return to Share Renting	-0.0057 (0.0082)	-0.0042 (0.0082)	0.0266* (0.0140)	0.0161 (0.0219)
<b>SONS' LITERACY (BINARY)</b>				
Return to White Collar Status	0.1680*** (0.0168)	0.2113*** (0.0217)	0.0505*** (0.0012)	0.1461*** (0.0041)
Return to Blue Collar Status	0.0872*** (0.0089)	0.1035*** (0.0120)	0.0394*** (0.0011)	0.1131*** (0.0043)
Return to Farm Ownership	0.1178*** (0.0171)	0.1240*** (0.0180)	0.0416*** (0.0030)	0.1221*** (0.0081)
Return to Cash Renting	0.0410** (0.0174)	0.0451*** (0.0175)	0.0272** (0.0132)	0.0641** (0.0319)
Return to Share Renting	0.0346** (0.0138)	0.0390*** (0.0140)	0.0292*** (0.0081)	0.0884*** (0.0179)
<b>SONS' LABORER STATUS (BINARY)</b>				
Return to White Collar Status	-0.1135*** (0.0243)	-0.0864*** (0.0292)	-0.2457*** (0.0026)	-0.2619*** (0.0062)
Return to Blue Collar Status	-0.0886*** (0.0112)	-0.0570*** (0.0137)	-0.1851*** (0.0025)	-0.1729*** (0.0062)
Return to Farm Ownership	-0.1086*** (0.0183)	-0.1148*** (0.0189)	-0.0824*** (0.0080)	-0.0984*** (0.0145)
Return to Cash Renting	-0.0675*** (0.0172)	-0.0655*** (0.0173)	-0.0179 (0.0289)	-0.0698 (0.0444)
Return to Share Renting	-0.0856*** (0.0138)	-0.0831*** (0.0139)	-0.0236 (0.0179)	-0.1101*** (0.0270)
<b>N</b>	<b>25,025</b>	<b>20,128</b>	<b>282,938</b>	<b>44,456</b>

*Notes:* Table contains regression coefficient for occupation dummies, representing fathers' 1880 occupation category, and are interpreted relative laborer fathers. Sample includes all fathers with a recorded occupation. Regressions also include fixed effects for sons' age in 1900, fathers' age in 1880, fathers' urban status, and fathers' 1880 state fixed effects.

**TABLE A4: RETURN TO FATHERS' OCCUPATION CATEGORIES  
(RELATIVE TO LABORERS), OTHER OUTCOMES**

	BLACK		WHITE	
	U.S.	South Only	U.S.	South Only
<b>SONS' STATE-TO-STATE MIGRATION</b>				
Return to White Collar Status	0.0775*** (0.0209)	0.0712*** (0.0254)	0.0672*** (0.0025)	0.0480*** (0.0058)
Return to Blue Collar Status	0.0176** (0.0088)	0.0285*** (0.0110)	0.0262*** (0.0022)	0.0322*** (0.0055)
Return to Farm Ownership	-0.0270** (0.0124)	-0.0228* (0.0127)	-0.0441*** (0.0065)	-0.0398*** (0.0114)
Return to Cash Renting	-0.0210* (0.0118)	-0.0183 (0.0118)	0.0094 (0.0251)	0.0491 (0.0411)
Return to Share Renting	-0.0351*** (0.0090)	-0.0308*** (0.0090)	0.0078 (0.0157)	0.0169 (0.0237)
<b>SONS' INTER-REGIONAL MIGRATION</b>				
Return to White Collar Status	0.0256* (0.0134)	0.0253 (0.0167)	0.0405*** (0.0017)	0.0406*** (0.0039)
Return to Blue Collar Status	0.0197*** (0.0065)	0.0317*** (0.0085)	0.0157*** (0.0014)	0.0265*** (0.0038)
Return to Farm Ownership	-0.0159** (0.0074)	-0.0122 (0.0075)	-0.0087** (0.0042)	-0.0042 (0.0073)
Return to Cash Renting	-0.0188*** (0.0059)	-0.0169*** (0.0059)	0.0094 (0.0156)	0.0225 (0.0218)
Return to Share Renting	-0.0147*** (0.0055)	-0.0125** (0.0055)	-0.0104 (0.0092)	0.0044 (0.0118)
<b>SONS' WHITE COLLAR STATUS</b>				
Return to White Collar Status	0.1026*** (0.0165)	0.0872*** (0.0185)	0.4329*** (0.0026)	0.4518*** (0.0061)
Return to Blue Collar Status	0.0150*** (0.0045)	0.0098** (0.0050)	0.1008*** (0.0022)	0.0958*** (0.0051)
Return to Farm Ownership	0.0061 (0.0056)	0.0072 (0.0059)	0.0287*** (0.0064)	0.0311*** (0.0101)
Return to Cash Renting	-0.0017 (0.0043)	-0.0019 (0.0044)	0.0122 (0.0193)	0.0325 (0.0292)
Return to Share Renting	-0.0009 (0.0035)	-0.0011 (0.0035)	0.0008 (0.0116)	-0.0031 (0.0166)
<b>SONS' BLUE COLLAR STATUS</b>				
Return to White Collar Status	0.0624*** (0.0206)	0.0692*** (0.0239)	-0.1199*** (0.0028)	-0.0257*** (0.0055)
Return to Blue Collar Status	0.1488*** (0.0101)	0.1463*** (0.0123)	0.1332*** (0.0026)	0.1950*** (0.0058)
Return to Farm Ownership	-0.0099 (0.0113)	-0.0029 (0.0116)	-0.1508*** (0.0065)	-0.0486*** (0.0100)
Return to Cash Renting	-0.0272*** (0.0095)	-0.0249*** (0.0095)	-0.1150*** (0.0222)	-0.0162 (0.0316)
Return to Share Renting	-0.0145* (0.0082)	-0.0119 (0.0082)	-0.1122*** (0.0142)	-0.0062 (0.0207)

<b>SONS' MARITAL STATUS</b>				
Return to White Collar Status	-0.0374*	-0.0527**	-0.0692***	-0.1419***
	(0.0213)	(0.0256)	(0.0027)	(0.0065)
Return to Blue Collar Status	-0.0249**	-0.0273**	-0.0190***	-0.0596***
	(0.0101)	(0.0125)	(0.0025)	(0.0061)
Return to Farm Ownership	-0.0308*	-0.0362**	-0.0015	-0.0448***
	(0.0162)	(0.0166)	(0.0078)	(0.0138)
Return to Cash Renting	0.0124	0.0091	-0.0091	-0.0121
	(0.0147)	(0.0148)	(0.0265)	(0.0365)
Return to Share Renting	0.0129	0.0090	-0.0014	-0.0135
	(0.0119)	(0.0120)	(0.0170)	(0.0276)
<b>SONS' NUMBER OF CHILDREN</b>				
Return to White Collar Status	-0.2119***	-0.3182***	-0.2800***	-0.4733***
	(0.0654)	(0.0778)	(0.0070)	(0.0183)
Return to Blue Collar Status	-0.1172***	-0.1629***	-0.0876***	-0.2042***
	(0.0315)	(0.0410)	(0.0064)	(0.0174)
Return to Farm Ownership	0.1097*	0.1168*	-0.0301	-0.0476
	(0.0633)	(0.0661)	(0.0200)	(0.0396)
Return to Cash Renting	0.1085*	0.1012*	-0.0906	-0.0922
	(0.0583)	(0.0585)	(0.0730)	(0.1288)
Return to Share Renting	0.1593***	0.1505***	-0.0477	-0.0673
	(0.0473)	(0.0474)	(0.0440)	(0.0737)
<b>N</b>	<b>25,025</b>	<b>20,128</b>	<b>282,938</b>	<b>44,456</b>

Notes: See notes to Table A3.

## APPENDIX II. DATA APPENDIX

### a. LINKED SAMPLE

We begin with the 1880 full count sample, originally provided by the North American Population Project (MPC 2017; Ruggles 2015). We restrict the sample to all males aged 0 to 17 in 1880 living in the same household as their father or stepfather, regardless of whether the father was a household head. We do restrict the sample to include only sons who are household heads in 1900.

Our linking methodology follows the advice of Bailey et al. (2020). We link the 1880 sons to the 1900 selves using two distinct methods and then define our analysis sample as the intersection of those two census-to-census linked samples. From Bailey et al., “to the extent that different methods make errors for different reasons, taking the set of common links helps avoid idiosyncratic reasons for errors” (Bailey et al. (2020), p. 1038).

The two linking methods we employ are well-known in this literature. The first, based on Ferrie (1996) and executed using code from Bailey and Cole (2019), links on name, place of birth, race, and age. Links are accepted if ages match within a 5-year window, and we use actual names rather than NYSIIS variants. Matches are accepted in this methodology if they are unique and exact based on last name, first four letters of first name, and (if provided) middle initial. This method prioritizes matches based on age; for a given individual, if there are more than 10 potential matches based on name, birthplace, and race, the case is dropped. If there are 10 or fewer potential matches but greater than one potential match, the potential match which most closely matches on age is selected. Ties result in no match.

The second methodology is based on Abramitzky, Boustan, and Eriksson (2014), executed using code provided by the authors (hereafter, ABE).<sup>1</sup> In this method, matches are made based on the same set of characteristics as before: name, place of birth, race, and age, but we employ NYSIIS coding of names rather than actual names.<sup>2</sup> We use the ABE methodology for robustness, which restricts the set of linked names to those that are unique within +/- 2 years of age. Matches are exact matches only, meaning they must match on last, first, and (if provided) middle initial NYSIIS codes to be included in the linked sample.

The analysis sample for this paper is the intersection of these two matched samples. Samples constructed in this way carry a substantially lower match rate than single-method match processes

---

<sup>1</sup> Available at <https://ranabr.people.stanford.edu/matching-codes>.

<sup>2</sup> By incorporating NYSIIS codes in one methodology and actual name strings in the other, our approach may be more robust to false matches.

alone. Black match rates are 8.40% using the Ferrie method and 7.60% using the ABE method. The intersected sample of Black sons has a match rate of 3.40%. White match rates are 13.8% using the Ferrie method and 15.3% using the ABE method. The intersected sample of White sons has a match rate of 9.50%. The Data Appendix to Collins and Wanamaker (2022) provides additional discussion of the qualities of the intersected sample.

To generate a final analysis sample that is representative of the 1880 population from which our initial data derive, we use propensity-based reweighting techniques, as recommended by Bailey et al. (2020). We employ a simple probit estimator for whether an individual was successfully matched to the 1900 Census where the probability is specified as a function of region of residence dummies, urban residence, 1880 city population, and age. The analysis sample is then re-weighted using inverse probability weights that downweight observations with a high probability of being matched to the 1900 census and upweight observations with a low probability of being matched. (See Table A.1. of Collins and Wanamaker (2022) for additional information.)

#### b. IMPUTING MISSING OCCUPATIONS FOR SONS OBSERVED IN 1900

The 100 percent Census samples provided to NBER From IPUMS have incomplete occupation coding. 13.1 percent of occupation codes are missing in the 1900 matched sample, but far fewer are missing occupation strings in the transcribed data. These omissions appear to be related to the presence or absence of an industry string or industry code. We assigned occupation codes to occupations using an algorithm based on the modal codes for other individuals with the same occupation string. Our method uses plurality-based coding, assigning occupation codes to occupation strings based on the most common IPUMS-assigned code for other individuals with the same occupation string. Observations with occupations still uncoded were then dropped from the analysis sample.

#### c. 1880 AGRICULTURE CENSUS LINK

The 1880 Census of Agriculture manuscripts contain a wealth of information that informs our analysis. For each farm property, the farm's operator is listed along with information on whether that operator is an owner, "rents for fixed money rental", or "rents for share of product". Acreage for each property is listed as tilled or unimproved, and acreage is further split between tilled acres that are in rotation that which is "permanent meadow, permanent pasture". Unimproved acreage is divided between "woodland and forest" and "other unimproved". Enumerators elicited the value of the farm itself (including land, fences, and buildings), the value of farming implements and machinery, and the

value of livestock. The enumerators also inquired about the estimated value of farm production, the production volume for each crop, the amount of labor hired, and costs of inputs. We use information on tenure (owner versus renter categories), farm value, livestock value, and farm acreage to derive the results in this paper. Assuming that the tenure designation is correct, we would argue that small errors in reported acreage and value are unlikely to confound our analysis very much since we array the data by quintiles. To the extent families are sorted into an incorrect quintile at random, it would tend to flatten the observed relationship between sons' and fathers' status.

This requires making links of farmers (who are fathers to sons 0-17) from the 1880 Census of Population records to the Census of Agriculture records, which are separate. Specifically, using the 1880 to 1900 linked sample of sons as the analysis sample, we take the additional step of linking a random subset of farming fathers to the 1880 Census of Agriculture. Only individuals who reported an occupation of “farmer” are included in the Census of Agriculture linkage process, and we restrict our work to fathers living in states with indexed, searchable Census of Agriculture files on Ancestry.com. The sets of searchable and unsearchable states are similar on a variety of agriculture sector metrics, including share of owner-operated farms (Collins and Wanamaker 2022).

Before searching for matches in the Census of Agriculture, we randomly sorted the farmers in our analysis sample for linking, and we oversampled Black farmers. Our linked sample includes 5,708 white farmers and 3,605 black farmers. To account for the random sampling procedure among farmers, we add an additional component to the sample probability weights described above. Each black farmer in our Ag Census linked sample represents  $B/3,605$  farmers and each White farmer represents  $W/5,708$  farmers where B and W represent the number of Black and White farmers, respectively, in the 1880 population census to 1900 population census linked sample.<sup>3</sup> We multiply the probability sampling weights by these representation weights to generate final analysis weights.

There are several reasons one might not find a farmer in the agricultural census. Names might be transcribed differently from the population and agricultural census records; “farmers” in the population census might not have been the sole operator (the agricultural census lists only one); enumerators might not have completed a schedule for the farm or it might not have been preserved; or we might have missed a farmer who is actually in the manuscripts somewhere. Our procedure implicitly assumes individuals who are unsuccessfully linked to the 1880 Census of Agriculture are missing at random.

---

<sup>3</sup> Our analysis sample is further reduced by incomplete/illegible agriculture census images (145 in total) and by missing occupations for fathers or for sons. The final sample sizes are contained in Tables A1 (farmers only) and A3 (all father occupations).

#### d. ASSIGNMENT OF INCOME SCORES AND INCOME SCORE RANKS

One measure of sons' outcomes in 1900 is based on an occupational income score, which is then ranked. We assign this score based on the individual's occupation, region of residence, gender, and race. Income scores are derived from a combination of information on earnings from the 1940 (full count) and 1960 (5% sample) Census of Population data. Note that the 1940 Census was the first to inquire about annual earnings. If imputations would be based on fewer than 50 underlying earnings observations, we drop the location/region covariate from the procedure. If the problem persists, we create cells based on one-digit rather than three-digit occupation. Our approach is similar in spirit to the *occscore* variable that is widely available in IPUMS samples (Ruggles et. al 2015), but incorporates the additional considerations of race, gender, farm tenure, and location. These additional covariates provide valuable information for understanding intergenerational transmission of status (e.g., Inwood, Minns, and Summerfield 2019; Ward 2023, Saavedra and Twinam 2020).

The 1940 Census provides a valuable, but incomplete, record of earnings. Unfortunately, earnings from self-employment, including farming, were excluded. We assign incomes to self-employed individuals who are *not* farmers using the 1960 Census enumeration where earnings of wage and salary workers *and* self-employed workers are available. The ratio of income for self-employed relative to wage and salary workers within an occupation in the 1960 Census gives us a multiplier, which we then apply to wage and salary income for each race/gender/region/occupation cell from the 1940 Census to estimate income for self-employed workers in that year.

For farmers, our method reflects both differences between renters and owners and the value of perquisites enjoyed by farm residents. For farming sons in 1900, we use the *home ownership* variable in the Census of Population to identify likely farm owners. To rank farmers' income (both owners and renters) relative to individuals in our sample with non-farming occupations, we require a 1940-based occupational income score. We derive an estimate, separately for owners and renters and by race, gender, and region of residence, by assuming that the ratio of farmer (owning and non-owning) income to farm laborer income is constant between 1940 and 1960, and then exploiting the fact that farm laborer income is available in 1940 and 1960 and farmer income (owning and non-owning) is available in 1960. In estimating these values, we use the all-inclusive income of farmers and farm laborers, including the value of perquisites as discussed below.

*Farm laborer* incomes are reported in the 1940 and 1960 Census of Population enumerations. For the 1940 income observations, we scale up reported incomes for farm laborers by the ratio of perquisites and cash wages to cash wages alone in 1939 (equal to 1.26), as reported in the 1957 USDA report *Major Statistical Series of the U.S. Department of Agriculture: Volume 3, Gross and Net Farm*

*Income* (Table 7).<sup>4</sup> We use this scaling factor to scale up the 1940-based occupation income score for each race, gender, and region of residence. We repeat this exercise for 1960 farm laborer income observations using the last available year of the same USDA report, 1956. For the 1960 calculation, the ratio is 1.19.

*Farmer* incomes in 1960 are available in the population schedules, and we derive perquisite values (1956 based) from the same USDA report cited above. Farmer perquisite rates are more volatile year-to-year, so we use additional sources to smooth the rate between 1956 and 1959. In the end, we calculate a perquisite-included income ratio for farmers in the 1960 Census of 1.35. Combining all of this information allows us to estimate the 1940 income of owning and non-owning farmers, inclusive of perquisites, to anchor an occupational income score. For farm managers in all years, we estimate a perquisite rate as the average of farmers and farm laborers: 34% in 1940 and 27% in 1960.

---

<sup>4</sup> We divide “Value of Perquisites – TOTAL” by “Cash Wages + Value of Perquisites – TOTAL” to calculate the perquisite rate. See Volume 3, Table 7.



## REFERENCES

- Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2014. "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration." *Journal of Political Economy* 122, 3: 467-506.
- Bailey, Martha J. and Connor Cole. *Autolink.ado*. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [distributor], 2019-06-13.
- Bailey, Martha, Connor Cole, Morgan Henderson, and Catherine Massey. 2020. "How Well Do Automated Linking Methods Perform in Historical Samples? Evidence from US Historical Data." *Journal of Economic Literature* 58, 4: 997-1044.
- Collins, William and Marianne H. Wanamaker. 2022. "African American Intergenerational Mobility since 1880." *American Economic Journal: Applied* 14, 3: 84-117.
- Ferrie, Joseph P. 1996. "A New Sample of Males Linked from the Public Use Microdata Sample of the 1850 US Federal Census of Population to the 1860 US Federal Census Manuscript Schedules." *Historical Methods* 29, 4: 141-156.
- Inwood, Kris, Chris Minns, and Fraser Summerfield. 2019. "Occupational Earnings Scores and Immigrant Assimilation: Evidence from the Canadian Census." *Explorations in Economic History* 72 (April): 114-122.
- Minnesota Population Center. 2017. *North Atlantic Population Project: Complete Count Microdata. Version 2.3* [dataset]. Minneapolis: Minnesota Population Center. <http://doi.org/10.18128/D040.V2.3>.
- Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2015. *Integrated Public Use Microdata Series: Version 6.0* [database]. Minneapolis: University of Minnesota. <http://doi.org/10.18128/D010.V6.0>.
- Saavedra, Martin and Tate Twinam. 2020. "A Machine Learning Approach to Improving Occupational Income Scores." *Explorations in Economic History* 75.
- U.S. Department of Agriculture. 1957. *Major Statistical Series of the U.S. Department of Agriculture*. Volumes 3 and 7. Washington, DC: Government Printing Office.
- Ward, Zachary. 2023. "Intergenerational Mobility in American History: Accounting for Race and Measurement Error." NBER Working Paper 29256.