*Long-run Impacts of Agricultural Shocks on Educational Attainment: Evidence from the Boll Weevil*

Richard B. Baker, John Blanchette, and Katherine Eriksson

Online Appendix A

A.1 Matching Algorithm Description

We construct our linked census data by linking individuals on the basis of first name, last name, birth year, race, and birth state using the following procedure developed by Abramitzky, Boustan, and Eriksson (2012):[[1]](#footnote-1)

1. In each dataset, we replace common nicknames with proper first names and then standardize first and last names using the NYSIIS procedure (Atack, Bateman, and Gregson 1992).
2. Starting with 1940 we match backwards to each of the earlier censuses in the following way:
   1. Restrict the 1940 dataset to individuals who are unique by standardized name, birth year, race, and birth state.
   2. For each remaining observation in 1940, look for exact matches on the specified matching variables in the earlier dataset. If there is one and only one match, consider this observation matched. If there is more than one match, drop the observation in 1940. Otherwise, leave the observation unmatched.
   3. For each remaining unmatched observation in 1940, look for matches one year off in birth year in either direction. Follow the same decision rules as (b).
   4. For each observation in 1940 that is still unmatched and has not been dropped, look for a unique match two years off in birth year in either direction. If a unique match is found, consider this observation matched. Otherwise, drop the 1940 observation.
3. Conduct the matching process described in step 2 above, but starting with each of the earlier census years and matching forwards to 1940.
4. Take the intersection of the resulting backward- and forward-matched samples.

References

Abramitzky, Ran, Leah Platt Boustan, Katherine Eriksson, James Feigenbaum, and Santiago Perez. “Automated Linking of Historical Data.” Working paper, 2018.

Atack, Jeremy, Fred Bateman, and Mary Eschelbach Gregson. “‘Matchmaker, Matchmaker, Make Me a Match’: A General Personal Computer-Based Matching Program for Historical Research.” *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 25, no. 2 (1992): 53–65.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Table A.1 | | | | | | |
| ROBUSTNESS OF ESTIMATES TO MEASURES OF THE BOLL WEEVIL INFESTATION | | | | | | |
|  | Years of Schooling | | | Completed 8th Grade | | |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
|  |  |  |  |  |  |  |
| *Panel (a): White* | | | | | | |
| Age exposed: |  |  |  |  |  |  |
| 4–6 | 0.2669\*\*\* | 0.2630\*\*\* | 0.2566\*\*\* | 0.0177\* | 0.0260\*\* | 0.0173+ |
|  | (0.0608) | (0.0637) | (0.0642) | (0.0086) | (0.0084) | (0.0090) |
| 7–9 | 0.2364\*\*\* | 0.2515\*\*\* | 0.2087\*\*\* | 0.0263\*\*\* | 0.0298\*\*\* | 0.0243\*\*\* |
|  | (0.0500) | (0.0514) | (0.0502) | (0.0069) | (0.0069) | (0.0070) |
| 10–12 | 0.1481\*\*\* | 0.1815\*\*\* | 0.1374\*\*\* | 0.0175\*\*\* | 0.0238\*\*\* | 0.0166\*\* |
|  | (0.0378) | (0.0394) | (0.0387) | (0.0052) | (0.0053) | (0.0053) |
| 13–15 | 0.1019\*\*\* | 0.0948\*\* | 0.0972\*\*\* | 0.0140\*\*\* | 0.0147\*\*\* | 0.0150\*\*\* |
|  | (0.0287) | (0.0296) | (0.0289) | (0.0040) | (0.0041) | (0.0040) |
| 16–18 | 0.0452\* | 0.0630\*\* | 0.0326 | 0.0058+ | 0.0090\*\* | 0.0052 |
|  | (0.0226) | (0.0239) | (0.0232) | (0.0031) | (0.0032) | (0.0031) |
|  |  |  |  |  |  |  |
| Observations | 429,757 | 429,390 | 429,048 | 429,757 | 429,390 | 429,048 |
| R-squared | 0.1627 | 0.1623 | 0.1628 | 0.1169 | 0.1168 | 0.1171 |
|  |  |  |  |  |  |  |
| *Panel (b): Black* | | | | | | |
| Age exposed: |  |  |  |  |  |  |
| 4–6 | 0.3579\*\*\* | 0.3109\*\*\* | 0.3457\*\*\* | 0.0327\*\*\* | 0.0258\*\* | 0.0257\*\* |
|  | (0.0724) | (0.0760) | (0.0752) | (0.0093) | (0.0094) | (0.0096) |
| 7–9 | 0.2427\*\*\* | 0.1977\*\* | 0.2425\*\*\* | 0.0249\*\*\* | 0.0170\* | 0.0228\*\* |
|  | (0.0601) | (0.0636) | (0.0616) | (0.0074) | (0.0076) | (0.0077) |
| 10–12 | 0.1514\*\* | 0.1392\*\* | 0.1585\*\* | 0.0128\* | 0.0085 | 0.0100+ |
|  | (0.0467) | (0.0491) | (0.0488) | (0.0058) | (0.0060) | (0.0059) |
| 13–15 | 0.1409\*\*\* | 0.0945\* | 0.1236\*\* | 0.0116\* | 0.0064 | 0.0087+ |
|  | (0.0390) | (0.0407) | (0.0403) | (0.0047) | (0.0049) | (0.0047) |
| 16–18 | 0.0609\* | 0.0349 | 0.0645\* | 0.0035 | -0.0009 | 0.0034 |
|  | (0.0304) | (0.0335) | (0.0311) | (0.0036) | (0.0039) | (0.0038) |
|  |  |  |  |  |  |  |
| Observations | 170,839 | 170,596 | 170,429 | 170,839 | 170,596 | 170,429 |
| R-squared | 0.0908 | 0.0898 | 0.0911 | 0.0566 | 0.0558 | 0.0570 |
| Measure of boll weevil arrival: |  |  |  |  |  |  |
| Year of infestation? | YES | NO | NO | YES | NO | NO |
| First arrival year? | NO | YES | NO | NO | YES | NO |
| Complete infestation year? | NO | NO | YES | NO | NO | YES |
| *Notes*: The dependent variables are given in the column headings. Standard errors adjusted for clustering by childhood county of residence are in parentheses. All specifications include year of birth fixed effects, childhood county of residence fixed effects, and controls for family background. Family background controls include childhood household head's occupational score, homeownership status, and literacy, as well as indicators for urban location and farm residence. +, \*, \*\*, and \*\*\* indicate significance at the 10, 5, 1, and 0.1 percent levels, respectively. | | | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Table A.2 | | | | | | |
| ROBUSTNESS OF THE BOLL WEEVIL'S EFFECT ON YEARS OF SCHOOLING TO SAMPLE SELECTION | | | | | | |
|  | White | | | | | |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| Age exposed: |  |  |  |  |  |  |
| 4–6 | 0.2669\*\*\* | 0.3100\*\*\* | 0.2843\*\*\* | 0.3061\*\* | 0.2380\*\*\* | 0.2704\*\*\* |
|  | (0.0608) | (0.0544) | (0.0580) | (0.1062) | (0.0646) | (0.0622) |
| 7–9 | 0.2364\*\*\* | 0.2542\*\*\* | 0.2340\*\*\* | 0.2206\* | 0.2066\*\*\* | 0.2296\*\*\* |
|  | (0.0500) | (0.0444) | (0.0479) | (0.0880) | (0.0527) | (0.0516) |
| 10–12 | 0.1481\*\*\* | 0.1602\*\*\* | 0.1499\*\*\* | 0.1154 | 0.1231\*\* | 0.1346\*\*\* |
|  | (0.0378) | (0.0337) | (0.0363) | (0.0722) | (0.0406) | (0.0388) |
| 13–15 | 0.1019\*\*\* | 0.0894\*\*\* | 0.0970\*\*\* | 0.0660 | 0.0742\* | 0.0930\*\* |
|  | (0.0287) | (0.0257) | (0.0278) | (0.0521) | (0.0308) | (0.0301) |
| 16–18 | 0.0452\* | 0.0308 | 0.0353 | -0.0165 | 0.0342 | 0.0463+ |
|  | (0.0226) | (0.0198) | (0.0218) | (0.0393) | (0.0243) | (0.0238) |
| County-level sample requirements: |  |  |  |  |  |  |
| Full infestation prior to 1923? | YES | YES | YES | YES | YES | YES |
| No full retreat of boll weevil? | YES | NO | YES | YES | YES | YES |
| Full infestation in ≤ 4 years? | YES | NO | NO | YES | YES | YES |
| Infestation prior to 1915? | NO | NO | NO | YES | NO | NO |
| No border changes? | NO | NO | NO | NO | YES | NO |
| 1889 acres in cotton ≥ 100? | NO | NO | NO | NO | NO | YES |
| Observations | 429,757 | 551,409 | 451,746 | 145,230 | 365,212 | 405,937 |
| R-squared | 0.1627 | 0.1589 | 0.1611 | 0.1604 | 0.1589 | 0.1624 |
|  | Black | | | | | |
|  | (7) | (8) | (9) | (10) | (11) | (12) |
| Age exposed: |  |  |  |  |  |  |
| 4–6 | 0.3579\*\*\* | 0.3557\*\*\* | 0.3652\*\*\* | 0.4769\*\* | 0.3683\*\*\* | 0.3650\*\*\* |
|  | (0.0724) | (0.0694) | (0.0725) | (0.1470) | (0.0783) | (0.0730) |
| 7–9 | 0.2427\*\*\* | 0.2425\*\*\* | 0.2492\*\*\* | 0.3713\*\* | 0.2354\*\*\* | 0.2398\*\*\* |
|  | (0.0601) | (0.0579) | (0.0603) | (0.1171) | (0.0647) | (0.0612) |
| 10–12 | 0.1514\*\* | 0.1583\*\*\* | 0.1526\*\* | 0.1970\* | 0.1638\*\* | 0.1429\*\* |
|  | (0.0467) | (0.0448) | (0.0467) | (0.0878) | (0.0499) | (0.0474) |
| 13–15 | 0.1409\*\*\* | 0.1289\*\*\* | 0.1379\*\*\* | 0.1707\* | 0.1569\*\*\* | 0.1445\*\*\* |
|  | (0.0390) | (0.0368) | (0.0387) | (0.0739) | (0.0421) | (0.0399) |
| 16–18 | 0.0609\* | 0.0522+ | 0.0593+ | 0.0558 | 0.0768\* | 0.0641\* |
|  | (0.0304) | (0.0287) | (0.0302) | (0.0493) | (0.0319) | (0.0311) |
| County-level sample requirements: |  |  |  |  |  |  |
| Full infestation prior to 1923? | YES | YES | YES | YES | YES | YES |
| No full retreat of boll weevil? | YES | NO | YES | YES | YES | YES |
| Full infestation in ≤ 4 years? | YES | NO | NO | YES | YES | YES |
| Infestation prior to 1915? | NO | NO | NO | YES | NO | NO |
| No border changes? | NO | NO | NO | NO | YES | NO |
| 1889 acres in cotton ≥ 100? | NO | NO | NO | NO | NO | YES |
| Observations | 170,839 | 187,083 | 172,938 | 65,408 | 143,547 | 164,526 |
| R-squared | 0.0908 | 0.0978 | 0.0910 | 0.0812 | 0.0914 | 0.0900 |
| *Notes*: The dependent variable is years of schooling. Standard errors adjusted for clustering by childhood county of residence are in parentheses. All specifications include year of birth fixed effects, childhood county of residence fixed effects, and controls for family background. Family background controls include childhood household head's occupational score, homeownership status, and literacy, as well as indicators for urban location and farm residence. Columns (1) and (7) repeat the baseline results shown in Table 4. +, \*, \*\*, and \*\*\* indicate significance at the 10, 5, 1, and 0.1 percent levels, respectively. | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table A.3 | | | | | | | | | | | | | | |
| ROBUSTNESS OF THE BOLL WEEVIL'S EFFECT ON EIGHTH GRADE COMPLETION TO SAMPLE SELECTION | | | | | | | | | | | | | | |
|  | | | White | | | | | | | | | | | |
|  | | | (1) | | (2) | | (3) | | (4) | | (5) | | (6) | |
| Age exposed: | | |  | |  | |  | |  | |  | |  | |
| 4–6 | | | 0.0177\* | | 0.0240\*\* | | 0.0194\* | | 0.0340\* | | 0.0145 | | 0.0205\* | |
|  | | | (0.0086) | | (0.0079) | | (0.0083) | | (0.0149) | | (0.0091) | | (0.0089) | |
| 7–9 | | | 0.0263\*\*\* | | 0.0259\*\*\* | | 0.0244\*\*\* | | 0.0288\* | | 0.0237\*\* | | 0.0275\*\*\* | |
|  | | | (0.0069) | | (0.0062) | | (0.0067) | | (0.0120) | | (0.0073) | | (0.0071) | |
| 10–12 | | | 0.0175\*\*\* | | 0.0146\*\* | | 0.0159\*\* | | 0.0142 | | 0.0149\*\* | | 0.0176\*\* | |
|  | | | (0.0052) | | (0.0046) | | (0.0050) | | (0.0098) | | (0.0055) | | (0.0054) | |
| 13–15 | | | 0.0140\*\*\* | | 0.0108\*\* | | 0.0122\*\* | | 0.0106 | | 0.0106\* | | 0.0139\*\* | |
|  | | | (0.0040) | | (0.0036) | | (0.0040) | | (0.0072) | | (0.0042) | | (0.0043) | |
| 16–18 | | | 0.0058+ | | 0.0024 | | 0.0034 | | -0.0030 | | 0.0040 | | 0.0069\* | |
|  | | | (0.0031) | | (0.0028) | | (0.0030) | | (0.0053) | | (0.0032) | | (0.0033) | |
| County-level sample requirements: | | |  | |  | |  | |  | |  | |  | |
| Full infestation prior to 1923? | | | YES | | YES | | YES | | YES | | YES | | YES | |
| No full retreat of boll weevil? | | | YES | | NO | | YES | | YES | | YES | | YES | |
| Full infestation in ≤ 4 years? | | | YES | | NO | | NO | | YES | | YES | | YES | |
| Infestation prior to 1915? | | | NO | | NO | | NO | | YES | | NO | | NO | |
| No border changes? | | | NO | | NO | | NO | | NO | | YES | | NO | |
| 1889 acres in cotton ≥ 100? | | | NO | | NO | | NO | | NO | | NO | | YES | |
| Observations | | | 429,757 | | 551,409 | | 451,746 | | 145,230 | | 365,212 | | 405,937 | |
| R-squared | | | 0.1169 | | 0.1124 | | 0.1148 | | 0.1138 | | 0.1147 | | 0.1165 | |
|  | | | Black | | | | | | | | | | | |
|  | | | (7) | | (8) | | (9) | | (10) | | (11) | | (12) | |
| Age exposed: | | |  | |  | |  | |  | |  | |  | |
| 4–6 | | | 0.0327\*\*\* | | 0.0283\*\* | | 0.0331\*\*\* | | 0.0559\*\*\* | | 0.0328\*\* | | 0.0356\*\*\* | |
|  | | | (0.0093) | | (0.0089) | | (0.0092) | | (0.0164) | | (0.0101) | | (0.0092) | |
| 7–9 | | | 0.0249\*\*\* | | 0.0232\*\* | | 0.0254\*\*\* | | 0.0456\*\*\* | | 0.0238\*\* | | 0.0245\*\* | |
|  | | | (0.0074) | | (0.0072) | | (0.0074) | | (0.0122) | | (0.0080) | | (0.0074) | |
| 10–12 | | | 0.0128\* | | 0.0144\*\* | | 0.0128\* | | 0.0164+ | | 0.0134\* | | 0.0121\* | |
|  | | | (0.0058) | | (0.0055) | | (0.0057) | | (0.0096) | | (0.0061) | | (0.0058) | |
| 13–15 | | | 0.0116\* | | 0.0119\*\* | | 0.0119\* | | 0.0140+ | | 0.0134\*\* | | 0.0120\* | |
|  | | | (0.0047) | | (0.0045) | | (0.0046) | | (0.0080) | | (0.0050) | | (0.0048) | |
| 16–18 | | | 0.0035 | | 0.0031 | | 0.0036 | | 0.0004 | | 0.0052 | | 0.0035 | |
|  | | | (0.0036) | | (0.0034) | | (0.0036) | | (0.0055) | | (0.0037) | | (0.0037) | |
| County-level sample requirements: | | |  | |  | |  | |  | |  | |  | |
| Full infestation prior to 1923? | | | YES | | YES | | YES | | YES | | YES | | YES | |
| No full retreat of boll weevil? | | | YES | | NO | | YES | | YES | | YES | | YES | |
| Full infestation in ≤ 4 years? | | | YES | | NO | | NO | | YES | | YES | | YES | |
| Infestation prior to 1915? | | | NO | | NO | | NO | | YES | | NO | | NO | |
| No border changes? | | | NO | | NO | | NO | | NO | | YES | | NO | |
| 1889 acres in cotton ≥ 100? | | | NO | | NO | | NO | | NO | | NO | | YES | |
| Observations | | | 170,839 | | 187,083 | | 172,938 | | 65,408 | | 143,547 | | 164,526 | |
| R-squared | | | 0.0566 | | 0.0635 | | 0.0569 | | 0.0461 | | 0.0562 | | 0.0555 | |
| *Notes*: The dependent variable is an indicator for reporting completion of eight or more years of schooling in the 1940 census. Standard errors adjusted for clustering by childhood county of residence are in parentheses. All specifications include year of birth fixed effects, childhood county of residence fixed effects, and controls for family background. Family background controls include childhood household head's occupational score, homeownership status, and literacy, as well as indicators for urban location and farm residence. Columns (1) and (7) repeat the baseline results shown in Table 4 for ease of comparison. +, \*, \*\*, and \*\*\* indicate significance at the 10, 5, 1, and 0.1 percent levels, respectively. | | | | | | | | | | | | | | |
| Table A.4 | | | | | | | | | | | | | | | |
| ESTIMATES OF THE BOLL WEEVIL'S EFFECT ON LN(ENROLLMENT) IN GEORGIA COUNTIES ONLY | | | | | | | | | | | | | | | |
|  | White | | | | | | | Black | | | | | | | |
|  | (1) | (2) | | (3) | | (4) | | (5) | | (6) | | (7) | | (8) | |
|  |  |  | |  | |  | |  | |  | |  | |  | |
| boll weevil | 0.018 | 0.004 | | 0.015 | | 0.006 | | 0.057\*\* | | 0.055\*\* | | 0.050\* | | 0.051\*\* | |
|  | (0.015) | (0.014) | | (0.015) | | (0.014) | | (0.020) | | (0.018) | | (0.021) | | (0.018) | |
| ln(teachers)? | NO | YES | | NO | | YES | | NO | | YES | | NO | | YES | |
| Time trends? | NO | NO | | YES | | YES | | NO | | NO | | YES | | YES | |
|  |  |  | |  | |  | |  | |  | |  | |  | |
| Observations | 1,887 | 1,887 | | 1,887 | | 1,887 | | 1,886 | | 1,880 | | 1,886 | | 1,880 | |
| Number of counties | 118 | 118 | | 118 | | 118 | | 118 | | 118 | | 118 | | 118 | |
| R-squared | 0.9735 | 0.9806 | | 0.9838 | | 0.9864 | | 0.9618 | | 0.9741 | | 0.9757 | | 0.9806 | |
| *Notes*: The dependent variable is the natural log of enrollment. Standard errors adjusted for clustering by county in parentheses. All regressions include year and county fixed effects. Columns (2), (4), (6), and (8) include the natural log of the number of teachers employed as a control for the supply of education. Columns (3), (4), (7), and (8) also include county-specific linear time trends. +, \*, \*\*, and \*\*\* indicate significance at the 10, 5, 1, and 0.1 percent levels, respectively. | | | | | | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table A.5 | | | | | | | | |
| ROBUSTNESS OF ESTIMATES TO RESTRICTING THE SAMPLE TO COUNTIES INFESTED AFTER SPECIFIED YEARS | | | | | | | | |
|  | Years of Schooling | | | | Completed 8th Grade | | | |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|  |  |  |  |  |  |  |  |  |
| *Panel (a): White* | | | | | | | | |
| Age exposed: |  |  |  |  |  |  |  |  |
| 4–6 | 0.2669\*\*\* | 0.2546\*\*\* | 0.2191\*\* | 0.0785 | 0.0177\* | 0.0085 | 0.0035 | -0.0124 |
|  | (0.0608) | (0.0708) | (0.0727) | (0.0959) | (0.0086) | (0.0094) | (0.0097) | (0.0128) |
| 7–9 | 0.2364\*\*\* | 0.2404\*\*\* | 0.2025\*\*\* | 0.1522\* | 0.0263\*\*\* | 0.0235\*\* | 0.0185\* | 0.0121 |
|  | (0.0500) | (0.0570) | (0.0588) | (0.0695) | (0.0069) | (0.0079) | (0.0082) | (0.0099) |
| 10–12 | 0.1481\*\*\* | 0.1681\*\*\* | 0.1363\*\* | 0.0904+ | 0.0175\*\*\* | 0.0186\*\*\* | 0.0144\* | 0.0074 |
|  | (0.0378) | (0.0396) | (0.0416) | (0.0482) | (0.0052) | (0.0055) | (0.0059) | (0.0071) |
| 13–15 | 0.1019\*\*\* | 0.1168\*\*\* | 0.0963\*\* | 0.0698+ | 0.0140\*\*\* | 0.0147\*\*\* | 0.0114\* | 0.0073 |
|  | (0.0287) | (0.0302) | (0.0318) | (0.0367) | (0.0040) | (0.0043) | (0.0045) | (0.0055) |
| 16–18 | 0.0452\* | 0.0591\* | 0.0569\* | 0.0525+ | 0.0058+ | 0.0068\* | 0.0060+ | 0.0052 |
|  | (0.0226) | (0.0233) | (0.0240) | (0.0273) | (0.0031) | (0.0032) | (0.0033) | (0.0039) |
|  |  |  |  |  |  |  |  |  |
| Observations | 429,757 | 392,109 | 373,389 | 321,796 | 429,757 | 392,109 | 373,389 | 321,796 |
| R-squared | 0.1627 | 0.1682 | 0.1627 | 0.1641 | 0.1169 | 0.1217 | 0.1194 | 0.1175 |
|  |  |  |  |  |  |  |  |  |
| *Panel (b): Black* | | | | | | | | |
| Age exposed: |  |  |  |  |  |  |  |  |
| 4–6 | 0.3579\*\*\* | 0.3282\*\*\* | 0.2986\*\*\* | 0.4569\*\*\* | 0.0327\*\*\* | 0.0290\*\* | 0.0245\* | 0.0340\* |
|  | (0.0724) | (0.0776) | (0.0799) | (0.1068) | (0.0093) | (0.0102) | (0.0105) | (0.0135) |
| 7–9 | 0.2427\*\*\* | 0.2323\*\*\* | 0.1978\*\* | 0.2620\*\* | 0.0249\*\*\* | 0.0239\*\* | 0.0193\* | 0.0142 |
|  | (0.0601) | (0.0645) | (0.0672) | (0.0867) | (0.0074) | (0.0082) | (0.0086) | (0.0113) |
| 10–12 | 0.1514\*\* | 0.1521\*\* | 0.1310\* | 0.1728\* | 0.0128\* | 0.0141\* | 0.0100 | 0.0081 |
|  | (0.0467) | (0.0491) | (0.0533) | (0.0668) | (0.0058) | (0.0060) | (0.0067) | (0.0082) |
| 13–15 | 0.1409\*\*\* | 0.1396\*\*\* | 0.1297\*\* | 0.1648\*\* | 0.0116\* | 0.0121\* | 0.0107\* | 0.0095 |
|  | (0.0390) | (0.0407) | (0.0427) | (0.0530) | (0.0047) | (0.0049) | (0.0052) | (0.0065) |
| 16–18 | 0.0609\* | 0.0566+ | 0.0468 | 0.0656+ | 0.0035 | 0.0034 | 0.0018 | 0.0032 |
|  | (0.0304) | (0.0313) | (0.0325) | (0.0379) | (0.0036) | (0.0037) | (0.0039) | (0.0046) |
|  |  |  |  |  |  |  |  |  |
| Observations | 170,839 | 160,855 | 150,920 | 124,377 | 170,839 | 160,855 | 150,920 | 124,377 |
| R-squared | 0.0908 | 0.0913 | 0.0866 | 0.0897 | 0.0566 | 0.0570 | 0.0566 | 0.0591 |
| Includes only counties infested in or after: | | | | | | | | |
| 1900? | YES | NO | NO | NO | YES | NO | NO | NO |
| 1906? | NO | YES | NO | NO | NO | YES | NO | NO |
| 1909? | NO | NO | YES | NO | NO | NO | YES | NO |
| 1912? | NO | NO | NO | YES | NO | NO | NO | YES |
| *Notes*: The dependent variables are given in the column headings. Standard errors adjusted for clustering by childhood county of residence are in parentheses. All specifications include year of birth fixed effects, childhood county of residence fixed effects, and controls for family background. Family background controls include childhood household head's occupational score, homeownership status, and literacy, as well as indicators for urban location and farm residence. +, \*, \*\*, and \*\*\* indicate significance at the 10, 5, 1, and 0.1 percent levels, respectively. | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table A.6 | | | | | | | | | | | | |
| ESTIMATES OF THE BOLL WEEVIL'S EFFECT ON GRADE COMPLETION, WHITE | | | | | | | | | | | | |
| Grade: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| Age exposed: |  |  |  |  |  |  |  |  |  |  |  |  |
| 4–6 | -0.0020 | -0.0037 | -0.0000 | 0.0011 | 0.0012 | 0.0084 | 0.0108 | 0.0177\* | 0.0497\*\*\* | 0.0527\*\*\* | 0.0440\*\*\* | 0.0358\*\*\* |
|  | (0.0027) | (0.0030) | (0.0036) | (0.0045) | (0.0059) | (0.0067) | (0.0080) | (0.0086) | (0.0086) | (0.0084) | (0.0073) | (0.0069) |
| 7–9 | 0.0017 | 0.0006 | 0.0047 | 0.0065+ | 0.0072 | 0.0150\*\* | 0.0216\*\*\* | 0.0263\*\*\* | 0.0422\*\*\* | 0.0395\*\*\* | 0.0275\*\*\* | 0.0223\*\*\* |
|  | (0.0024) | (0.0026) | (0.0032) | (0.0037) | (0.0049) | (0.0057) | (0.0064) | (0.0069) | (0.0072) | (0.0068) | (0.0060) | (0.0056) |
| 10–12 | 0.0025 | 0.0011 | 0.0045+ | 0.0055+ | 0.0067+ | 0.0128\*\* | 0.0130\*\* | 0.0175\*\*\* | 0.0235\*\*\* | 0.0227\*\*\* | 0.0147\*\*\* | 0.0115\*\* |
|  | (0.0021) | (0.0023) | (0.0026) | (0.0030) | (0.0040) | (0.0046) | (0.0050) | (0.0052) | (0.0052) | (0.0050) | (0.0043) | (0.0040) |
| 13–15 | 0.0023 | 0.0016 | 0.0046\* | 0.0071\*\* | 0.0092\*\* | 0.0127\*\*\* | 0.0124\*\* | 0.0140\*\*\* | 0.0105\* | 0.0087\* | 0.0070\* | 0.0071\* |
|  | (0.0015) | (0.0016) | (0.0020) | (0.0024) | (0.0032) | (0.0035) | (0.0038) | (0.0040) | (0.0041) | (0.0039) | (0.0035) | (0.0033) |
| 16–18 | 0.0007 | 0.0003 | 0.0023 | 0.0030 | 0.0039 | 0.0051+ | 0.0051+ | 0.0058+ | 0.0046 | 0.0040 | 0.0024 | 0.0027 |
|  | (0.0012) | (0.0013) | (0.0015) | (0.0018) | (0.0024) | (0.0026) | (0.0028) | (0.0031) | (0.0031) | (0.0031) | (0.0029) | (0.0027) |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| Observations | 429,757 | 429,757 | 429,757 | 429,757 | 429,757 | 429,757 | 429,757 | 429,757 | 429,757 | 429,757 | 429,757 | 429,757 |
| R-squared | 0.0467 | 0.0511 | 0.0563 | 0.0655 | 0.0795 | 0.0933 | 0.1004 | 0.1169 | 0.1155 | 0.1163 | 0.1113 | 0.1039 |
| Dependent variable: | | | | | | | | | | | | |
| Mean | 0.9768 | 0.9696 | 0.9541 | 0.9212 | 0.8577 | 0.7887 | 0.7030 | 0.5735 | 0.4252 | 0.3464 | 0.2749 | 0.2319 |
| Std. Dev. | 0.1506 | 0.1716 | 0.2092 | 0.2695 | 0.3493 | 0.4082 | 0.4569 | 0.4946 | 0.4944 | 0.4758 | 0.4465 | 0.4220 |
| *Notes*: The dependent variable in each column takes a value of one if the individual has completed the grade indicated in the column heading, and zero otherwise. Standard errors adjusted for clustering by childhood county of residence are in parentheses. All specifications include year of birth fixed effects, childhood county of residence fixed effects, and controls for family background. Family background controls include childhood household head's occupational score, homeownership status, and literacy, as well as indicators for urban location and farm residence. +, \*, \*\*, and \*\*\* indicate significance at the 10, 5, 1, and 0.1 percent levels, respectively. | | | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table A.7 | | | | | | | | | | | | |
| ESTIMATES OF THE BOLL WEEVIL'S EFFECT ON GRADE COMPLETION, BLACK | | | | | | | | | | | | |
| Grade: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| Age exposed: |  |  |  |  |  |  |  |  |  |  |  |  |
| 4–6 | 0.0165\* | 0.0274\*\*\* | 0.0423\*\*\* | 0.0441\*\*\* | 0.0457\*\*\* | 0.0417\*\*\* | 0.0394\*\*\* | 0.0327\*\*\* | 0.0179\* | 0.0165\* | 0.0121\* | 0.0110\* |
|  | (0.0074) | (0.0082) | (0.0102) | (0.0111) | (0.0126) | (0.0121) | (0.0105) | (0.0093) | (0.0080) | (0.0071) | (0.0062) | (0.0056) |
| 7–9 | 0.0123\* | 0.0181\* | 0.0277\*\*\* | 0.0275\*\* | 0.0279\*\* | 0.0257\*\* | 0.0283\*\*\* | 0.0249\*\*\* | 0.0163\*\* | 0.0130\* | 0.0070 | 0.0069 |
|  | (0.0062) | (0.0070) | (0.0083) | (0.0089) | (0.0101) | (0.0100) | (0.0084) | (0.0074) | (0.0062) | (0.0055) | (0.0046) | (0.0045) |
| 10–12 | 0.0080 | 0.0143\* | 0.0262\*\*\* | 0.0264\*\*\* | 0.0209\*\* | 0.0157\* | 0.0155\* | 0.0128\* | 0.0063 | 0.0051 | 0.0014 | 0.0011 |
|  | (0.0052) | (0.0057) | (0.0068) | (0.0073) | (0.0077) | (0.0076) | (0.0067) | (0.0058) | (0.0044) | (0.0041) | (0.0035) | (0.0033) |
| 13–15 | 0.0090\* | 0.0131\*\* | 0.0193\*\*\* | 0.0228\*\*\* | 0.0209\*\* | 0.0142\* | 0.0160\*\* | 0.0116\* | 0.0046 | 0.0038 | 0.0022 | 0.0029 |
|  | (0.0040) | (0.0044) | (0.0052) | (0.0058) | (0.0064) | (0.0062) | (0.0052) | (0.0047) | (0.0037) | (0.0035) | (0.0031) | (0.0028) |
| 16–18 | 0.0039 | 0.0063+ | 0.0119\*\* | 0.0110\* | 0.0066 | 0.0056 | 0.0060 | 0.0035 | 0.0002 | 0.0016 | 0.0006 | 0.0010 |
|  | (0.0033) | (0.0036) | (0.0042) | (0.0047) | (0.0050) | (0.0050) | (0.0041) | (0.0036) | (0.0029) | (0.0025) | (0.0022) | (0.0021) |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| Observations | 170,839 | 170,839 | 170,839 | 170,839 | 170,839 | 170,839 | 170,839 | 170,839 | 170,839 | 170,839 | 170,839 | 170,839 |
| R-squared | 0.0330 | 0.0387 | 0.0462 | 0.0535 | 0.0608 | 0.0638 | 0.0622 | 0.0566 | 0.0525 | 0.0484 | 0.0419 | 0.0371 |
| Dependent variable: | | | | | | | | | | | | |
| Mean | 0.9049 | 0.8705 | 0.7914 | 0.6733 | 0.5126 | 0.3885 | 0.2787 | 0.1980 | 0.1067 | 0.0777 | 0.0577 | 0.0489 |
| Std. Dev. | 0.2934 | 0.3358 | 0.4063 | 0.4690 | 0.4998 | 0.4874 | 0.4483 | 0.3985 | 0.3087 | 0.2677 | 0.2332 | 0.2156 |
| *Notes*: The dependent variable in each column takes a value of one if the individual has completed the grade indicated in the column heading, and zero otherwise. Standard errors adjusted for clustering by childhood county of residence are in parentheses. All specifications include year of birth fixed effects, childhood county of residence fixed effects, and controls for family background. Family background controls include childhood household head's occupational score, homeownership status, and literacy, as well as indicators for urban location and farm residence. +, \*, \*\*, and \*\*\* indicate significance at the 10, 5, 1, and 0.1 percent levels, respectively. | | | | | | | | | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Table A.8 | | | | | | |
| ROBUSTNESS OF ESTIMATES TO VARIOUS AGE GROUPS AS THE OMITTED CATEGORY | | | | | | |
|  | Years of Schooling | | | Completed 8th Grade | | |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
|  |  |  |  |  |  |  |
| *Panel (a): White* | | | | | | |
| Age exposed: |  |  |  |  |  |  |
| 4–6 | 0.2669\*\*\* | 0.2052\*\* | 0.1689\* | 0.0177\* | 0.0127 | 0.0057 |
|  | (0.0608) | (0.0674) | (0.0826) | (0.0086) | (0.0098) | (0.0122) |
| 7–9 | 0.2364\*\*\* | 0.1831\*\* | 0.1523\* | 0.0263\*\*\* | 0.0219\*\* | 0.0158 |
|  | (0.0500) | (0.0556) | (0.0672) | (0.0069) | (0.0079) | (0.0098) |
| 10–12 | 0.1481\*\*\* | 0.1057\* | 0.0806 | 0.0175\*\*\* | 0.0138\* | 0.0088 |
|  | (0.0378) | (0.0431) | (0.0520) | (0.0052) | (0.0060) | (0.0074) |
| 13–15 | 0.1019\*\*\* | 0.0715\* | 0.0537 | 0.0140\*\*\* | 0.0113\* | 0.0080 |
|  | (0.0287) | (0.0324) | (0.0387) | (0.0040) | (0.0046) | (0.0056) |
| 16–18 | 0.0452\* | 0.0263 | 0.0161 | 0.0058+ | 0.0041 | 0.0023 |
|  | (0.0226) | (0.0242) | (0.0272) | (0.0031) | (0.0033) | (0.0038) |
|  |  |  |  |  |  |  |
| Observations | 429,757 | 401,629 | 361,203 | 429,757 | 401,629 | 361,203 |
| R-squared | 0.1627 | 0.1641 | 0.1663 | 0.1169 | 0.1172 | 0.1180 |
|  |  |  |  |  |  |  |
| *Panel (b): Black* | | | | | | |
| Age exposed: |  |  |  |  |  |  |
| 4–6 | 0.3579\*\*\* | 0.3790\*\*\* | 0.4063\*\*\* | 0.0327\*\*\* | 0.0300\*\* | 0.0477\*\* |
|  | (0.0724) | (0.0879) | (0.1153) | (0.0093) | (0.0112) | (0.0153) |
| 7–9 | 0.2427\*\*\* | 0.2592\*\*\* | 0.2764\*\* | 0.0249\*\*\* | 0.0229\* | 0.0370\*\* |
|  | (0.0601) | (0.0716) | (0.0943) | (0.0074) | (0.0090) | (0.0124) |
| 10–12 | 0.1514\*\* | 0.1643\*\* | 0.1731\* | 0.0128\* | 0.0111 | 0.0215\* |
|  | (0.0467) | (0.0568) | (0.0742) | (0.0058) | (0.0069) | (0.0094) |
| 13–15 | 0.1409\*\*\* | 0.1526\*\*\* | 0.1598\*\* | 0.0116\* | 0.0105+ | 0.0176\* |
|  | (0.0390) | (0.0457) | (0.0558) | (0.0047) | (0.0055) | (0.0072) |
| 16–18 | 0.0609\* | 0.0694\* | 0.0744\* | 0.0035 | 0.0032 | 0.0070 |
|  | (0.0304) | (0.0331) | (0.0369) | (0.0036) | (0.0040) | (0.0047) |
|  |  |  |  |  |  |  |
| Observations | 170,839 | 159,451 | 142,666 | 170,839 | 159,451 | 142,666 |
| R-squared | 0.0908 | 0.0923 | 0.0947 | 0.0566 | 0.0581 | 0.0599 |
| Exposure ages of omitted category: | | | | | | |
| 19–30? | YES | NO | NO | YES | NO | NO |
| 19–27? | NO | YES | NO | NO | YES | NO |
| 19–24? | NO | NO | YES | NO | NO | YES |
| *Notes*: The dependent variables are given in the column headings. Standard errors adjusted for clustering by childhood county of residence are in parentheses. All specifications include year of birth fixed effects, childhood county of residence fixed effects, and controls for family background. Family background controls include childhood household head's occupational score, homeownership status, and literacy, as well as indicators for urban location and farm residence. +, \*, \*\*, and \*\*\* indicate significance at the 10, 5, 1, and 0.1 percent levels, respectively. | | | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Table A.9 | | | | | | |
| ROBUSTNESS OF ESTIMATES OF THE BOLL WEEVIL'S EFFECT ON LONG-RUN EDUCATIONAL OUTCOMES | | | | | | |
|  | Years of Schooling | | | | Completed 8th Grade | |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
|  | White | Black | White | Black | White | Black |
|  |  |  |  |  |  |  |
| Age exposed: |  |  |  |  |  |  |
| 4–6 | 0.3111\*\*\* | 0.2980\*\* | 0.2157\*\*\* | 0.3473\*\*\* | 0.0246\* | 0.0345\*\* |
|  | (0.0801) | (0.1064) | (0.0518) | (0.0682) | (0.0112) | (0.0132) |
| 7–9 | 0.2756\*\*\* | 0.1900\* | 0.2153\*\*\* | 0.2356\*\*\* | 0.0324\*\*\* | 0.0265\* |
|  | (0.0704) | (0.0911) | (0.0433) | (0.0566) | (0.0096) | (0.0109) |
| 10–12 | 0.1813\*\* | 0.1068 | 0.1359\*\*\* | 0.1537\*\*\* | 0.0226\*\* | 0.0141 |
|  | (0.0565) | (0.0748) | (0.0327) | (0.0438) | (0.0076) | (0.0092) |
| 13–15 | 0.1291\*\* | 0.1046+ | 0.0971\*\*\* | 0.1405\*\*\* | 0.0183\*\* | 0.0127+ |
|  | (0.0441) | (0.0610) | (0.0251) | (0.0364) | (0.0060) | (0.0073) |
| 16–18 | 0.0663+ | 0.0328 | 0.0398\* | 0.0583\* | 0.0091+ | 0.0044 |
|  | (0.0350) | (0.0491) | (0.0194) | (0.0286) | (0.0047) | (0.0058) |
| Exclude 19–21 from omitted category? | YES | YES | NO | NO | YES | YES |
| Top code years of schooling at 12? | NO | NO | YES | YES | NO | NO |
|  |  |  |  |  |  |  |
| Observations | 429,757 | 170,839 | 429,757 | 170,839 | 429,757 | 170,839 |
| R-squared | 0.1627 | 0.0908 | 0.1595 | 0.0914 | 0.1169 | 0.0566 |
| *Notes*: Standard errors adjusted for clustering by childhood county of residence are in parentheses. All specifications include year of birth fixed effects, childhood county of residence fixed effects, and controls for family background. Family background controls include childhood household head's occupational score, homeownership status, and literacy, as well as indicators for urban location and farm residence. +, \*, \*\*, and \*\*\* indicate significance at the 10, 5, 1, and 0.1 percent levels, respectively. | | | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Table A.10 | | | | | | |
| ROBUSTNESS OF THE BOLL WEEVIL'S EFFECT ON EIGHTH GRADE COMPLETION TO MATCHING | | | | | | |
|  | White | | | | | |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| Age exposed: |  |  |  |  |  |  |
| 4–6 | 0.0177\* | 0.0189\* | 0.0152 | 0.0184\* | 0.0159 | 0.0103 |
|  | (0.0086) | (0.0076) | (0.0093) | (0.0091) | (0.0098) | (0.0095) |
| 7–9 | 0.0263\*\*\* | 0.0244\*\*\* | 0.0254\*\*\* | 0.0258\*\*\* | 0.0252\*\* | 0.0176\* |
|  | (0.0069) | (0.0060) | (0.0075) | (0.0073) | (0.0079) | (0.0073) |
| 10–12 | 0.0175\*\*\* | 0.0153\*\*\* | 0.0146\*\* | 0.0161\*\* | 0.0135\* | 0.0139\* |
|  | (0.0052) | (0.0046) | (0.0056) | (0.0055) | (0.0060) | (0.0057) |
| 13–15 | 0.0140\*\*\* | 0.0146\*\*\* | 0.0125\*\* | 0.0128\*\* | 0.0117\* | 0.0092\* |
|  | (0.0040) | (0.0037) | (0.0044) | (0.0043) | (0.0047) | (0.0044) |
| 16–18 | 0.0058+ | 0.0065\* | 0.0055+ | 0.0050 | 0.0049 | 0.0037 |
|  | (0.0031) | (0.0028) | (0.0033) | (0.0033) | (0.0035) | (0.0033) |
| Match requirements: |  |  |  |  |  |  |
| Unique within 3 years? | YES | NO | YES | YES | YES | YES |
| Unique within 5 years? | NO | NO | YES | NO | YES | NO |
| Difference in birth year ≤ 1? | NO | NO | NO | YES | YES | NO |
| Match on exact names? | NO | NO | NO | NO | NO | YES |
| Match Rate | 27.3749 | 35.7069 | 23.5834 | 25.0809 | 21.4880 | 24.6369 |
| Observations | 429,757 | 567,158 | 367,879 | 382,444 | 325,555 | 358,976 |
| R-squared | 0.1169 | 0.0995 | 0.1230 | 0.1212 | 0.1274 | 0.1174 |
|  | Black | | | | | |
|  | (7) | (8) | (9) | (10) | (11) | (12) |
| Age exposed: |  |  |  |  |  |  |
| 4–6 | 0.0327\*\*\* | 0.0215\*\* | 0.0323\*\* | 0.0309\*\* | 0.0312\* | 0.0323\*\* |
|  | (0.0093) | (0.0080) | (0.0106) | (0.0109) | (0.0123) | (0.0109) |
| 7–9 | 0.0249\*\*\* | 0.0193\*\* | 0.0200\* | 0.0309\*\*\* | 0.0274\*\* | 0.0267\*\* |
|  | (0.0074) | (0.0062) | (0.0085) | (0.0087) | (0.0101) | (0.0090) |
| 10–12 | 0.0128\* | 0.0109\* | 0.0114+ | 0.0151\* | 0.0145+ | 0.0162\* |
|  | (0.0058) | (0.0050) | (0.0066) | (0.0069) | (0.0081) | (0.0068) |
| 13–15 | 0.0116\* | 0.0077+ | 0.0098+ | 0.0107+ | 0.0081 | 0.0150\*\* |
|  | (0.0047) | (0.0041) | (0.0054) | (0.0055) | (0.0065) | (0.0054) |
| 16–18 | 0.0035 | 0.0011 | 0.0036 | 0.0057 | 0.0068 | 0.0058 |
|  | (0.0036) | (0.0030) | (0.0042) | (0.0042) | (0.0049) | (0.0043) |
| Match requirements: |  |  |  |  |  |  |
| Unique within 3 years? | YES | NO | YES | YES | YES | YES |
| Unique within 5 years? | NO | NO | YES | NO | YES | NO |
| Difference in birth year ≤ 1? | NO | NO | NO | YES | YES | NO |
| Match on exact names? | NO | NO | NO | NO | NO | YES |
| Match Rate | 18.5826 | 26.5591 | 15.1410 | 14.8900 | 11.8161 | 15.3941 |
| Observations | 170,839 | 254,015 | 136,112 | 129,810 | 100,120 | 133,857 |
| R-squared | 0.0566 | 0.0412 | 0.0672 | 0.0622 | 0.0753 | 0.0617 |
| *Notes*: The dependent variable is an indicator for reporting completion of eight or more years of schooling in the 1940 census. Standard errors adjusted for clustering by childhood county of residence are in parentheses. All specifications include year of birth fixed effects, childhood county of residence fixed effects, and controls for family background. Family background controls include childhood household head's occupational score, homeownership status, and literacy, as well as indicators for urban location and farm residence. Columns (1) and (7) repeat the baseline results shown in Table 4 for ease of comparison. +, \*, \*\*, and \*\*\* indicate significance at the 10, 5, 1, and 0.1 percent levels, respectively. | | | | | | |

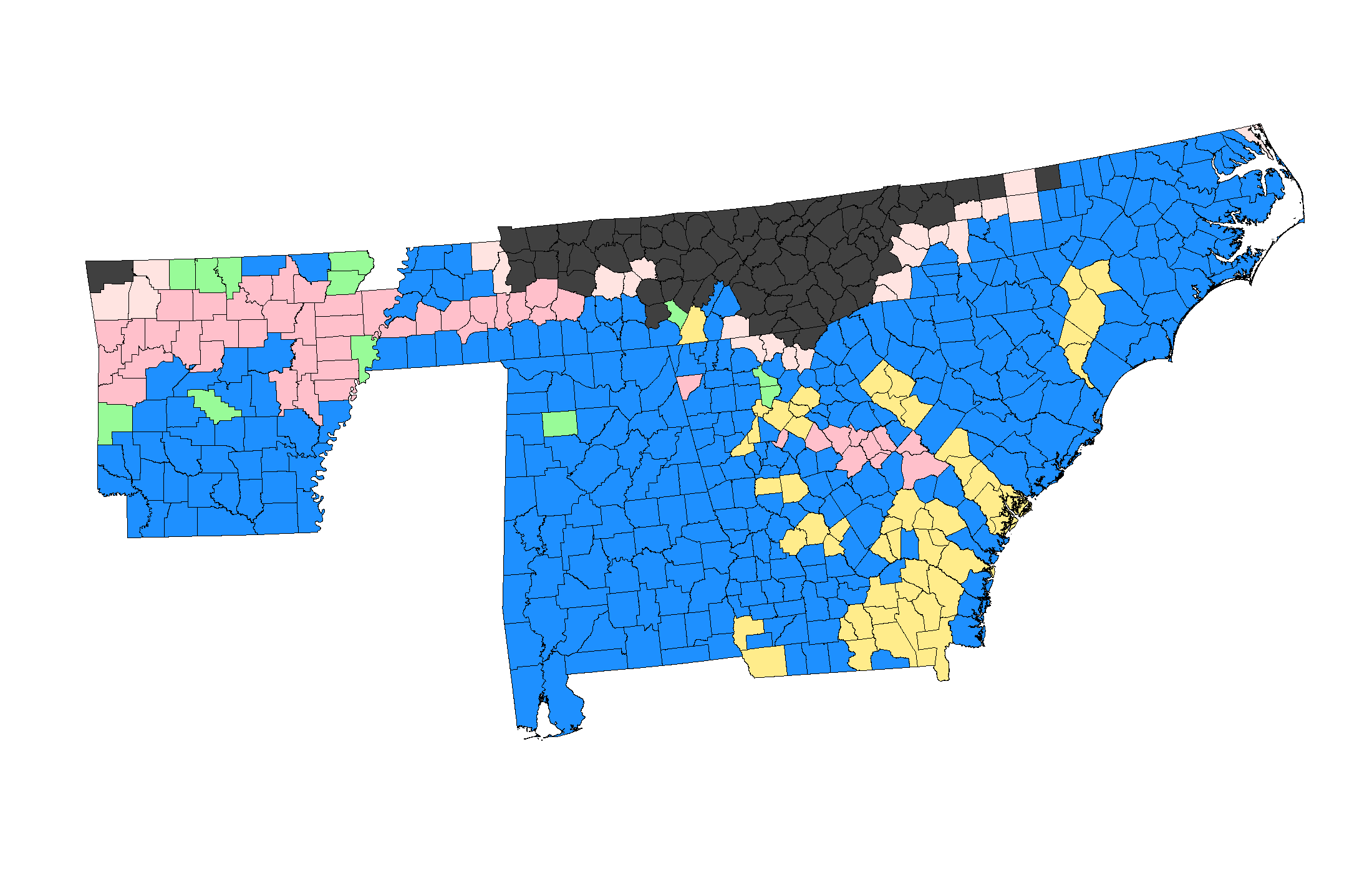


Figure A.1

Sample for Contemporaneous Schooling Outcomes: Enrollment and Attendance

*Notes*: The figure displays all counties in Alabama, Arkansas, Georgia, North Carolina, South Carolina, and Tennessee according to their 1920 borders. The 327 counties shaded in blue make up our regression sample for the contemporaneous outcomes of enrollment and attendance. Other colors represent counties excluded from the sample due to various restrictions. Those shaded in charcoal were not in the Cotton Belt; light pink either not infested or only partially infested by the boll weevil in 1922; pink experienced a full retreat of the boll weevil before being later reinfested; green have little to no black children; and yellow experienced significant border changes during the relevant period.



Figure A.2

Sample for Long-Run Schooling Outcomes: Years of Schooling and Grade Completion

*Notes*: The figure displays all counties, according to their 1920 borders, in Southern states with territory in the Cotton Belt (Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, and Virginia). The counties shaded in blue represent the childhood locations of individuals included in our regression sample for the long-run outcomes of years of schooling and grade completion. Other colors represent counties excluded from the sample due to various restrictions. Those shaded in charcoal were not in the Cotton Belt; light pink either not infested or only partially infested by the boll weevil in 1922; pink experienced a full retreat of the boll weevil before being later reinfested; violet the difference between *first arrival year* and *complete infestation year* is greater than four years; and green the *year of infestation* was prior to 1900.

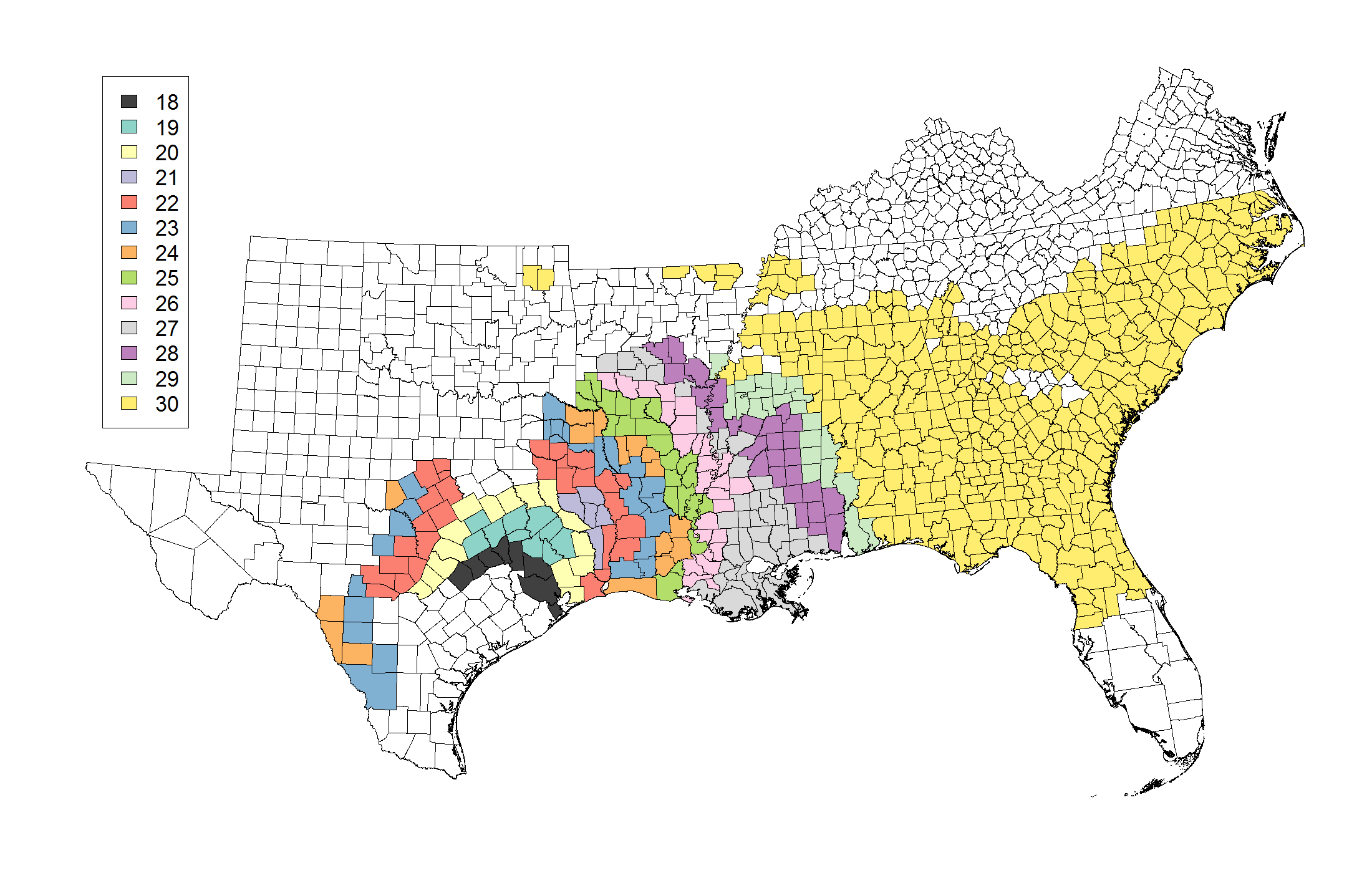


Figure A.3

Oldest Age at Exposure Cohort by County for Long-Run Outcomes Sample

*Notes*: The figure displays all counties, according to their 1920 borders, in Southern states with territory in the Cotton Belt (Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, and Virginia). The shaded counties represent the childhood locations of individuals included in our regression sample for the long-run outcomes of years of schooling and grade completion. The fill color for each county indicates the oldest age at exposure cohort included in our sample from that location. Since we match those ages 3–18 in each census year, beginning with 1900, the oldest age at exposure cohort varies by county from 18 in counties where the boll weevil arrived in 1900 to 30 in counties where the boll weevil arrived after 1911.

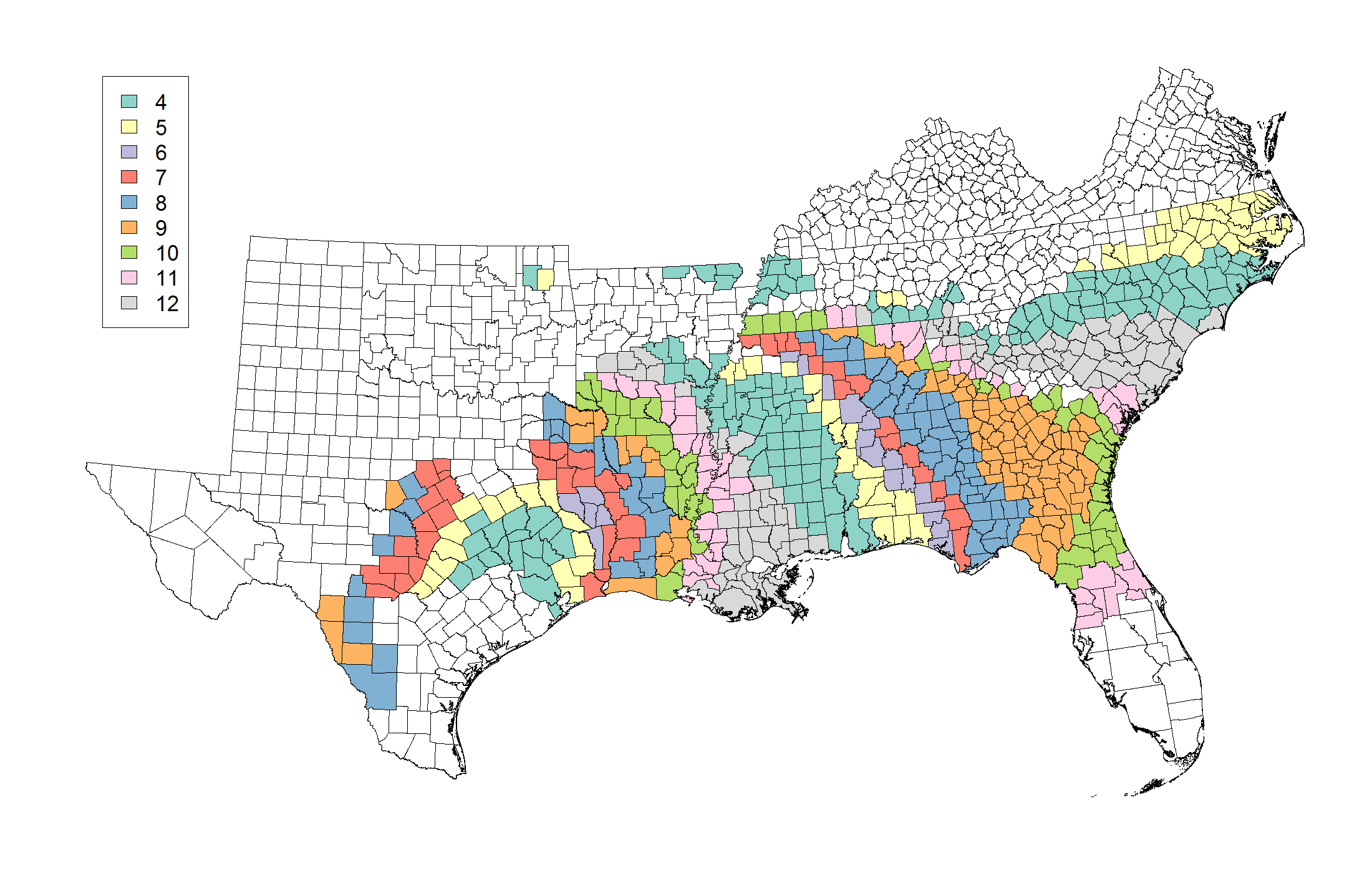
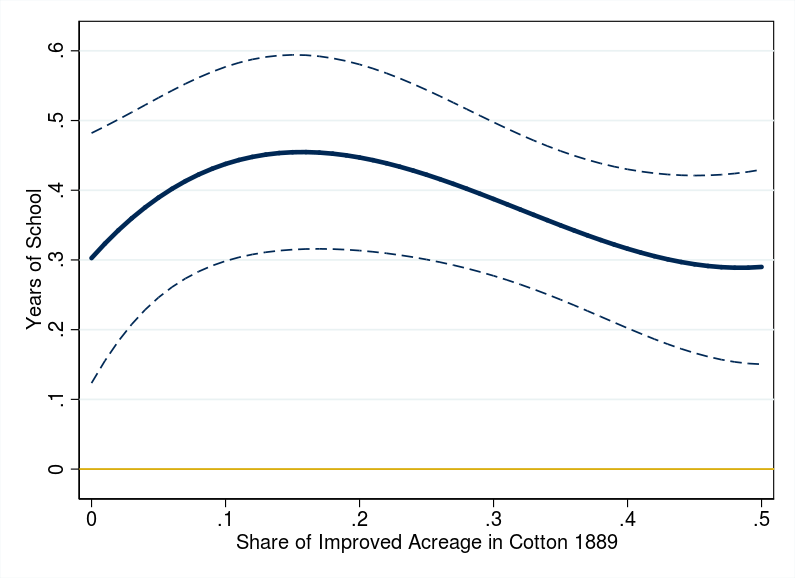
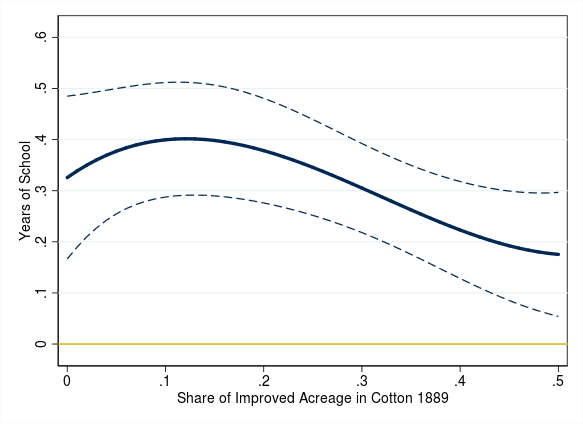


Figure A.4

Youngest Age at Exposure Cohort by County for Long-Run Outcomes Sample

*Notes*: The figure displays all counties, according to their 1920 borders, in Southern states with territory in the Cotton Belt (Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, and Virginia). The shaded counties represent the childhood locations of individuals included in our regression sample for the long-run outcomes of years of schooling and grade completion. The fill color for each county indicates the youngest age at exposure cohort included in our sample from that location. Since we match those ages 3–18 in each census year to 1940 and we restrict the sample to those for whom we observe childhood location prior to the boll weevil's arrival, the youngest age at exposure cohort varies by county from 3 in counties where the boll weevil arrived in a census year to 12 in counties where the boll weevil arrived in the year prior to a census year.

a. Children exposed at 4–6 b. Children exposed at 7–9

c. Children exposed at 10–12 d. Children exposed at 13–15

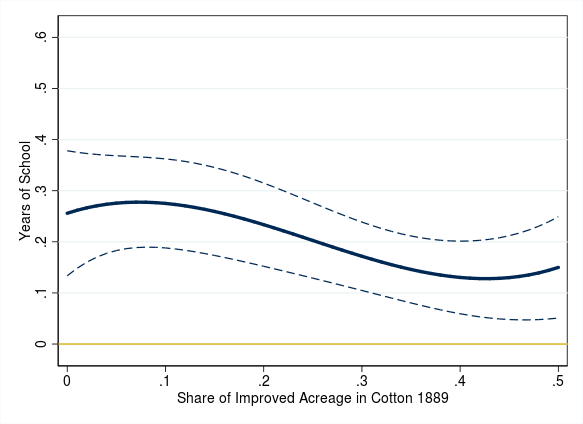
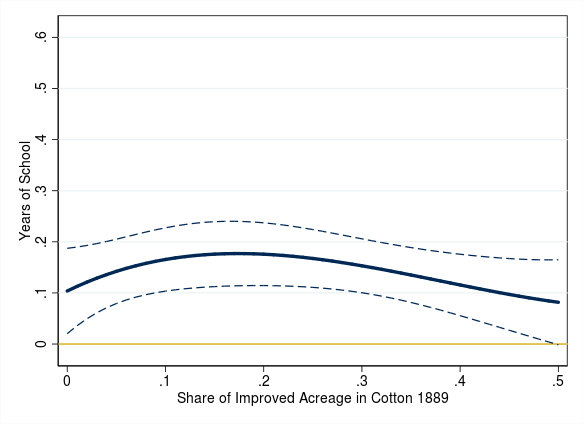
 

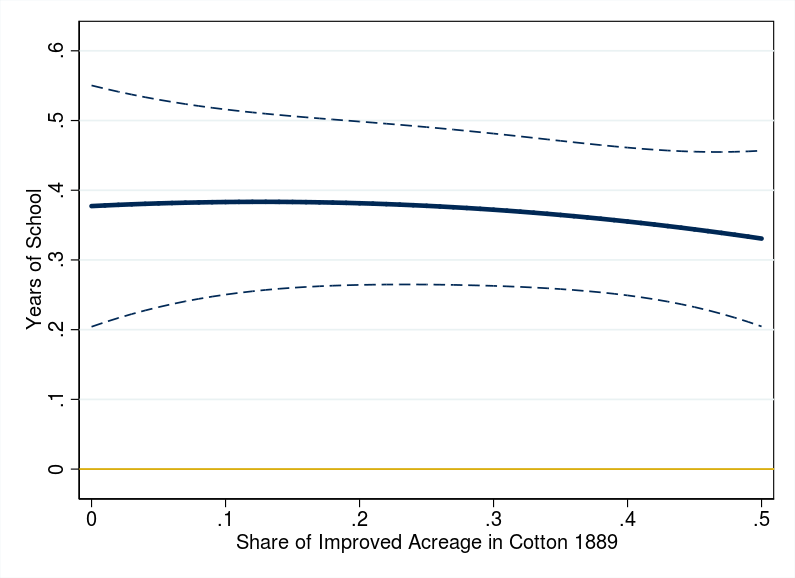
Figure A.5

Impact of the Boll Weevil on Years of Schooling with Third Order Polynomial in Intensity of Cotton Farming in 1889

*Notes*: The y-axis shows the difference in years of schooling relative to individuals exposed to the boll weevil between the ages of 19 and 30, inclusive. The dashed lines indicate 95 percent confidence intervals. Adjusted 0.2038.

*Sources*: See the text.

a. Children exposed at 4–6 b. Children exposed at 7–9

c. Children exposed at 10–12 d. Children exposed at 13–15

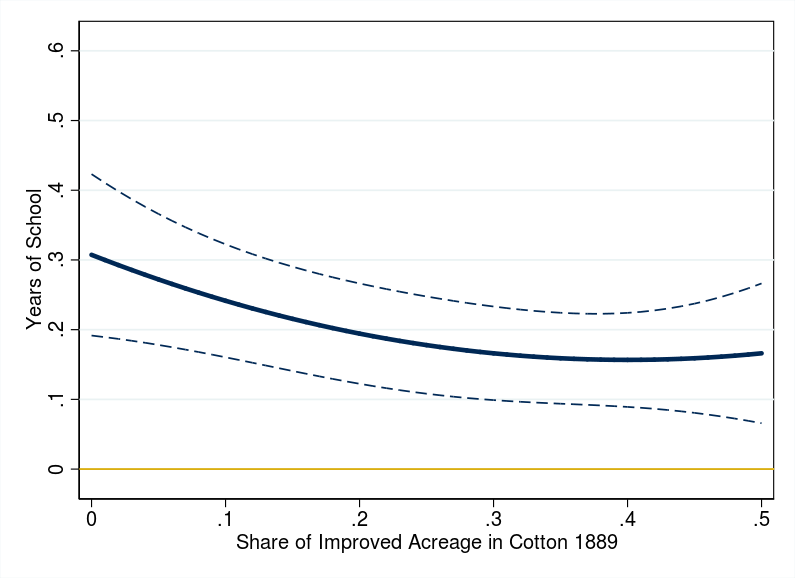
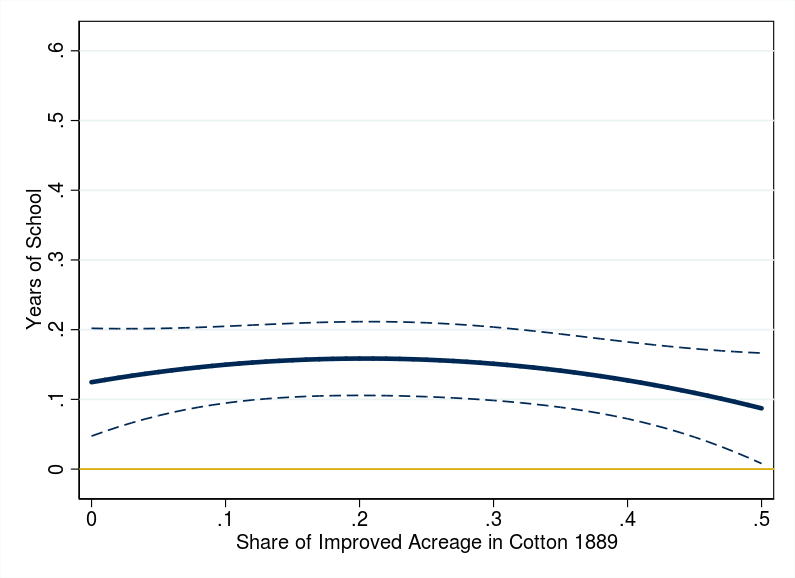
 

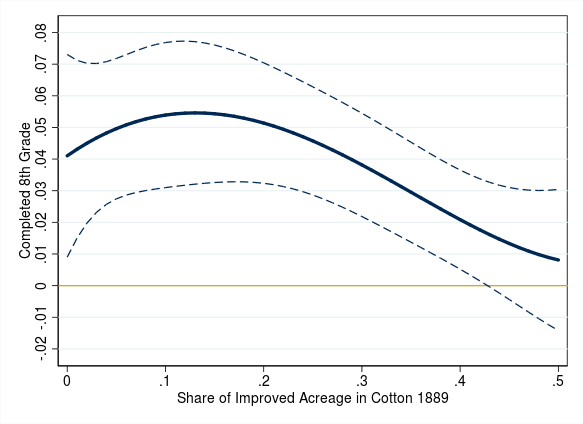
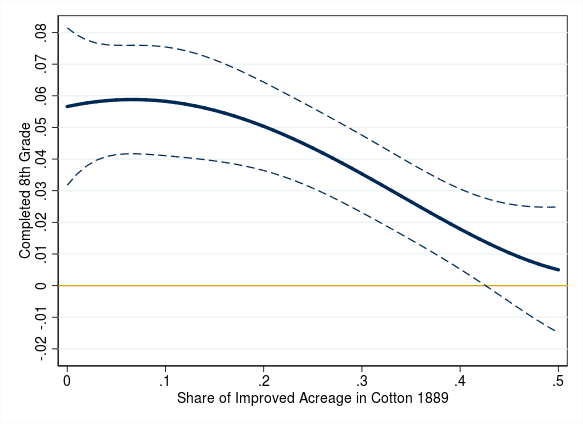
Figure A.6

Impact of the Boll Weevil on Years of Schooling with Second Order Polynomial in Intensity of Cotton Farming in 1889

*Notes*: The y-axis shows the difference in years of schooling relative to individuals exposed to the boll weevil between the ages of 19 and 30, inclusive. The dashed lines indicate 95 percent confidence intervals. Adjusted 0.2038.

*Sources*: See the text.

a. Children exposed at 4–6 b. Children exposed at 7–9

c. Children exposed at 10–12 d. Children exposed at 13–15

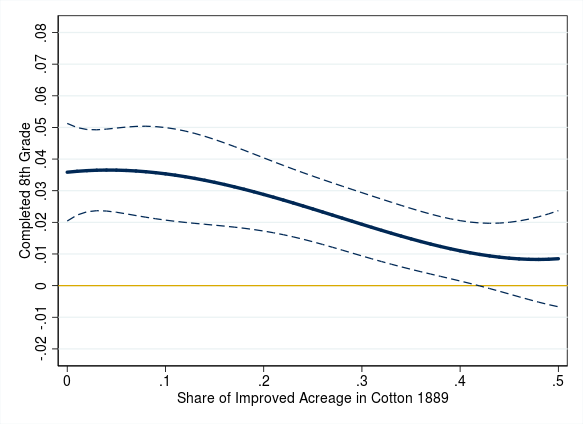
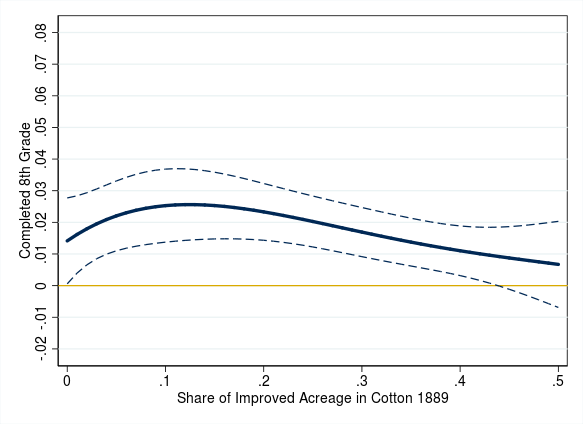
 

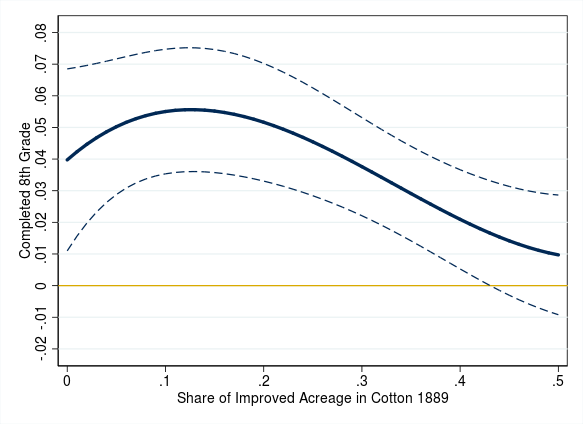
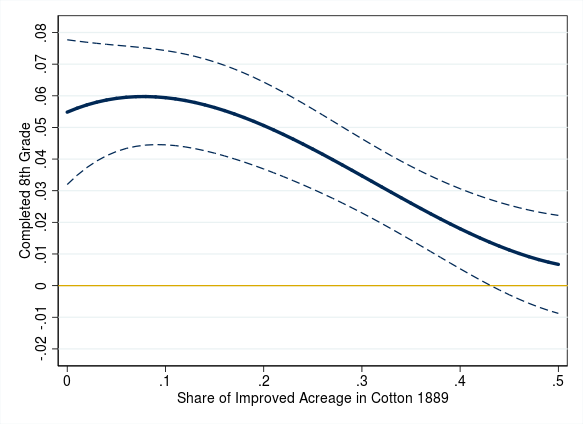
Figure A.7

Impact of the Boll Weevil on Eighth Grade Completion with Fourth Order Polynomial in Intensity of Cotton Farming in 1889

*Notes*: The y-axis shows the difference in likelihood of completing eighth grade relative to individuals exposed to the boll weevil between the ages of 19 and 30, inclusive. The dashed lines indicate 95 percent confidence intervals. Adjusted 0.1484.

*Sources*: See the text.

a. Children exposed at 4–6 b. Children exposed at 7–9

c. Children exposed at 10–12 d. Children exposed at 13–15

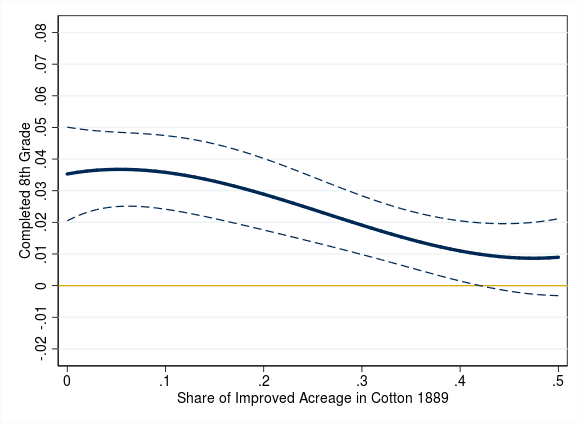
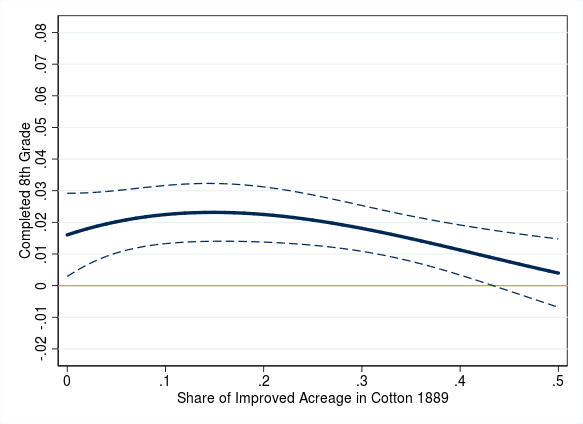
 

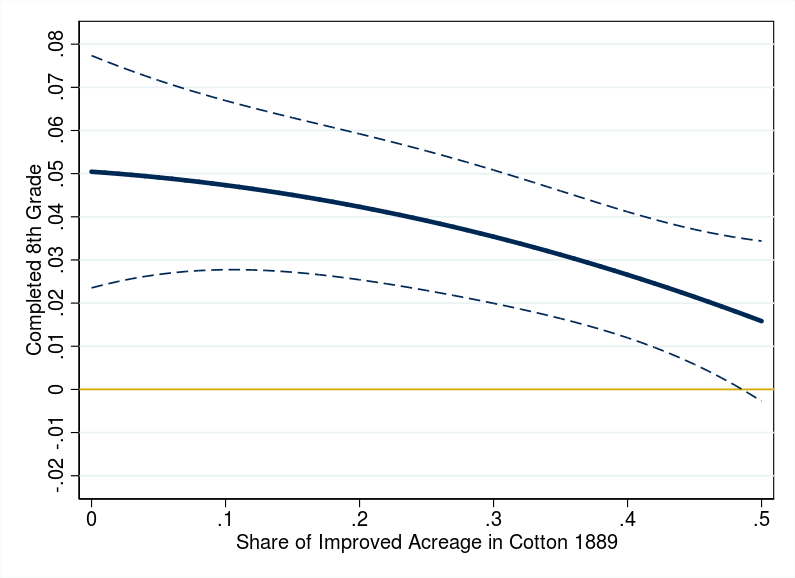
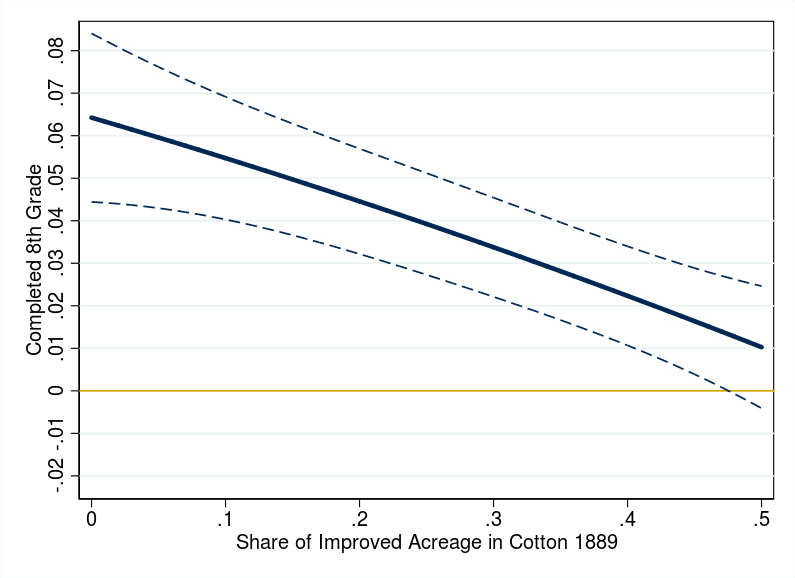
Figure A.8

Impact of the Boll Weevil on Eighth Grade Completion with Third Order Polynomial in Intensity of Cotton Farming in 1889

*Notes*: The y-axis shows the difference in likelihood of completing eighth grade relative to individuals exposed to the boll weevil between the ages of 19 and 30, inclusive. The dashed lines indicate 95 percent confidence intervals. Adjusted 0.1484.

*Sources*: See the text.

a. Children exposed at 4–6 b. Children exposed at 7–9

c. Children exposed at 10–12 d. Children exposed at 13–15

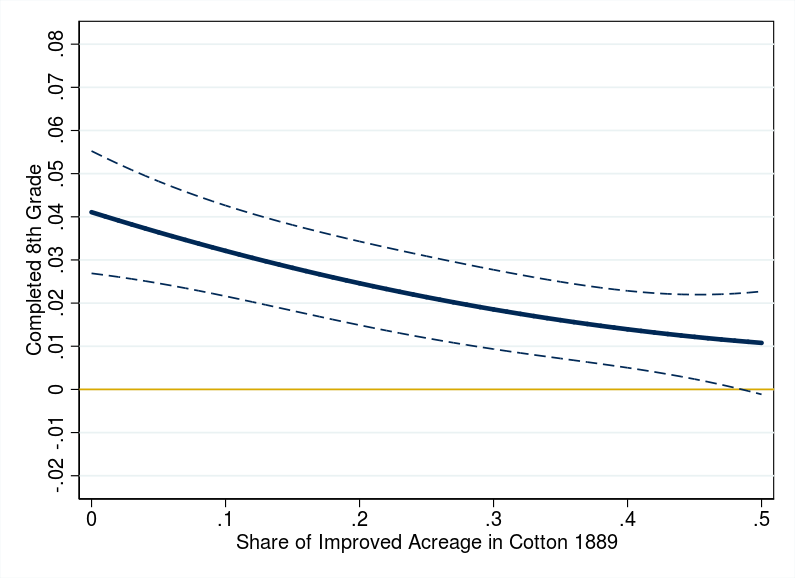
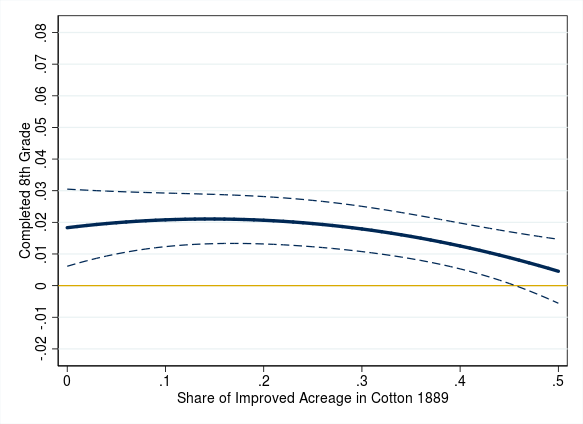
 

Figure A.9

Impact of the Boll Weevil on Eighth Grade Completion with Second Order Polynomial in Intensity of Cotton Farming in 1889

*Notes*: The y-axis shows the difference in likelihood of completing eighth grade relative to individuals exposed to the boll weevil between the ages of 19 and 30, inclusive. The dashed lines indicate 95 percent confidence intervals. Adjusted 0.1484.

*Sources*: See the text.

1. See Abramitzky et al. (2018) for a description of the small differences between the matching method used in this paper and the original method described by Abramitzky, Boustan, and Eriksson (2012). Specifically, we use the abematch command provided at https://ranabr.people.stanford.edu/matching-codes. [↑](#footnote-ref-1)