*Online Appendix A: Supplementary Tables and Additional Results*

Appendix Table 1

SUMMARY STATISTICS

|  |  |  |
| --- | --- | --- |
|  | Mean | Standard Deviation |
| Percent of public housing units 1970 | 0.8085 | 1.246 |
| Percent of public housing units 1960 | 0.3230 | 0.7960 |
| Percent of public housing units 1950 | 0.07604 | 0.3758 |
| Ln median property value 1970 | 9.275 | 0.3580 |
| Ln median family income 1970 | 8.888 | 0.2530 |
| Percent of families with < $3,000 income 1970 | 16.67 | 8.396 |
| Percent of female-headed households 1970 | 8.999 | 3.201 |
| Ln density 1970 | 3.487 | 1.499 |
| Percent of high school grads 1970 | 44.64 | 12.54 |
| Percent of lf in manufacturing 1940 | 10.64 | 10.54 |
| Percent of lf in agriculture 1940 | 23.31 | 13.18 |
| Percent pop urban 1940 | 23.08 | 24.55 |
| Percent units owner occupied 1940 | 49.94 | 11.67 |
| Median persons per rental unit 1940 | 3.476 | 0.4067 |
| Percent units good 1940 | 68.22 | 12.70 |
| Percent units electricity 1940 | 53.64 | 24.22 |
| Percent units water 1940 | 40.74 | 24.22 |
| Ln median property value 1940 | 7.225 | 0.5807 |
| Percent of high school graduates 1940 | 20.33 | 7.701 |
| Percent of college graduates 1940 | 3.295 | 1.648 |
| Ln density 1940 | 3.377 | 1.311 |
| Percent black 1940 | 10.66 | 17.78 |
| Percent votes for Roosevelt 1940 | 60.50 | 20.25 |
| Percent pop Baptist 1950 | 10.15 | 11.49 |
| Percent pop Catholic 1950 | 11.11 | 15.73 |
| Total major war supply contracts ($000s) 1940–1945 per capita | 0.3925 | 1.800 |
| Total major war facilities projects ($000s) 1940–1945 per capita | 0.1750 | 0.7347 |
|  |  |
| *N* |  2,973 |

*Sources*: County population and housing data are from Haines (2004). Data on 1940 presidential election results are from Leip (2009). Public housing data are from the *Consolidated Development Directory* (HUD 1973).

Appendix Table 2

RESULTS USING 1950 CONTROLS, COMPARABLE TO TABLE 1 IN MAIN TEXT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Ln Median Property Value | Ln Median Family Income | Percent of Families with < $3,000 Income | Ln Population Density |
| *Panel A: Basic specification, with state-fixed effects* |
| 1. Entire sample (*N* = 2,970) | –0.01552\*\*\* (0.003749) | –0.01314\*\*\* (0.001352) | 0.3566\*\*\* (0.05425) | –0.03066\*\*\* (0.004012) |
| *R2* | 0.77 | 0.83 | 0.82 | 0.98 |
| 2. Counties with public housing in 1970 (*N* = 1,377) | –0.01618\*\*\*(0.005305) | –0.01654\*\*\*(0.002534) | 0.4731\*\*\*(0.09346) | –0.03784\*\*\*(0.00559) |
| *R2* | 0.82 | 0.86 | 0.84 | 0.98 |
| *Panel B: Basic specification, with state economic area fixed effects* |
| 3. Entire sample (*N* = 2,970) | –0.009944\*\*\* (0.003234) | –0.008953\*\*\* (0.001538) | 0.2708\*\*\* (0.05391) | –0.02018\*\*\* (0.005087) |
| *R2* | 0.84 | 0.88 | 0.87 | 0.99 |
| 4. Counties with public housing in 1970 (*N* = 1,377) | –0.01303\*\*(0.006313) | –0.01215\*\*\*(0.001983) | 0.3714\*\*\*(0.08970) | –0.02952\*\*\*(0.005394) |
| *R2* | 0.89 | 0.92 | 0.90 | 0.99 |
| *Panel C: Additional controls for pre-1940 trends, with state-fixed effects* |
| 5. Entire sample (*N* = 2,336) | –0.01207\*\*\* (0.003528) | –0.01195\*\*\* (0.001543) | 0.3148\*\*\* (0.05870) | –0.02538\*\*\* (0.005362) |
| *R2* | 0.80 | 0.85 | 0.84 | 0.98 |
| 6. Entire sample without trends, for comparison (*N* = 2,336) | –0.1484\*\*\*(0.004105) | –0.01293\*\*\*(0.001586) | 0.3462\*\*\*(0.06161) | –0.03000\*\*\*(0.005254) |
| *R2* | 0.79 | 0.85 | 0.84 | 0.98 |
| 7. Counties with public housing in 1970 (*N* = 1,188) | –0.01531\*\*\*(0.005057) | –0.01486\*\*\*(0.002836) | 0.4292\*\*\*(0.1024) | –0.03057\*\*\*(0.007750) |
| *R2* | 0.83 | 0.88 | 0.87 | 0.98 |
| 8. Counties with public housing in 1970 without trends, for comparison (*N* = 1,188) | –0.01765\*\*\*(0.005714) | –0.01552\*\*\*(0.003090) | 0.4548\*\*\*(0.1112) | –0.03535\*\*\*(0.007351) |
| *R2* | 0.82 | 0.87 | 0.86 | 0.98 |
| *Panel D: No pre-program controls* |
| 9. Entire sample (*N* = 2,970) | 0.008205(0.01330) | 0.004770(0.008867) | -0.1149(0.3144) | 0.2108\*\*\*(0.04210) |
| *R2* | 0.34 | 0.44 | 0.45 | 0.46 |
| *Ratio (compare rows 1 and 9)* | –.65 | –.73 | –.76 | –.13 |
| 10. Counties with public housing in 1970 (*N* = 1,377) | –0.02161(0.01502) | –0.02372\*\*(0.009586) | 0.6957\*\*(0.3287) | 0.004900(0.04339) |
| *R2* | 0.42 | 0.48 | 0.47 | 0.40 |
| *Ratio(compare rows 2 and 10)* | 3.0 | 2.3 | 2.1 | –0.89 |

\* = Significant at the 10 percent level.

\*\* = Significant at the 5 percent level.

\*\*\* = Significant at the 1 percent level.

Appendix Table 2 — continued

*Notes*: Each coefficient is from a separate OLS regression of an economic outcome on
public housing intensity and control variables, where public housing intensity is defined as
(public housing units in 1970 / total occupied units in 1970 \* 100). Preprogram controls included in Panels A-C include: 1950 log population density, percent urban, percent nonwhite, percent nonwhite squared, percent of dwelling units with running water and toilet (not dilapidated), percent of units owner-occupied, median persons per occupied unit, log median property value, percent of high school graduates, percent of college graduates, percent of labor force in manufacturing, percent of labor force in agriculture, percent of families with less than $2,000 income, log median family income, percent Baptists, and percent Catholics; percent of votes for Roosevelt in 1940 and per capita total war contracts and total war facilities 1940–1945. Panel C includes additional controls for the percent change in total population, black population, and number of dwelling units between 1900–1910, 1910–1920, 1920–1930, and 1930–1940. State- fixed effects are included in panels A, C, and D. State economic area fixed effects are included
in panel B. In all panels, standard errors are reported in parentheses and are clustered by state. Ratios reported in panel D are calculated as *βC* / (*βNC* – *βC*), where *βC* is the estimated coefficient
of the percent of public housing units in 1970 in a regression with controls (panel A) and *βNC* is the estimated coefficient of the percent of public housing units in 1970 in a regression with no controls (panel D).

*Sources*: County population and housing data are from Haines (2004). Data on 1940 presidential election results are from Leip (2009). Public housing data are from the *Consolidated Development Directory* (HUD 1973).

Appendix Table 3

RESULTS USING 1970–1950 DIFFERENCES AS THE DEPENDENT VARIABLE AND 1950 CONTROLS, COMPARABLE TO TABLE 1 IN MAIN TEXT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Ln Median Property Value | Ln Median Family Income | Percent of Families with < $3,000 Income | Ln Population Density |
| *Panel A: Basic specification, with state-fixed effects* |
| 1. Entire sample (*N* = 2,970) | –0.01248\*\*\* (0.003784) | –0.01644\*\*\* (0.001550) | 0.1589\* (0.08554) | –0.03035\*\*\* (0.003921) |
| *R2* | 0.43 | 0.78 | 0.89 | 0.57 |
| 2. Counties with public housing in 1970 (*N* = 1,377) | –0.01184\*\*(0.005342) | –0.02167\*\*\*(0.003059) | 0.1919(0.1157) | –0.03769\*\*\*(0.005619) |
| *R2* | 0.43 | 0.81 | 0.91 | 0.65 |
| *Panel B: Basic specification, with state economic area fixed effects* |
| 3. Entire sample (*N* =2,970) | –0.008009\*\* (0.003828) | –0.01192\*\*\* (0.001837) | 0.09039(0.08977) | –0.02348\*\*\* (0.005532) |
| *R2* | 0.59 | 0.85 | 0.92 | 0.72 |
| 4. Counties with public housing in 1970 (*N*=1377) | –0.01262(0.007685) | –0.01546\*\*\*(0.002659) | 0.1783(0.1287) | –0.03074\*\*\*(0.005298) |
| *R2* | 0.63 | 0.88 | 0.94 | 0.81 |
| *Panel C: Additional controls for pre-1940 trends, with state-fixed effects* |
| 5. Entire sample (*N* = 2,336) | –0.006759\*\* (0.003265) | –0.01492\*\*\* (0.001898) | 0.1266(0.08105) | –0.02483\*\*\* (0.005173) |
| *R2* | 0.44 | 0.78 | 0.90 | 0.61 |
| 6. Entire sample without trends, for comparison (*N* = 2,336) | –0.009234\*\*(0.003770) | –0.01577\*\*\*(0.001911) | 0.1693\*(0.08771) | –0.02919\*\*\*(0.005074) |
| *R2* | 0.41 | 0.77 | 0.89 | 0.59 |
| 7. Counties with public housing in 1970 (*N* = 1,188) | –0.007964(0.005085) | –0.01994\*\*\*(0.003140) | 0.1682(0.1164) | –0.03086\*\*\*(0.007614) |
| *R2* | 0.42 | 0.81 | 0.92 | 0.66 |
| 8. Counties with public housing in 1970 without trends, for comparison (*N* = 1,188) | –0.009834\*(0.005613) | –0.02079\*\*\*(0.003457) | 0.1786(0.1247) | –0.03537\*\*\*(0.007286) |
| *R2* | 0.39 | 0.80 | 0.92 | 0.63 |
| *Panel D: No preprogram controls* |
| 9. Entire sample (*N* = 2,970) | –0.02324\*\*\*(0.003857) | –0.02320\*\*\*(0.005505) | 1.0828\*\*\*(0.3598) | –0.01984\*\*(0.009203) |
| *R2* | 0.29 | 0.54 | 0.56 | 0.24 |
| *Ratio (compare rows 1 and 9)* | 1.2 | 2.4 | 0.17 | –2.9 |
| 10. Counties with public housing in 1970 (*N* = 1,377) | –0.01279\*\*\*(0.004665) | –0.02004\*\*\*(0.004017) | 0.3118(0.2880) | –0.05522\*\*\*(0.01116) |
| *R2* | 0.31 | 0.54 | 0.61 | 0.32 |
| *Ratio(compare rows 2 and 10)* | 12.5 | –13.3 | 1.6 | 2.2 |

\* = Significant at the 10 percent level.

\*\* = Significant at the 5 percent level.

\*\*\* = Significant at the 1 percent level.

Appendix Table 3 — continued

*Notes*: Each coefficient is from a separate OLS regression of an economic outcome on public housing intensity and control variables, where public housing intensity is defined as (public housing units in 1970 / total occupied units in 1970 \* 100). The dependent variables are 1970–1950 differences. See Appendix Table 2 for a complete list of 1950 controls. State-fixed effects are included in panels A, C, and D. State economic area fixed effects are included in panel B. In all panels, standard errors are reported in parentheses and are clustered by state. Ratios reported in panel D are calculated as *βC* / (*βNC* – *βC*), where *βC* is the estimated coefficient of the percent of public housing units in 1970 in a regression with controls (panel A) and *βNC* is the estimated coefficient of the percent of public housing units in 1970 in a regression with no controls (panel D).

*Sources*: See the text.

Appendix Table 4

RESULTS EXCLUDING COUNTIES IN THE 95TH PERCENTILE OF PUBLIC HOUSING INTENSITY

(comparable to panel A of Table 1 in the main text)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Ln Median Property Value | Ln Median Family Income | Percent of Families with < $3,000 Income | Ln Population Density |
| 1. Entire sample (*N* = 2,973) | –0.01618\*\*\*(0.004584) | –0.01626\*\*\*(0.002695) | 0.4241\*\*\*(0.1032) | –0.03585\*\*\*(0.006088) |
|  |  |  |  |  |
| 2. All counties with PH < 4.25 (*N* = 2,906) | –0.02494\*\*\*(0.005242) | –0.01868\*\*\*(0.003808) | 0.4335\*\*\*(0.1366) | –0.04481\*\*\*(0.008322) |
|  |  |  |  |  |
| 3. All counties with PH in 1970 (*N* = 1,371) | –0.01655\*\*\*(0.005322) | –0.01769\*\*\*(0.002649) | 0.4549\*\*\*(0.1130) | –0.04602\*\*\*(0.007224) |
|  |  |  |  |  |
| 4. All counties with 0 < PH < 4.25 in 1970 (*N* = 1,304) | –0.03426\*\*\*(0.006309) | –0.02469\*\*\*(0.004631) | 0.5529\*\*\*(0.1887) | –0.07074\*\*\*(0.01165) |

\* = Significant at the 10 percent level.

\*\* = Significant at the 5 percent level.

\*\*\* = Significant at the 1 percent level.

*Notes*: Notes: Each coefficient is from a separate OLS regression of an economic outcome
on public housing intensity and control variables, where public housing intensity is defined as (public housing units in 1970 / total occupied units in 1970 \* 100). Rows 1 and 3 replicate rows 1 and 2 from Table 1, Panel A for comparison. See notes to Table 1 or the text for a list of
the independent variables. State-fixed effects are included. Standard errors are reported in parentheses and are clustered by state.

*Sources*: See the text.

Appendix Table 5

REGRESSIONS WITH AND WITHOUT PRE-PROGRAM CONTROLS, SEPARATELY BY URBAN STATUS

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Ln Median Property Value | Ln Median Family Income | Percent of Families with < $3,000 Income | Ln Population Density |
| *Panel A: Nonurban counties* |  |
| 1. No preprogram controls (*N* = 1,747) | –0.01586(0.009961) | –0.01688\*\*(0.007390) | 0.5883\*\*(0.2638) | 0.04562(0.02744) |
| 2. Preprogram controls (*N* = 1,747) | –0.01495\*\*\*(0.004947) | –0.01689\*\*\*(0.004209) | 0.4889\*\*\*(0.1347) | –0.03650\*\*\*(0.006204) |
|  |  |  |  |  |
| Ratio | 16.4 1689 | 4.9 | –0.44 |
| *Only counties with PH in 1970* |
| 3. No preprogram controls (*N* = 646) | –0.01563\*(0.008307) | –0.02439\*\*\*(0.006913) | 0.7176\*\*\*(0.2630) | –0.07792\*\*\*(0.02481) |
| 4. Preprogram controls (*N* = 646) | –0.009781\*(0.005590) | –0.01638\*\*\*(0.002849) | 0.4488\*\*\*(0.1145) | –0.03706\*\*\*(0.005794) |
| Ratio | 1.7 | 2.0 | 1.7 | 0.91 |
| *Panel B: Urban counties* |
| 5. No preprogram controls (*N* = 1,226) | –0.02014(0.01206) | –0.01834\*\*(0.006900) | 0.6661\*\*\*(0.2198) | 0.1918\*\*\*(0.05972) |
| 6. Preprogram controls (*N* = 1,226) | –0.01424\*\*(0.005493) | –0.01741\*\*\*(0.002735) | 0.4569\*\*\*(0.09545) | –0.02409\*\*(0.01174) |
| Ratio | 2.4 | 18.7 | 2.2 | –0.1 |
| *Only counties with PH in 1970* |
| 7. No preprogram controls (*N* = 725) | –0.04911\*\*\*(0.01425) | –0.04591\*\*\*(0.01058) | 1.474\*\*\*(0.3284) | 0.01692(0.06906) |
| 8. Preprogram controls (*N* = 725) | –0.02360\*\*\*(0.007453) | –0.02522\*\*\*(0.005236) | 0.7635\*\*\*(0.1734) | –0.04966\*\*\*(0.01556) |
| Ratio | 0.93 | 1.2 | 1.07 | –0.75 |

Appendix Table 5 — continued

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Ln Median Property Value | Ln Median Family Income | Percent of Families with < $3,000 Income | Ln Population Density |
| *Panel C: Counties in SMAs in 1950* |
| 9. No preprogram controls (*N* = 266) | –0.03638\*\*(0.01523) | –0.04770\*\*\*(0.006963) | 1.296\*\*\*(0.2043) | 0.3063\*\*\*(0.07416) |
| 10. Preprogram controls (*N* = 266) | –0.01595(0.01147) | –0.02612\*\*\*(0.005427) | 0.7876\*\*\*(0.1731) | –0.08251\*\*\*(0.02885) |
| Ratio | 0.78 | 1.2 | 1.5 | –0.21 |
| *Only counties with PH in 1970* |
| 11. No preprogram controls (*N* = 230) | –0.02914(0.01817) | –0.04761\*\*\*(0.008816) | 1.283\*\*\*(0.2414) | 0.3089\*\*\*(0.08803) |
| 12. Preprogram controls (*N* = 230) | –0.01393(0.01136) | –0.02414\*\*\*(0.005288) | 0.7151\*\*\*(0.1582) | –0.08731\*\*\*(0.02559) |
| Ratio | 0.92 | 1.0 | 1.3 | –0.22 |

\* = Significant at the 10 percent level.

\*\* = Significant at the 5 percent level.

\*\*\* = Significant at the 1 percent level.

*Notes*: Each coefficient is from a separate OLS regression of an economic outcome on
public housing intensity and control variables, where public housing intensity is defined as (public housing units in 1970 / total occupied units in 1970 \* 100). “Nonurban” (“Urban”) is defined as having less (more) than 25 percent urban population in 1940. See notes to Table 1 or the text for a list of the independent variables. State fixed effects are included. Standard errors are reported in parentheses and are clustered by state. Ratios are calculated as *βC* / (*βNC* – *βC*), where *βC* is the estimated coefficient of the percent of public housing units in 1970 in a regression with controls and β*NC* is the estimated coefficient of the percent of public housing units in 1970 in a regression with no controls. Rows 2, 4, 6, 8, 10, and 12 replicate results reported in Table 2 of the text, for comparison.

*Sources*: See the text.

*Online Appendix B: Data Sources and Merging*

*Data Sources*

 Public housing data comes from the *Consolidated Development Directory (CDD)*, published by the U.S. Department of Housing and Urban Development (1973).
The *CDD* contains project-level information on the state and locality (generally
this is the name of a city, although it is occasionally reported at the county level);
project number; type of program at the latest stage of construction (e.g., Acquisition, Conventional new construction, Leased housing, Turnkey, PWA, Housing Act of 1937, National Defense housing (converted for low-income use), etc.);
total and elderly units planned; total and elderly units completed; dates of initial and full occupancy. Housing built under the Indian Housing Program is included in the *CDD*, but excluded from the analysis. I assign each project to its county and aggregate the data in 1940, 1950, 1960, and 1970 by summing the total units of low-income housing in projects with full occupancy on or before those years.

 The data on population density, median family income, median property values,
and the percent of low-income families are from the 1947, 1952, 1962, and 1972 *County Data Book* files compiled by Haines (2004, file numbers 70, 72, 74, and 76). These files also include data on the percent of owner-occupied housing units in 1940 and 1970, median persons per room in rental units, percent of units in good condition, the percent of units with electricity, percent of units with water, percent urban, percent black, percent of the labor force in agriculture and manufacturing in 1940, the value
of World War II contracts and facilities projects between 1940–1945, Standard Metropolitan Area (SMA) status and State Economic Area (SEA) in 1950, and the percent of high school graduates in 1960 and 1970.[[1]](#footnote-1) Data on adult educational attainment used to construct the percent of high school graduates and the percent of college graduates in 1940 come from the 1940 Census (Haines 2004, file number 32). Data on the percent of Baptists and Catholics comes from the 1952 *Survey of Churches and Church Membership* (Haines 2004, file number 57). Data on the percent of votes of Roosevelt in 1940 come from Leip (2009). Data on the population, black population, and number of dwelling units in 1900, 1910, 1920, and 1930 come from the 1900, 1910, 1920, and 1930 Census (Haines 2004, file numbers 20, 22, 24,
and 26). County-level data on the percent of recent migrants in 1960 and 1970 come from the printed volumes of the Population Census (Characteristics of the Population) of 1960 (table 35) and 1970 (table 43). County-group data on the percent of recent migrants in 1970, the percent of recent migrants with high school degrees, and the percent of stayers with high school degrees were calculated using the 1% 1970 Metro Form 2 Sample downloaded from IPUMS (Ruggles *et al.* 2010). County-group data on 1970 public housing intensity, 1960 log population density, and 1960 percent of high school graduates is aggregated from county-level data.

*Boundary Changes and Data Merging*

 There were several instances in which there were significant county boundary changes between 1940 and 1970. For these cases, I combined counties (or independent cities and counties) to form consistent boundaries. Most of the cases occurred
in Virginia, as a number of cities became independent cities during this period.
A detailed description of combined localities is below.

VIRGINIA

Bedford City into Bedford County (1970): Bedford City became independent from Bedford County in 1968. I combined Bedford City and Bedford County in 1970.

Colonial Heights City into Chesterfield County (1950, 1960, 1970): Colonial Heights City became independent from Chesterfield County in 1948. I combined Colonial Heights City and Chesterfield County in 1950, 1960, and 1970.

Covington City into Alleghany County (1960, 1970): Covington City became independent from Alleghany County in 1952. I merged the two in 1960 and 1970.

Emporia City into Greensville County (1970): Emporia City became independent from Greensville County in 1967. I merged the two in 1970.

Fairfax City into Fairfax County (1970) and Falls Church City into Fairfax County (1950, 1960, 1970): Falls Church City and Fairfax City were both part of Fairfax County in 1940. Falls Church City became independent in 1948 and Fairfax City became independent in 1961. I merged Falls Church City and Fairfax County in 1950 and 1960 and merged Falls Church City, Fairfax City, and Fairfax County in 1970.

Franklin City into Southampton County (1970): Franklin City became independent from Southampton County in 1961. I merged Franklin City and Southampton County in 1970.

Lexington City into Rockbridge County (1970): Lexington City became independent from Rockbridge County in 1965. I merged the two in 1970.

Norton City into Wise County (1960, 1970): Norton City became independent from Wise County in 1954. I merged the two in 1960 and 1970.

Virginia Beach into Princess Anne County (1960, 1970): Virginia Beach became independent from Princess Anne County in 1952. In 1963, Princess Anne County was consolidated into Virginia Beach. I assigned Princess Anne to Virginia Beach in 1940 and 1950 and merged the two in 1960.

Salem City into Roanoke County (1970): Salem City became independent of Roanoke County in 1968. I merged the two in 1970.

South Boston City into Halifax County (1960, 1970): South Boston City became independent from Halifax County in 1960. I merged the two in 1960 and 1970.

Waynesboro City into Augusta County (1970): Waynesboro City became independent of Augusta County in 1948. I merged the two in 1950, 1960, and 1970.

Chesapeake City, Norfolk County, and South Norfolk City (1940, 1950, 1960):
The City of South Norfolk and Norfolk County were consolidated to form Chesapeake City in 1962. Norfolk County and South Norfolk were merged in 1940, 1950, and 1960 and assigned to Chesapeake City.

Newport News City and Warwick County (1940, 1950): Warwick County was consolidated with Newport News in 1958. I merged the two in 1940 and 1950.

Galax City, Carroll County, and Grayson County (1940, 1950, 1960, 1970): Galax City became an independent city in 1954 and was formed from parts of Carroll and Grayson Counties. All three were merged for every period in the data set.

Warwick County and Newport News City (1940, 1950): Warwick County was consolidated into Newport News City in 1958. I merged Warwick County and Newport News in 1940 and 1950.

NEW MEXICO

Los Alamos County, Santa Fe County, Sandoval County (1940, 1950, 1960, 1970): Los Alamos County was formed in 1949 from parts of Santa Fe and Sandoval Counties. All three were merged for every period in the data set.

LARGE CITIES

New York City reports public housing at the city level, and is located in Bronx County, Kings County, New York County, Queens County, and Richmond County.
I merged these 5 counties for my analysis for all years.

Kansas City, MO reports public housing at the city level, and is located in Cass County, Clay County, Jackson County, and Platte County. I merged these four counties for my analysis for all years.

Details on Merging: When consolidating counties, I used the weighted average of
the counties’ values. For example, I weighted median owner-occupied property values by the number of owner-occupied units, median family income and the percent of
low-income families by the number of families, the percent of the population 25 and up with a high school degree by the population 25 years and up, and the percent of owner-occupied units by total occupied units. Data for the variables used as weights come from the *County Data Books* (Haines 2004).

*Online Appendix C: Sensitivity of OLS Results to Omitted Variable Bias*

 I assess the sensitivity of my results to omitted variable bias by adopting
a technique formulated by Joseph G. Altonji, Todd E. Elder, and Christopher R.
Taber (2005). In Altonji, Elder, and Taber (2005), the authors adopt a bivariate normal framework and assess the sensitivity of their results to omitted variables by running regressions with and without observable controls. The basic idea is that if the coefficient of the variable of interest is not sensitive to the inclusion (exclusion) of observable controls, then it is unlikely that the coefficient is sensitive to the inclusion (exclusion) of *unobservable* controls. John Bellows and Edward Miguel (2009) adapt this technique to fit a linear framework, and I use their structure in this article.

 Assume that 1970 community outcomes, *Y*­­1970­­­­, are a function of public housing intensity, *PH*, and an index of community characteristics, *Z*. I would like to estimate the function

*Y*­­1970­­­­ = *α* + *β1PH* + *β2Z* + *ε* (C1)

If I estimate this regression without controlling for *Z*, the estimated *β1* will suffer from omitted variable bias and

plim $\hat{β}$1,*NC* = *β1* + *β2* cov(*PH*,*Z*) / var(*PH*) (C2)

where “*NC*” indicates a regression run with no controls.

 Suppose, however, that *Z* consists of observable and unobservable components, such that

 *Z* = *Zobs.* + *Zunobs*­­­ (C3)

If I estimate

 *Y*­­1970­­­­ = *α* + *β1,CPH* + *β2Zob****s*.** + *ε* (C4)

the probability limit of $\hat{β}$*1,C­­* (where “C” indicates a regression run with controls for observable characteristics in index *Zobs.*) is equal to the true value of *β1* plus the omitted variable bias from the exclusion of *Zunobs*

 plim $\hat{β}$1,*C­­* = *β1* + *β2*cov(*PH*,*Zunobs*) / var(*PH*) (C5)

While I cannot rule out omitted variable bias when I estimate *β1,*C in equation (C4),
I would like to know how strong the covariance between *PH* and Z*unobs* would have
to be in order for omitted variable bias to be able to explain away the entire point estimate. If the true value of *β1* is 0, then the ratio of the covariances of *PH* and the observable and unobservable portions of *Z* can be written as

 plim $\hat{β}$1*,*C / (plim $\hat{β}$*­­*1*,NC* – plim $\hat{β}$1*,*C ) = cov(*PH*,*Zunobs*) / cov(*PH*, *Zobs*) (C6)

which I refer to as the sensitivity ratio. The left-hand side can be estimated using estimated $\hat{β}$1s from regressions with and without controls. The right-hand side is a measure of the necessary relative size of the covariance between *PH* and the *unobservable* portion of *Z* to the covariance between *PH* and the *observable* portion of *Z* in order for the true effect of *PH* to be zero. More simply, it is the ratio of selection on *unobservables* relative to selection on observables, if the true effect of *PH* is zero (Altonji, Elder, and Taber 2005, p. 177).[[2]](#footnote-2)

 Altonji, Elder, and Taber (2005) argue that “…the ratio of selection on unobservables relative to selection on observables is likely to be less than one…
(pp. 176–77).” Therefore, if the sensitivity ratio is between zero and one, it is possible that selection on unobservables can explain the result. On the other hand, if the sensitivity ratio is *greater* than 1, then it is unlikely that omitted variable bias can explain the entire result.

REFERENCES

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. “Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools.” *Journal of Political Economy* 113, no. 1 (2005): 151–84.

Bellows, John, and Edward Miguel. “War and Local Collective Action in Sierra Leone.” *Journal of Public Economics* 93, no. 11 (2009): 1144–57.

Haines, Michael R. *Historical, Demographic, Economic, and Social Data: The United States, 1790–2000* [computer file]. ICPSR Study 2896. Ann Arbor, MI: Inter-university Consortium for Political and Social Research, 2004.

Leip, David. “Dave Leip’s Atlas of US Presidential Elections.” Electronic database: http://uselectionatlas.org/, 2009.

Ruggles, Steven J., Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. *Integrated Public Use Microdata Series: Version 5.0* [Machine-readable database]. Minneapolis: University of Minnesota, 2010.

United States Department of Housing and Urban Development. *Consolidated Development Directory*. Washington, DC: GPO, 1973.

1. I use the percent of the labor force working on farms as a proxy for the percent of the labor force in agriculture in 1940. [↑](#footnote-ref-1)
2. The sensitivity ratio presented in Altonji, Elder, and Taber (2005) contains a ratio of the variances of observable and unobservable characteristics. Implicit in Bellows and Miguel’s framework is an assumption that the relative variance of observables and unobservables is equal to 1. [↑](#footnote-ref-2)