Online Appendix for

**International Migration Responses to Modern Europe's Most Destructive Earthquake**
**Messina and Reggio Calabria, 1908**

Yannay Spitzer
The Hebrew University of Jerusalem
& CEPR

Gaspare Tortorici
LISER

Ariell Zimran
Vanderbilt University
& NBER

March 2, 2024

# B Additional tables and figures


Figure B.1: Map of districts in the affected area

*Note:* Districts in the provinces of Messina and Reggio Calabria are labeled. The earthquake epicenter is indicated by the large dot.

Table B.1: Mercalli severity scale

| Grade | Description |
|---|---|
| I. Not Felt | Not felt except by very few under especially favorable conditions. |
| II. Weak | Felt only by a few people at rest, especially on upper floors of buildings. |
| III. Weak | Felt quite noticeably by people indoors, especially on upper floors of buildings. Many people do not recognize it as an earthquake. Standing motor cars may rock slightly. Vibrations similar to the passing of a truck. Duration estimated. |
| IV. Light | Felt indoors by many, outdoors by few during the day. At night, some awakened. Dishes, windows, doors disturbed; walls make cracking sound. Sensation like heavy truck striking building. Standing motor cars rocked noticeably. |
| V. Moderate | Felt by nearly everyone; many awakened. Some dishes, windows broken. Unstable objects overturned. Pendulum clocks may stop. |
| VI. Strong | Felt by all, many frightened. Some heavy furniture moved; a few instances of fallen plaster. Damage slight. |
| VII. Very Strong | Damage negligible in buildings of good design and construction; slight to moderate in well-built ordinary structures; considerable damage in poorly built or badly designed structures; some chimneys broken. |
| VIII. Severe | Damage slight in specially designed structures; considerable damage in ordinary substantial buildings with partial collapse. Damage great in poorly built structures. Fall of chimneys, factory stacks, columns, monuments, walls. Heavy furniture overturned. |
| IX. Violent | Damage considerable in specially designed structures; well-designed frame structures thrown out of plumb. Damage great in substantial buildings, with partial collapse. Buildings shifted off foundations. Liquefaction. |
| X. Extreme | Some well-built wooden structures destroyed; most masonry and frame structures destroyed with foundations. Rails bent. |
| XI. Extreme | Few, if any, (masonry) structures remain standing. Bridges destroyed. Broad fissures in ground. Underground pipe lines completely out of service. Earth slumps and land slips in soft ground. Rails bent greatly. |
| XII. Extreme | Damage total. Waves seen on ground surfaces. Lines of sight and level distorted. Objects thrown upward into the air. |

*Source*: US Geological Survey (2019)

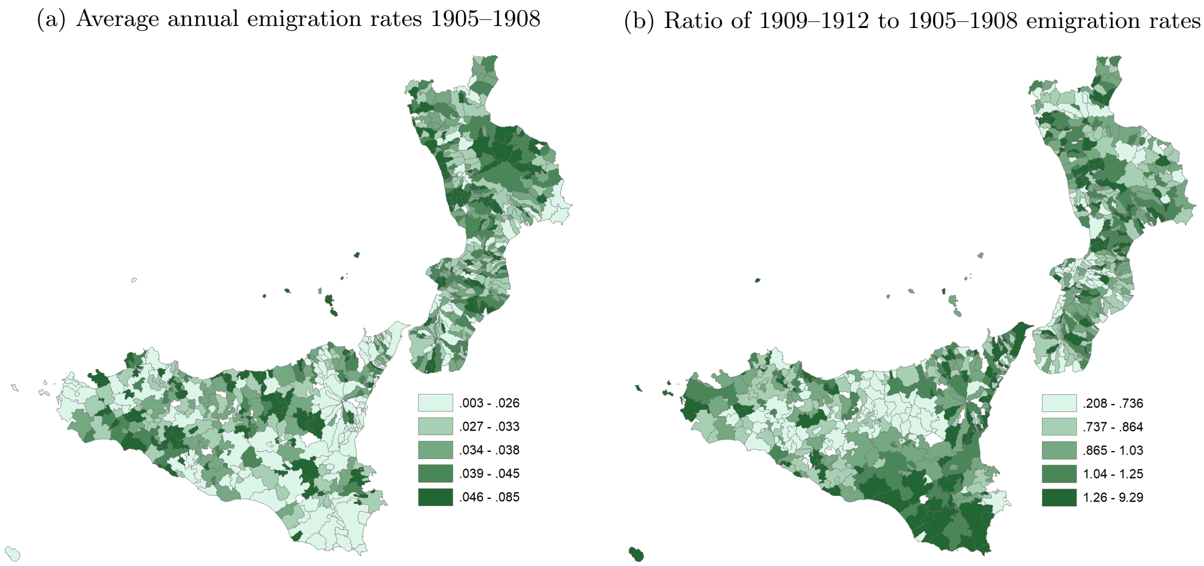



Figure B.2: Municipality-level emigration rates, official data

*Note:* Data are from Italian official statistics. Panel (a) shows average annual emigration rates for 1905–1908. Panel (b) shows the ratio of the average annual emigration rate for 1909–1912 to that for 1905–1908. Both map scales are based on quantiles of the distribution.

Table B.2: Summary statistics for additional municipality and district characteristics

| | Sicily and Calabria | | Calabria | | Sicily | | Italy |
|---|---|---|---|---|---|---|---|
| | Severe | Not | Severe | Not | Severe | Not | All |
| *Variable* | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| *Jacini Agricultural Data* | | | | | | | |
| Share of Area Cultivated | | | | | 0.592$^b$ | 0.714 | |
| | | | | | (0.345) | (0.267) | |
| Share of Cultivated Area Devoted to Trees | | | | | 0.561 | 0.329 | |
| | | | | | (0.267) | (0.293) | |
| Share of Cultivated Area Devoted to Citrus | | | | | 0.089 | 0.028 | |
| | | | | | (0.148) | (0.071) | |
| Cultivated Hectares per Capita | | | | | 0.269 | 0.692 | |
| | | | | | (0.270) | (0.893) | |
| *District Characteristics* | | | | | | | |
| District fraction under age 15 (1901) | 0.231$^b$ | 0.241 | 0.229 | 0.230 | 0.237$^b$ | 0.251 | 0.240 |
| | (0.006) | (0.016) | (0.005) | (0.011) | (0.006) | (0.013) | (0.015) |
| District adult male literacy (1901) | 0.327 | 0.331 | 0.317 | 0.323 | 0.356 | 0.339 | 0.611 |
| | (0.047) | (0.051) | (0.037) | (0.030) | (0.058) | (0.065) | (0.208) |
| *Municipality Characteristics* | | | | | | | |
| Police station present (1901) | 0.364$^a$ | 0.594 | 0.358 | 0.408 | 0.379 | 0.770 | |
| | (0.483) | (0.492) | (0.482) | (0.492) | (0.494) | (0.422) | |
| Post office present (1901) | 0.573$^a$ | 0.828 | 0.543$^c$ | 0.722 | 0.655 | 0.929 | |
| | (0.497) | (0.377) | (0.501) | (0.449) | (0.484) | (0.256) | |
| Telegraph present (1901) | 0.382$^a$ | 0.635 | 0.346 | 0.460 | 0.483$^b$ | 0.801 | |
| | (0.488) | (0.482) | (0.479) | (0.499) | (0.509) | (0.400) | |
| Rail present (1908) | 0.318 | 0.415 | 0.272 | 0.345 | 0.448 | 0.482 | 0.354 |
| | (0.468) | (0.493) | (0.448) | (0.476) | (0.506) | (0.500) | (0.478) |
| Horsepower per capita (1911) | 0.011 | 0.011 | 0.014 | 0.010 | 0.005 | 0.011 | 0.062 |
| | (0.018) | (0.029) | (0.018) | (0.016) | (0.015) | (0.037) | (0.739) |
| Population density (1901) | 0.079$^c$ | 0.055 | 0.063$^b$ | 0.040 | 0.125 | 0.070 | 0.070 |
| | (0.094) | (0.091) | (0.072) | (0.060) | (0.127) | (0.112) | (0.224) |
| Pledge bank assets per capita (1896) | 0.621 | 0.344 | 0.355 | 0.272 | 1.363 | 0.412 | 0.555 |
| | (2.620) | (1.508) | (2.235) | (1.487) | (3.416) | (1.527) | (3.760) |
| Pledge bank assets per capita (25km) | 0.042 | 0.035 | 0.030 | 0.024 | 0.077 | 0.046 | |
| | (0.037) | (0.056) | (0.023) | (0.028) | (0.044) | (0.071) | |
| Savings bank credit per capita (1875) | 0.145 | 0.071 | 0.000 | 0.048 | 0.550 | 0.093 | 2.289 |
| | (1.519) | (1.055) | (0.000) | (0.850) | (2.959) | (1.220) | (20.775) |
| Savings bank credit per capita (25km) | 0.048 | 0.026 | 0.023 | 0.010 | 0.119 | 0.042 | |
| | (0.092) | (0.075) | (0.064) | (0.021) | (0.118) | (0.101) | |

*Notes*: Standard deviations in parentheses. Observations are at the municipality level. Pledge bank assets and savings bank credit are transformed into logarithms for analysis.

*Significance levels, district-clustered b.s.*: Columns (1), (3), and (5) indicate whether the differences between the severe and non-severe groups are statistically significant based on the district-clustered bootstrap. Notation is $^a$ p<0.01, $^b$ p<0.05, $^c$ p<0.1. For instance, an $^a$ in column (1) indicates that the value in column (1) is statistically significantly different from the value in column (2) at the 1-percent level.

Table B.3: Difference-in-differences results, alternative standard errors

| | Ellis Island | | | | Official Data | |
|---|---|---|---|---|---|---|
| | | | Prime-Age Only | | | |
| | (1) All | (2) Deaths | (3) All | (4) Deaths | (5) All | (6) Deaths |
| *Panel A: Sicily and Calabria* | −0.071 | −0.054 | −0.086 | −0.070 | −0.085 | −0.053 |
| Municipality-Clustered SE | (0.084) | (0.083) | (0.081) | (0.080) | (0.048) | (0.049) |
| Conley SE | [0.045] | [0.045] | [0.053] | [0.053] | [0.046] | [0.046] |
| Observations | 5,604 | 5,604 | 5,543 | 5,543 | 6,109 | 6,109 |
| R-squared | 0.751 | 0.751 | 0.741 | 0.741 | 0.618 | 0.617 |
| Municipalities | 737 | 737 | 735 | 735 | 742 | 742 |
| *Panel B: Italy* | −0.057 | −0.040 | −0.076 | −0.060 | −0.068 | −0.037 |
| Municipality-Clustered SE | (0.083) | (0.082) | (0.080) | (0.079) | (0.047) | (0.048) |
| Conley SE | [0.042] | [0.042] | [0.052] | [0.051] | [0.048] | [0.048] |
| Observations | 39,525 | 39,525 | 37,333 | 37,333 | 61,160 | 61,160 |
| R-squared | 0.799 | 0.799 | 0.795 | 0.795 | 0.788 | 0.788 |
| Municipalities | 6257 | 6257 | 6097 | 6097 | 7151 | 7151 |
| *Panel C: Sicily* | −0.123 | −0.105 | −0.124 | −0.106 | −0.069 | −0.037 |
| Municipality-Clustered SE | (0.176) | (0.174) | (0.166) | (0.165) | (0.065) | (0.063) |
| Conley SE | [0.090] | [0.089] | [0.088] | [0.087] | [0.091] | [0.091] |
| Observations | 2,708 | 2,708 | 2,678 | 2,678 | 2,815 | 2,815 |
| R-squared | 0.767 | 0.767 | 0.752 | 0.752 | 0.672 | 0.672 |
| Municipalities | 350 | 350 | 350 | 350 | 352 | 352 |
| *Panel D: Calabria* | −0.035 | −0.022 | −0.065 | −0.052 | −0.145 | −0.117 |
| Municipality-Clustered SE | (0.082) | (0.082) | (0.083) | (0.083) | (0.064) | (0.064) |
| Conley SE | [0.065] | [0.066] | [0.076] | [0.077] | [0.057] | [0.057] |
| Observations | 2,896 | 2,896 | 2,865 | 2,865 | 3,294 | 3,294 |
| R-squared | 0.733 | 0.733 | 0.728 | 0.728 | 0.516 | 0.512 |
| Municipalities | 387 | 387 | 385 | 385 | 390 | 390 |

*Notes*: This table presents difference-in-differences coefficients for severe-times-post. Standard errors clustered at the municipality level in parentheses. Conley standard errors permitting correlation within 3 lags and 150 kilometers are in square brackets. Columns (1), (3), and (5) use unadjusted base populations to compute migration rates; columns (2), (4), and (6) use populations adjusted for earthquake deaths. All regressions include municipality fixed effects, province-year indicators, and distance from the nearest emigration epicenter-times-post. Observations limited to 1905–1912.

Table B.4: Difference-in-differences results for migrant characteristics

| | (1) Age | (2) Male | (3) Prime | (4) Child |
|---|---|---|---|---|
| *Panel A: Sicily and Calabria* | −0.290 | −0.012 | −0.007 | 0.001 |
| District-Clustered b.s. p | (0.411) | (0.428) | (0.747) | (0.931) |
| Randomization Inference tail mass | {0.262} | {0.238} | {0.310} | {0.436} |
| District-Clustered b.s. 95% CI | [-1.155 0.649] | [-0.047 0.026] | [-0.057 0.034] | [-0.027 0.025] |
| Observations | 5,617 | 5,615 | 5,615 | 5,617 |
| R-squared | 0.241 | 0.470 | 0.470 | 0.376 |
| Districts | 35 | 35 | 35 | 35 |
| *Panel B: Italy* | −0.340 | −0.014 | −0.011 | 0.004 |
| District-Clustered p | (0.229) | (0.231) | (0.484) | (0.682) |
| Randomization Inference tail mass | {0.226} | {0.230} | {0.328} | {0.386} |
| District-Clustered 95% CI | [-0.895 0.215] | [-0.038 0.009] | [-0.040 0.019] | [-0.014 0.021] |
| Observations | 40,077 | 40,089 | 40,069 | 40,098 |
| R-squared | 0.245 | 0.323 | 0.325 | 0.281 |
| Districts | 284 | 284 | 284 | 284 |
| *Panel C: Sicily* | 0.054 | 0.007 | 0.029 | −0.007 |
| District-Clustered b.s. p | (0.897) | (0.742) | (0.298) | (0.682) |
| Randomization Inference tail mass | {0.408} | {0.404} | {0.148} | {0.334} |
| District-Clustered b.s. 95% CI | [-1.259 1.147] | [-0.074 0.129] | [-0.058 0.126] | [-0.072 0.056] |
| Observations | 2,718 | 2,718 | 2,718 | 2,718 |
| R-squared | 0.252 | 0.478 | 0.494 | 0.396 |
| Districts | 24 | 24 | 24 | 24 |
| *Panel D: Calabria* | −0.533 | −0.020 | −0.027 | 0.005 |
| District-Clustered b.s. p | (0.250) | (0.360) | (0.180) | (0.532) |
| Randomization Inference tail mass | {0.126} | {0.102} | {0.054} | {0.310} |
| District-Clustered b.s. 95% CI | [-1.603 1.585] | [-0.061 0.032] | [-0.077 0.012] | [-0.037 0.027] |
| Observations | 2,899 | 2,897 | 2,897 | 2,899 |
| R-squared | 0.190 | 0.361 | 0.358 | 0.270 |
| Districts | 11 | 11 | 11 | 11 |

*Significance levels, district-clustered b.s.*: $^a$ p<0.01, $^b$ p<0.05, $^c$ p<0.1

*Notes*: This table presents difference-in-differences coefficients for severe-times-post with the migrant characteristic in the header as the outcome. P-values from a wild bootstrap clustering at the district level are in parentheses. Tail mass values from a randomization inference test are in curly braces. Ninety-five percent confidence intervals from the wild bootstrap clustered at the district level are in square braces. In panel B, the district-clustered p-values and confidence intervals are based on large-sample approximations, not the bootstrap. Columns (1), (3), and (5) use unadjusted base populations to compute migration rates; columns (2), (4), and (6) use populations adjusted for earthquake deaths. All regressions include municipality fixed effects, province-year indicators, and distance from the nearest emigration epicenter-times-post. Observations limited to 1905–1912.

Table B.5: Difference-in-differences results with Mercalli scores

| | Ellis Island | | | | Official Data | |
|---|---|---|---|---|---|---|
| | | | Prime-Age Only | | | |
| | (1) All | (2) Deaths | (3) All | (4) Deaths | (5) All | (6) Deaths |
| *Panel A: Sicily and Calabria* | $-0.060^c$ | $-0.051^c$ | $-0.061^c$ | $-0.052^c$ | $-0.034$ | $-0.016$ |
| District-Clustered b.s. p | (0.074) | (0.094) | (0.083) | (0.099) | (0.234) | (0.696) |
| Randomization Inference tail mass | {0.000} | {0.004} | {0.002} | {0.002} | {0.118} | {0.444} |
| District-Clustered b.s. 95% CI | [-0.123 0.006] | [-0.105 0.006] | [-0.125 0.009] | [-0.111 0.012] | [-0.086 0.017] | [-0.075 0.037] |
| Observations | 5,604 | 5,604 | 5,543 | 5,543 | 6,109 | 6,109 |
| R-squared | 0.751 | 0.751 | 0.742 | 0.742 | 0.618 | 0.617 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 |
| *Panel B: Italy* | $-0.031^c$ | $-0.028^c$ | $-0.034^b$ | $-0.031^c$ | 0.011 | 0.018 |
| District-Clustered p | (0.070) | (0.086) | (0.043) | (0.053) | (0.517) | (0.314) |
| Randomization Inference tail mass | {0.000} | {0.000} | {0.000} | {0.000} | {0.086} | {0.016} |
| District-Clustered 95% CI | [-0.065 0.003] | [-0.060 0.004] | [-0.068 -0.001] | [-0.063 0.000] | [-0.023 0.046] | [-0.017 0.052] |
| Observations | 39,525 | 39,525 | 37,333 | 37,333 | 61,160 | 61,160 |
| R-squared | 0.799 | 0.799 | 0.795 | 0.795 | 0.788 | 0.788 |
| Districts | 284 | 284 | 284 | 284 | 284 | 284 |
| *Panel C: Sicily* | $-0.084^c$ | $-0.078^c$ | $-0.085^c$ | $-0.080^c$ | $-0.020$ | $-0.010$ |
| District-Clustered b.s. p | (0.098) | (0.088) | (0.096) | (0.073) | (0.365) | (0.593) |
| Randomization Inference tail mass | {0.006} | {0.008} | {0.014} | {0.018} | {0.202} | {0.352} |
| District-Clustered b.s. 95% CI | [-0.167 0.017] | [-0.152 0.015] | [-0.175 0.019] | [-0.158 0.009] | [-0.062 0.041] | [-0.047 0.039] |
| Observations | 2,708 | 2,708 | 2,678 | 2,678 | 2,815 | 2,815 |
| R-squared | 0.767 | 0.768 | 0.753 | 0.753 | 0.672 | 0.672 |
| Districts | 24 | 24 | 24 | 24 | 24 | 24 |
| *Panel D: Calabria* | $-0.021$ | $-0.008$ | $-0.021$ | $-0.009$ | $-0.061$ | $-0.035$ |
| District-Clustered b.s. p | (0.599) | (0.821) | (0.658) | (0.795) | (0.380) | (0.909) |
| Randomization Inference tail mass | {0.162} | {0.284} | {0.158} | {0.276} | {0.022} | {0.138} |
| District-Clustered b.s. 95% CI | [-0.105 0.066] | [-0.081 0.069] | [-0.130 0.088] | [-0.110 0.092] | [-0.187 0.055] | [-0.183 0.063] |
| Observations | 2,896 | 2,896 | 2,865 | 2,865 | 3,294 | 3,294 |
| R-squared | 0.733 | 0.733 | 0.728 | 0.728 | 0.516 | 0.512 |
| Districts | 11 | 11 | 11 | 11 | 11 | 11 |

*Significance levels, district-clustered b.s.*: $^a$ p<0.01, $^b$ p<0.05, $^c$ p<0.1

*Notes*: This table presents difference-in-differences coefficients with continuous treatment based on Mercalli intensity, imputed when it is not available. P-values from a wild bootstrap clustering at the district level are in parentheses. Tail mass values from a randomization inference test are in curly braces. Ninety-five percent confidence intervals from the wild bootstrap clustered at the district level are in square braces. In panel B, the district-clustered p-values and confidence intervals are based on large-sample approximations, not the bootstrap. Columns (1), (3), and (5) use unadjusted base populations to compute migration rates; columns (2), (4), and (6) use populations adjusted for earthquake deaths. All regressions include municipality fixed effects, province-year indicators, and distance from the nearest emigration epicenter-times-post. Observations limited to 1905–1912.

Table B.6: Difference-in-differences results with distance from epicenter

| | Ellis Island | | | | Official Data | |
|---|---|---|---|---|---|---|
| | | | Prime-Age Only | | | |
| | (1) All | (2) Deaths | (3) All | (4) Deaths | (5) All | (6) Deaths |
| *Panel A: Sicily and Calabria* | −0.149 | −0.135 | −0.208 | −0.195 | −0.235 | −0.203 |
| District-Clustered b.s. p | (0.377) | (0.451) | (0.210) | (0.264) | (0.107) | (0.187) |
| Randomization Inference tail mass | {0.228} | {0.248} | {0.188} | {0.204} | {0.360} | {0.488} |
| District-Clustered b.s. 95% CI | [-0.454 0.204] | [-0.455 0.214] | [-0.526 0.156] | [-0.487 0.171] | [-0.552 0.071] | [-0.516 0.099] |
| Observations | 5,604 | 5,604 | 5,543 | 5,543 | 6,109 | 6,109 |
| R-squared | 0.751 | 0.751 | 0.742 | 0.742 | 0.620 | 0.618 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 |
| *Panel B: Italy* | −0.030 | −0.027 | −0.033 | −0.030 | $-0.094^{b}$ | $-0.090^{c}$ |
| District-Clustered p | (0.503) | (0.537) | (0.455) | (0.488) | (0.049) | (0.061) |
| Randomization Inference tail mass | {0.480} | {0.448} | {0.028} | {0.034} | {0.078} | {0.098} |
| District-Clustered 95% CI | [-0.118 0.058] | [-0.114 0.059] | [-0.121 0.054] | [-0.117 0.056] | [-0.188 -0.001] | [-0.184 0.004] |
| Observations | 39,525 | 39,525 | 37,333 | 37,333 | 61,160 | 61,160 |
| R-squared | 0.799 | 0.799 | 0.795 | 0.795 | 0.788 | 0.789 |
| Districts | 284 | 284 | 284 | 284 | 284 | 284 |
| *Panel C: Sicily* | $-0.490^{a}$ | $-0.483^{a}$ | $-0.532^{a}$ | $-0.525^{a}$ | $-0.214^{c}$ | $-0.201^{c}$ |
| District-Clustered b.s. p | (0.008) | (0.007) | (0.005) | (0.005) | (0.093) | (0.092) |
| Randomization Inference tail mass | {0.030} | {0.032} | {0.062} | {0.068} | {0.298} | {0.324} |
| District-Clustered b.s. 95% CI | [-0.762 -0.173] | [-0.755 -0.173] | [-0.816 -0.206] | [-0.800 -0.218] | [-0.616 0.055] | [-0.610 0.042] |
| Observations | 2,708 | 2,708 | 2,678 | 2,678 | 2,815 | 2,815 |
| R-squared | 0.769 | 0.769 | 0.755 | 0.755 | 0.673 | 0.673 |
| Districts | 24 | 24 | 24 | 24 | 24 | 24 |
| *Panel D: Calabria* | 0.201 | 0.228 | 0.132 | 0.158 | −0.164 | −0.100 |
| District-Clustered b.s. p | (0.242) | (0.203) | (0.470) | (0.403) | (0.646) | (0.875) |
| Randomization Inference tail mass | {0.254} | {0.290} | {0.178} | {0.228} | {0.038} | {0.134} |
| District-Clustered b.s. 95% CI | [-0.189 0.763] | [-0.164 0.733] | [-0.259 0.736] | [-0.241 0.746] | [-0.865 0.466] | [-0.781 0.491] |
| Observations | 2,896 | 2,896 | 2,865 | 2,865 | 3,294 | 3,294 |
| R-squared | 0.733 | 0.733 | 0.728 | 0.728 | 0.515 | 0.512 |
| Districts | 11 | 11 | 11 | 11 | 11 | 11 |

*Significance levels, district-clustered b.s.*: $^{a}$ p<0.01, $^{b}$ p<0.05, $^{c}$ p<0.1

*Notes*: This table presents difference-in-differences coefficients with the negative of a municipality's distance from the earthquake epicenter (in 100s of kilometers) as the measure of treatment. P-values from a wild bootstrap clustering at the district level are in parentheses. Tail mass values from a randomization inference test are in curly braces. Ninety-five percent confidence intervals from the wild bootstrap clustered at the district level are in square braces. In panel B, the district-clustered p-values and confidence intervals are based on large-sample approximations, not the bootstrap. Columns (1), (3), and (5) use unadjusted base populations to compute migration rates; columns (2), (4), and (6) use populations adjusted for earthquake deaths. All regressions include municipality fixed effects, province-year indicators, and distance from the nearest emigration epicenter-times-post. Observations limited to 1905–1912.

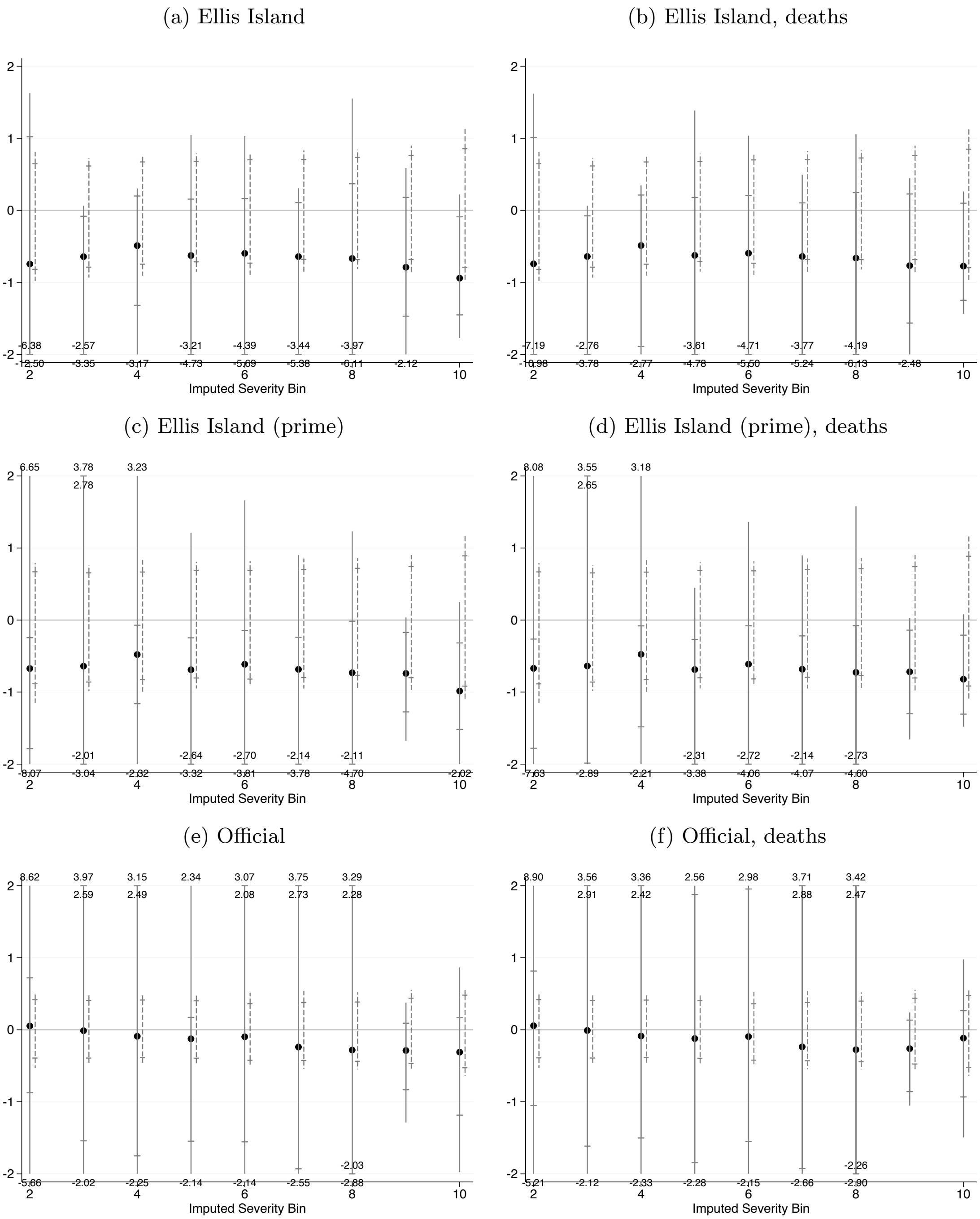



Figure B.3: Effects by Mercalli score bin

*Note:* This figure presents the coefficients from a version of equation (1) in which the severe indicator is replaced by a series of indicators for severity bins 2–10, omitting 1. No municipalities in the sample have severity 0. Severity is imputed when missing. The sample includes all municipalities in the regions of Sicily and Calabria. Solid bars are 90- and 95-percent confidence intervals from the wild bootstrap, clustered on the district level. dashed bars are the middle 90- and 95-percent of results from the randomization inference permutations. When the absolute value of the extreme of a confidence interval exceeds 0.75, the interval is truncated and the true value listed either outside the interval (95-percent) or inside the interval (90-percent).

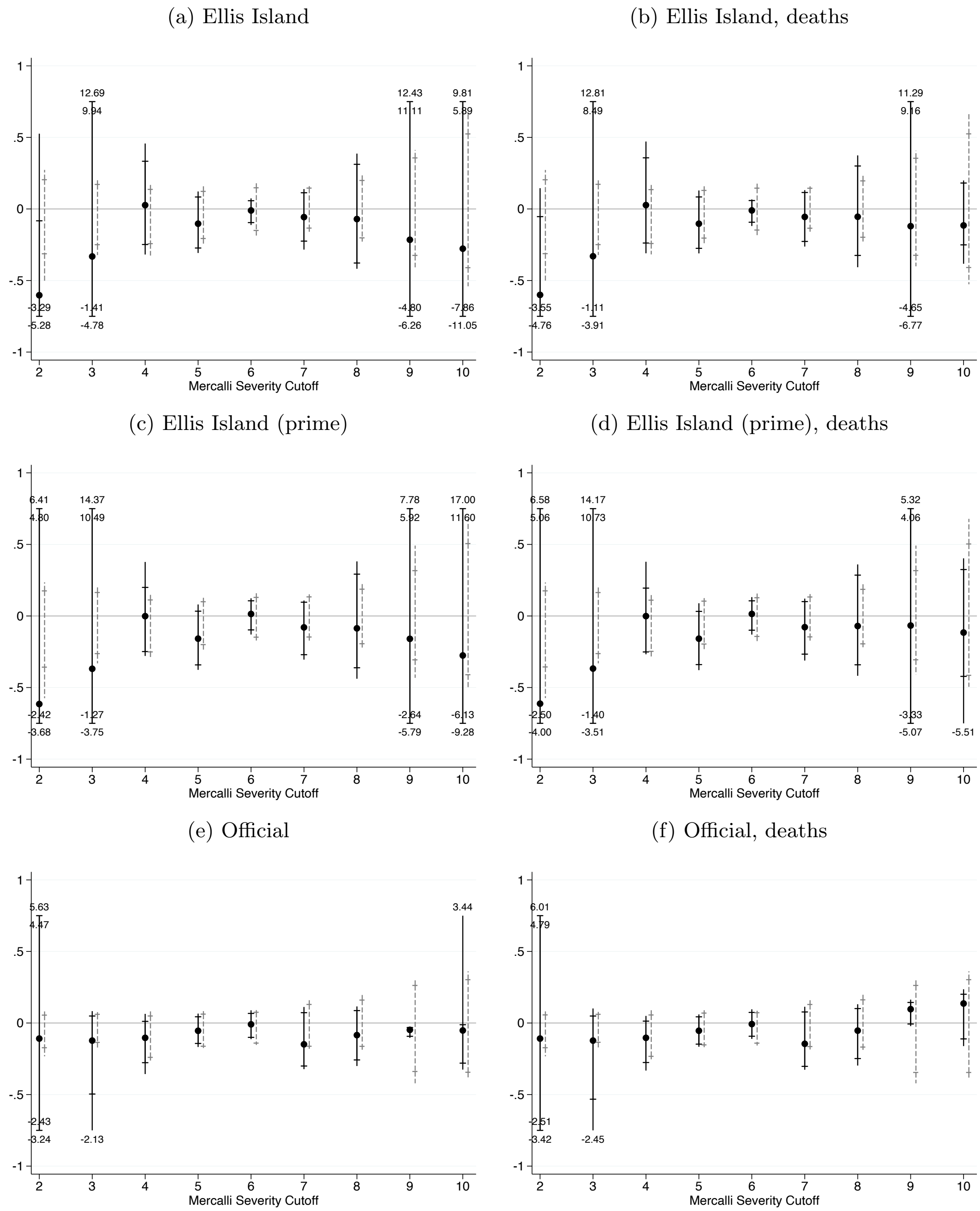



Figure B.4: Difference-in-differences results varying Mercalli cutoffs for severity, imputing severity

*Note:* These figures present the estimated difference-in-differences coefficient from estimating equation (1) with the value listed on the $x$-axis as the Mercalli score cutoff for defining severe damage, imputing the score where it is unavailable. Sample includes all municipalities in the regions of Sicily and Calabria. Solid bars are 90- and 95-percent confidence intervals from the wild bootstrap, clustered on the district level. Dashed bars are the middle 90- and 95-percent of results from the randomization inference permutations. When the absolute value of the extreme of a confidence interval exceeds 0.75, the interval is truncated and the true value listed either outside the interval (95-percent) or inside the interval (90-percent).

Table B.7: Additional difference-in-differences results with share of district employment in agriculture and agricultural day labor

| | Ellis Island | | | | Official Data | |
|---|---|---|---|---|---|---|
| | | | Prime-Age Only | | | |
| | (1) All | (2) Deaths | (3) All | (4) Deaths | (5) All | (6) Deaths |
| Severe x Post | 0.003 | 0.015 | −0.019 | −0.007 | −0.041 | −0.019 |
| District-Clustered b.s. p | (0.981) | (0.904) | (0.868) | (0.955) | (0.434) | (0.744) |
| Randomization Inference tail mass | {0.488} | {0.484} | {0.434} | {0.474} | {0.392} | {0.458} |
| District-Clustered b.s. 95% CI | [-0.185 0.306] | [-0.166 0.304] | [-0.194 0.287] | [-0.173 0.253] | [-0.190 0.083] | [-0.208 0.109] |
| Severe x Post x Ag Lab Share | 0.351[a] | 0.332[a] | 0.333[a] | 0.314[a] | 0.143[b] | 0.104[b] |
| District-Clustered b.s. p | (0.003) | (0.007) | (0.001) | (0.005) | (0.014) | (0.030) |
| Randomization Inference tail mass | {0.004} | {0.004} | {0.002} | {0.002} | {0.102} | {0.144} |
| District-Clustered b.s. 95% CI | [0.100 0.591] | [0.077 0.576] | [0.111 0.537] | [0.091 0.554] | [0.039 0.375] | [0.009 0.334] |
| Observations | 5,604 | 5,604 | 5,543 | 5,543 | 6,109 | 6,109 |
| R-squared | 0.752 | 0.752 | 0.743 | 0.743 | 0.619 | 0.618 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 |
| Severe x Post | −0.001 | 0.009 | −0.019 | −0.009 | −0.100 | −0.083 |
| District-Clustered b.s. p | (0.996) | (0.949) | (0.853) | (0.944) | (0.413) | (0.554) |
| Randomization Inference tail mass | {0.436} | {0.412} | {0.478} | {0.442} | {0.294} | {0.348} |
| District-Clustered b.s. 95% CI | [-0.152 0.317] | [-0.153 0.326] | [-0.186 0.282] | [-0.182 0.311] | [-0.322 0.124] | [-0.322 0.131] |
| Severe x Post x Ag Share | 0.195 | 0.177 | 0.185 | 0.168 | 0.154[c] | 0.119[c] |
| District-Clustered b.s. p | (0.150) | (0.253) | (0.114) | (0.172) | (0.051) | (0.073) |
| Randomization Inference tail mass | {0.108} | {0.122} | {0.090} | {0.104} | {0.138} | {0.162} |
| District-Clustered b.s. 95% CI | [-0.037 0.474] | [-0.054 0.440] | [-0.029 0.416] | [-0.039 0.391] | [-0.002 0.290] | [-0.023 0.243] |
| Observations | 5,604 | 5,604 | 5,543 | 5,543 | 6,109 | 6,109 |
| R-squared | 0.752 | 0.752 | 0.742 | 0.742 | 0.621 | 0.620 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 |
| Severe x Post | 0.007 | 0.018 | −0.014 | −0.004 | −0.079 | −0.061 |
| District-Clustered b.s. p | (0.951) | (0.899) | (0.890) | (0.973) | (0.275) | (0.357) |
| Randomization Inference tail mass | {0.436} | {0.396} | {0.480} | {0.446} | {0.262} | {0.328} |
| District-Clustered b.s. 95% CI | [-0.176 0.299] | [-0.182 0.308] | [-0.201 0.268] | [-0.181 0.293] | [-0.195 0.099] | [-0.187 0.103] |
| Severe x Post x Ag Lab Share | 0.339[b] | 0.347[c] | 0.321[b] | 0.328[b] | 0.155 | 0.174 |
| District-Clustered b.s. p | (0.043) | (0.056) | (0.038) | (0.040) | (0.591) | (0.531) |
| Randomization Inference tail mass | {0.012} | {0.010} | {0.020} | {0.018} | {0.156} | {0.134} |
| District-Clustered b.s. 95% CI | [0.018 0.861] | [-0.049 0.926] | [0.011 0.749] | [0.022 0.760] | [-0.271 0.620] | [-0.199 0.593] |
| Severe x Post x Ag Share | 0.002 | −0.020 | 0.002 | −0.019 | 0.064 | 0.019 |
| District-Clustered b.s. p | (0.984) | (0.870) | (0.982) | (0.844) | (0.633) | (0.886) |
| Randomization Inference tail mass | {0.338} | {0.430} | {0.382} | {0.456} | {0.178} | {0.242} |
| District-Clustered b.s. 95% CI | [-0.306 0.383] | [-0.403 0.388] | [-0.272 0.393] | [-0.327 0.387] | [-0.303 0.344] | [-0.346 0.260] |
| Observations | 5,604 | 5,604 | 5,543 | 5,543 | 6,109 | 6,109 |
| R-squared | 0.752 | 0.752 | 0.743 | 0.743 | 0.623 | 0.622 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 |

*Significance levels, district-clustered b.s.*: [a] p<0.01, [b] p<0.05, [c] p<0.1

*Notes*: Sample includes all municipalities in the regions of Sicily and Calabria. The interaction variables are all at the district level and are all standardized to have mean zero and standard deviation one. P-values from a wild bootstrap clustered at the district level in parentheses. Tail mass values from the randomization inference test are in curly braces. Ninety-five percent confidence intervals from the wild bootstrap clustered at the district level are in square braces. All regressions include municipality fixed effects, province-year indicators, and distance from the nearest emigration epicenter-times-post. Observations limited to 1905–1912.

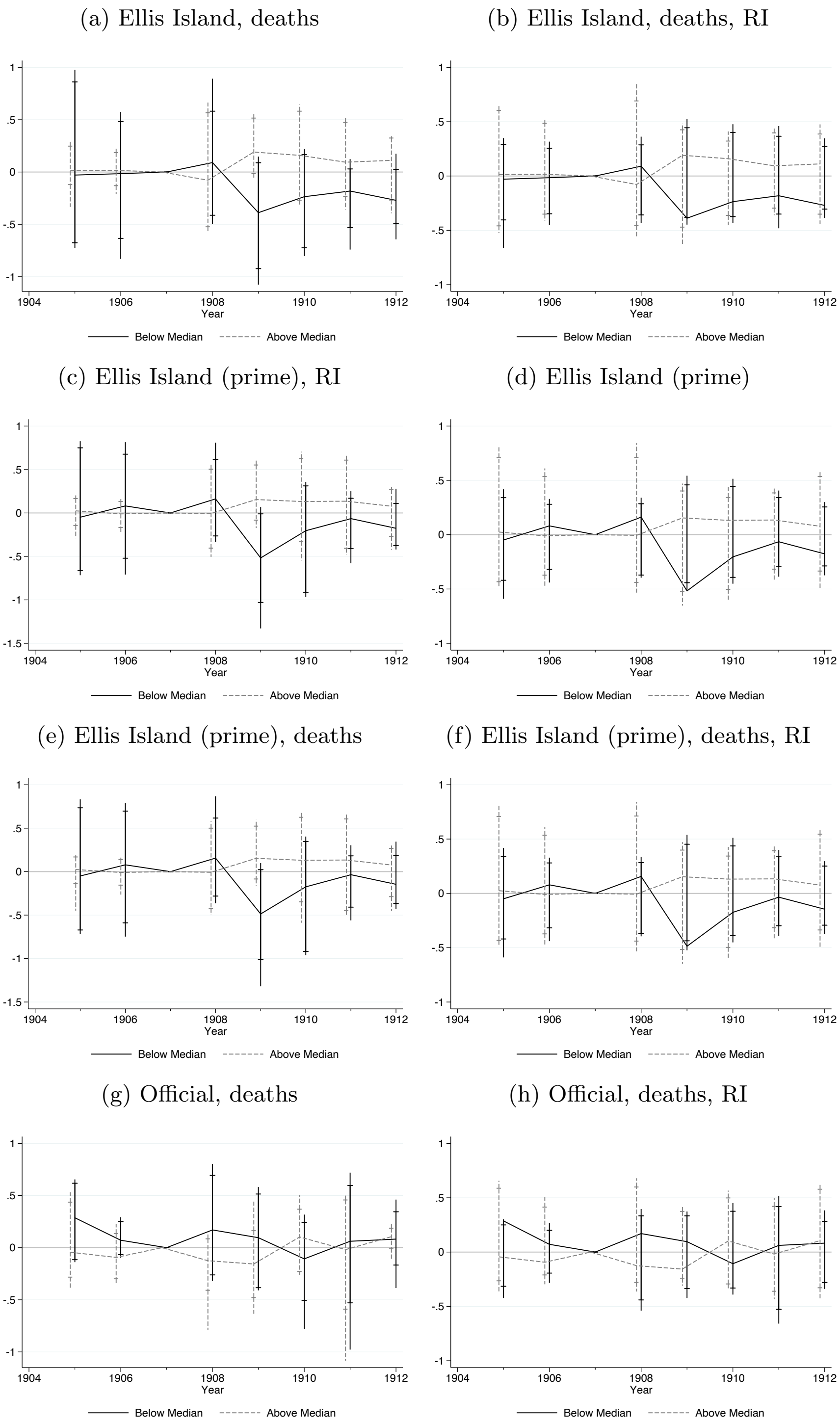



Figure B.5: Event studies divided by share of district employment in agricultural day labor

*Note:* Sample includes all municipalities in the regions of Sicily and Calabria. All event studies control for a year-specific function of distance from the nearest emigration epicenter and have 1907 as the base year. In panels (a) and (b), bars indicate 90- and 95-percent confidence intervals clustered on the district level, computed by a wild bootstrap. In panels (c) and (d), bars indicate the middle 90 and 95 percent of estimates from the randomization inference exercise. The division into "below median" and "above median" is based on the distribution of the share of employment in agricultural day labor in the sample. The measure on the $y$-axis is the effect in logs.

Table B.8: Additional difference-in-differences results with share of district employment in construction and agricultural day labor

| | Ellis Island | | | | Official Data | |
|---|---|---|---|---|---|---|
| | | | Prime-Age Only | | | |
| | (1) All | (2) Deaths | (3) All | (4) Deaths | (5) All | (6) Deaths |
| *Panel A* | | | | | | |
| Severe x Post | 0.018 | 0.025 | −0.004 | 0.002 | −0.082 | −0.072 |
| District-Clustered b.s. p | (0.957) | (0.930) | (0.981) | (0.993) | (0.643) | (0.701) |
| Randomization Inference tail mass | {0.428} | {0.448} | {0.382} | {0.394} | {0.250} | {0.278} |
| District-Clustered b.s. 95% CI | [-0.196 0.460] | [-0.194 0.447] | [-0.225 0.467] | [-0.218 0.415] | [-0.393 0.237] | [-0.381 0.237] |
| Severe x Post x Constr Share | −0.068 | −0.059 | −0.081 | −0.073 | −0.129[c] | −0.110 |
| District-Clustered b.s. p | (0.276) | (0.356) | (0.232) | (0.241) | (0.079) | (0.148) |
| Randomization Inference tail mass | {0.186} | {0.200} | {0.142} | {0.158} | {0.030} | {0.032} |
| District-Clustered b.s. 95% CI | [-0.165 0.111] | [-0.150 0.131] | [-0.186 0.079] | [-0.190 0.084] | [-0.509 0.311] | [-0.493 0.340] |
| Observations | 5,604 | 5,604 | 5,543 | 5,543 | 6,109 | 6,109 |
| R-squared | 0.751 | 0.751 | 0.742 | 0.742 | 0.620 | 0.619 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 |
| *Panel B* | | | | | | |
| Severe x Post | 0.010 | 0.016 | −0.016 | −0.009 | −0.069 | −0.059 |
| District-Clustered b.s. p | (0.933) | (0.854) | (0.888) | (0.940) | (0.555) | (0.610) |
| Randomization Inference tail mass | {0.444} | {0.452} | {0.406} | {0.434} | {0.246} | {0.268} |
| District-Clustered b.s. 95% CI | [-0.233 0.295] | [-0.205 0.275] | [-0.201 0.283] | [-0.216 0.294] | [-0.286 0.369] | [-0.271 0.319] |
| Severe x Post x Constr Share | 0.041 | 0.051 | 0.021 | 0.030 | −0.089 | −0.069 |
| District-Clustered b.s. p | (0.583) | (0.507) | (0.733) | (0.625) | (0.400) | (0.540) |
| Randomization Inference tail mass | {0.378} | {0.406} | {0.352} | {0.374} | {0.058} | {0.072} |
| District-Clustered b.s. 95% CI | [-0.151 0.265] | [-0.137 0.267] | [-0.161 0.202] | [-0.139 0.224] | [-0.478 0.121] | [-0.494 0.136] |
| Severe x Post x Ag Lab Share | 0.360[b] | 0.362[b] | 0.337[b] | 0.339[b] | 0.148 | 0.156 |
| District-Clustered b.s. p | (0.026) | (0.034) | (0.034) | (0.022) | (0.444) | (0.413) |
| Randomization Inference tail mass | {0.008} | {0.008} | {0.012} | {0.010} | {0.138} | {0.128} |
| District-Clustered b.s. 95% CI | [0.063 0.922] | [0.047 0.926] | [0.050 0.817] | [0.057 0.837] | [-0.269 0.434] | [-0.248 0.429] |
| Observations | 5,604 | 5,604 | 5,543 | 5,543 | 6,109 | 6,109 |
| R-squared | 0.752 | 0.752 | 0.743 | 0.743 | 0.621 | 0.620 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 |

*Significance levels, district-clustered b.s.*: [a] p<0.01, [b] p<0.05, [c] p<0.1
*Notes*: Sample includes all municipalities in the regions of Sicily and Calabria. The construction labor share and the agricultural labor share are both at the district level and are both standardized to have mean zero and standard deviation one. P-values from a wild bootstrap clustered at the district level in parentheses. Tail mass values from the randomization inference test are in curly braces. Ninety-five percent confidence intervals from the wild bootstrap clustered at the district level are in square braces. All regressions include commune fixed effects, province-year indicators, and distance from the nearest emigration epicenter-times-post. Observations limited to 1905–1912.

(a) Trends (b) Event Study

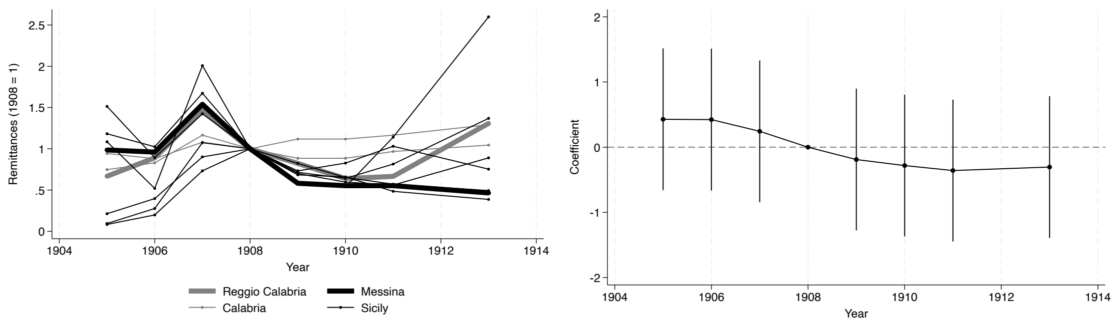



Figure B.6: Remittance patterns

*Note:* This figure studies province-level remittance data in Sicily and Calabria over the period 1907–1913 (data for 1912 are missing). Provinces in Sicily are in black and Messina is the bold line. Provinces in Calabria are in gray and Reggio Calabria is the bold line. Panel (b) performs an event study, with Messina and Reggio Calabria as the treated provinces and all others in Sicily and Calabria as controls. Standard errors in panel (b) are not adjusted in any way.

Table B.9: Additional difference-in-differences results with district property owners per capita and share of district employment in agricultural day labor

| | Ellis Island | | | | Official Data | |
|---|---|---|---|---|---|---|
| | | | Prime-Age Only | | | |
| | (1) All | (2) Deaths | (3) All | (4) Deaths | (5) All | (6) Deaths |
| *Panel A* | | | | | | |
| Severe x Post | −0.031 | −0.018 | −0.046 | −0.033 | −0.080 | −0.057 |
| District-Clustered b.s. p | (0.528) | (0.719) | (0.371) | (0.531) | (0.415) | (0.711) |
| Randomization Inference tail mass | {0.366} | {0.398} | {0.304} | {0.350} | {0.106} | {0.178} |
| District-Clustered b.s. 95% CI | [-0.155 0.106] | [-0.141 0.138] | [-0.175 0.086] | [-0.160 0.110] | [-0.276 0.083] | [-0.283 0.087] |
| Severe x Post x Building Owners per capita | 0.336 | 0.316 | 0.321[c] | 0.301 | 0.196[b] | 0.156[c] |
| District-Clustered b.s. p | (0.138) | (0.243) | (0.067) | (0.120) | (0.017) | (0.054) |
| Randomization Inference tail mass | {0.002} | {0.002} | {0.002} | {0.004} | {0.146} | {0.194} |
| District-Clustered b.s. 95% CI | [-0.051 0.666] | [-0.069 0.658] | [-0.012 0.636] | [-0.031 0.554] | [0.074 0.331] | [-0.006 0.281] |
| Observations | 5,604 | 5,604 | 5,543 | 5,543 | 6,109 | 6,109 |
| R-squared | 0.752 | 0.752 | 0.743 | 0.743 | 0.621 | 0.620 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 |
| *Panel B* | | | | | | |
| Severe x Post | −0.042 | −0.028 | −0.055 | −0.042 | −0.084 | −0.058 |
| District-Clustered b.s. p | (0.612) | (0.741) | (0.543) | (0.617) | (0.299) | (0.583) |
| Randomization Inference tail mass | {0.310} | {0.350} | {0.284} | {0.340} | {0.104} | {0.176} |
| District-Clustered b.s. 95% CI | [-0.264 0.151] | [-0.254 0.166] | [-0.275 0.152] | [-0.240 0.166] | [-0.247 0.073] | [-0.246 0.087] |
| Severe x Post x Land Owners Per Capita | 0.252 | 0.239 | 0.243 | 0.229 | 0.177[b] | 0.151[b] |
| District-Clustered b.s. p | (0.250) | (0.348) | (0.235) | (0.290) | (0.044) | (0.033) |
| Randomization Inference tail mass | {0.008} | {0.008} | {0.004} | {0.008} | {0.162} | {0.192} |
| District-Clustered b.s. 95% CI | [-0.107 0.616] | [-0.106 0.619] | [-0.087 0.556] | [-0.088 0.531] | [0.008 0.321] | [0.016 0.266] |
| Observations | 5,604 | 5,604 | 5,543 | 5,543 | 6,109 | 6,109 |
| R-squared | 0.752 | 0.752 | 0.742 | 0.742 | 0.621 | 0.619 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 |
| *Panel C* | | | | | | |
| Severe x Post | −0.027 | −0.014 | −0.038 | −0.026 | −0.074 | −0.050 |
| District-Clustered b.s. p | (0.615) | (0.809) | (0.490) | (0.663) | (0.537) | (0.711) |
| Randomization Inference tail mass | {0.396} | {0.422} | {0.348} | {0.378} | {0.136} | {0.222} |
| District-Clustered b.s. 95% CI | [-0.150 0.088] | [-0.141 0.100] | [-0.165 0.079] | [-0.162 0.106] | [-0.257 0.094] | [-0.241 0.110] |
| Severe x Post x Building Owners per capita | 0.748 | 0.694 | 0.703 | 0.651 | 0.220 | 0.101 |
| District-Clustered b.s. p | (0.228) | (0.273) | (0.206) | (0.254) | (0.612) | (0.738) |
| Randomization Inference tail mass | {0.016} | {0.018} | {0.016} | {0.016} | {0.154} | {0.272} |
| District-Clustered b.s. 95% CI | [-0.994 2.287] | [-1.098 1.633] | [-0.831 1.423] | [-0.848 1.323] | [-0.951 0.600] | [-1.020 0.481] |
| Severe x Post x Land Owners per capita | −0.410 | −0.376 | −0.382 | −0.350 | −0.028 | 0.051 |
| District-Clustered b.s. p | (0.294) | (0.301) | (0.315) | (0.359) | (0.918) | (0.808) |
| Randomization Inference tail mass | {0.058} | {0.062} | {0.072} | {0.082} | {0.442} | {0.408} |
| District-Clustered b.s. 95% CI | [-1.322 0.570] | [-1.512 0.706] | [-1.205 0.571] | [-1.194 0.628] | [-0.378 0.884] | [-0.291 0.944] |
| Observations | 5,604 | 5,604 | 5,543 | 5,543 | 6,109 | 6,109 |
| R-squared | 0.752 | 0.752 | 0.743 | 0.743 | 0.621 | 0.620 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 |
| *Panel D* | | | | | | |
| Severe x Post | −0.005 | 0.008 | −0.020 | −0.008 | −0.048 | −0.024 |
| District-Clustered b.s. p | (0.936) | (0.897) | (0.705) | (0.910) | (0.304) | (0.634) |
| Randomization Inference tail mass | {0.448} | {0.494} | {0.414} | {0.450} | {0.218} | {0.298} |
| District-Clustered b.s. 95% CI | [-0.156 0.141] | [-0.160 0.164] | [-0.147 0.115] | [-0.162 0.139] | [-0.177 0.068] | [-0.163 0.080] |
| Severe x Post x Building Owners per capita | 0.506 | 0.453 | 0.473 | 0.423 | 0.044 | −0.076 |
| District-Clustered b.s. p | (0.144) | (0.180) | (0.103) | (0.148) | (0.880) | (0.828) |
| Randomization Inference tail mass | {0.064} | {0.078} | {0.094} | {0.116} | {0.352} | {0.394} |
| District-Clustered b.s. 95% CI | [-0.383 2.161] | [-0.414 2.215] | [-0.175 1.895] | [-0.289 2.111] | [-1.223 1.298] | [-1.119 1.269] |
| Severe x Post x Land Owners per capita | −0.281 | −0.249 | −0.259 | −0.228 | 0.067 | 0.146 |
| District-Clustered b.s. p | (0.227) | (0.289) | (0.252) | (0.298) | (0.804) | (0.641) |
| Randomization Inference tail mass | {0.152} | {0.174} | {0.212} | {0.238} | {0.356} | {0.264} |
| District-Clustered b.s. 95% CI | [-1.585 0.564] | [-1.596 0.708] | [-1.557 0.447] | [-1.650 0.437] | [-1.101 0.920] | [-1.246 0.876] |
| Severe x Post x Ag Lab Share | 0.184[b] | 0.183[b] | 0.173[c] | 0.172[c] | 0.119 | 0.121 |
| District-Clustered b.s. p | (0.038) | (0.030) | (0.066) | (0.069) | (0.391) | (0.456) |
| Randomization Inference tail mass | {0.084} | {0.088} | {0.122} | {0.124} | {0.106} | {0.106} |
| District-Clustered b.s. 95% CI | [0.034 0.330] | [0.040 0.337] | [-0.051 0.321] | [-0.037 0.331] | [-0.094 0.367] | [-0.105 0.375] |
| Observations | 5,604 | 5,604 | 5,543 | 5,543 | 6,109 | 6,109 |
| R-squared | 0.753 | 0.752 | 0.743 | 0.743 | 0.622 | 0.621 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 |

*Significance levels, district-clustered b.s.:* [a] p<0.01, [b] p<0.05, [c] p<0.1

*Notes:* Sample includes all municipalities in the regions of Sicily and Calabria. The interaction variables are all at the district level and are all standardized to have mean zero and standard deviation one. P-values from a wild bootstrap clustered at the district level in parentheses. Tail mass values from the randomization inference test are in curly braces. Ninety-five percent confidence intervals from the wild bootstrap clustered at the district level are in square braces. All regressions include commune fixed effects, province-year indicators, and distance from the nearest emigration epicenter-times-post. Observations limited to 1905–1912.

Table B.10: Additional difference-in-differences results with heterogeneity

| | Ellis Island | | | Official | | |
|---|---|---|---|---|---|---|
| | No Ag. Lab. | With Ag. Lab. | | No Ag. Lab. | With Ag. Lab. | |
| | (1) Inter. | (2) Inter. | (3) Ag. Lab. Inter. | (4) Inter. | (5) Inter. | (6) Ag. Lab. Inter. |
| Credit Employment Share | −0.182 | −0.021 | $0.329^b$ | $-0.138^c$ | 0.017 | $0.310^a$ |
| District-Clustered b.s. p | (0.186) | (0.749) | (0.049) | (0.058) | (0.821) | (0.003) |
| Randomization Inference tail mass | {0.034} | {0.224} | {0.006} | {0.112} | {0.324} | {0.048} |
| District-Clustered b.s. 95% CI | [-0.402 0.053] | [-0.275 0.467] | [0.002 1.190] | [-0.383 0.009] | [-0.477 0.458] | [0.096 0.856] |
| Observations | 5,604 | 5,604 | | 6,109 | 6,109 | |
| R-squared | 0.751 | 0.752 | | 0.623 | 0.626 | |
| Districts | 35 | 35 | | 35 | 35 | |
| Credit per capita | $-0.111^b$ | $-0.130^c$ | $0.367^a$ | $-0.072^b$ | $-0.081^b$ | $0.155^b$ |
| District-Clustered b.s. p | (0.020) | (0.056) | (0.002) | (0.030) | (0.048) | (0.015) |
| Randomization Inference tail mass | {0.050} | {0.032} | {0.002} | {0.034} | {0.026} | {0.100} |
| District-Clustered b.s. 95% CI | [-0.175 -0.038] | [-0.203 0.010] | [0.113 0.557] | [-0.129 -0.032] | [-0.108 -0.002] | [0.032 0.393] |
| Observations | 5,604 | 5,604 | | 6,109 | 6,109 | |
| R-squared | 0.751 | 0.752 | | 0.619 | 0.620 | |
| Districts | 35 | 35 | | 35 | 35 | |
| Credit per capita (25km) | −0.273 | $-0.183^b$ | $0.307^a$ | −0.135 | −0.094 | $0.123^c$ |
| District-Clustered b.s. p | (0.461) | (0.018) | (0.008) | (0.761) | (0.872) | (0.065) |
| Randomization Inference tail mass | {0.004} | {0.040} | {0.004} | {0.012} | {0.064} | {0.122} |
| District-Clustered b.s. 95% CI | [-1.401 0.640] | [-0.599 -0.091] | [0.079 0.442] | [-0.612 0.684] | [-1.067 0.697] | [-0.014 0.344] |
| Observations | 5,604 | 5,604 | | 6,109 | 6,109 | |
| R-squared | 0.751 | 0.752 | | 0.619 | 0.620 | |
| Districts | 35 | 35 | | 35 | 35 | |
| Assets per capita | $-0.108^b$ | $-0.126^c$ | $0.366^a$ | $-0.072^b$ | $-0.081^c$ | $0.154^b$ |
| District-Clustered b.s. p | (0.023) | (0.067) | (0.005) | (0.039) | (0.053) | (0.013) |
| Randomization Inference tail mass | {0.052} | {0.034} | {0.002} | {0.028} | {0.024} | {0.100} |
| District-Clustered b.s. 95% CI | [-0.180 -0.035] | [-0.191 0.032] | [0.105 0.601] | [-0.121 -0.025] | [-0.108 0.009] | [0.035 0.399] |
| Observations | 5,604 | 5,604 | | 6,109 | 6,109 | |
| R-squared | 0.751 | 0.752 | | 0.619 | 0.620 | |
| Districts | 35 | 35 | | 35 | 35 | |
| Assets per capita (25km) | −0.291 | $-0.190^b$ | $0.299^a$ | −0.138 | −0.091 | $0.122^c$ |
| District-Clustered b.s. p | (0.454) | (0.040) | (0.004) | (0.615) | (0.902) | (0.065) |
| Randomization Inference tail mass | {0.002} | {0.042} | {0.006} | {0.008} | {0.068} | {0.124} |
| District-Clustered b.s. 95% CI | [-1.225 0.695] | [-0.639 -0.067] | [0.087 0.427] | [-0.432 0.555] | [-0.953 0.739] | [-0.011 0.338] |
| Observations | 5,604 | 5,604 | | 6,109 | 6,109 | |
| R-squared | 0.751 | 0.752 | | 0.619 | 0.620 | |
| Districts | 35 | 35 | | 35 | 35 | |
| Postal Credit per capita | −0.028 | −0.043 | $0.354^a$ | 0.005 | −0.003 | $0.147^b$ |
| District-Clustered b.s. p | (0.552) | (0.370) | (0.003) | (0.916) | (0.940) | (0.018) |
| Randomization Inference tail mass | {0.414} | {0.330} | {0.004} | {0.388} | {0.448} | {0.104} |
| District-Clustered b.s. 95% CI | [-0.127 0.080] | [-0.162 0.046] | [0.110 0.599] | [-0.080 0.154] | [-0.084 0.138] | [0.045 0.365] |
| Observations | 5,604 | 5,604 | | 6,109 | 6,109 | |
| R-squared | 0.751 | 0.752 | | 0.619 | 0.620 | |
| Districts | 35 | 35 | | 35 | 35 | |
| Postal Credit per capita (25km) | −0.094 | 0.031 | $0.381^a$ | −0.048 | 0.005 | $0.162^a$ |
| District-Clustered b.s. p | (0.423) | (0.707) | (0.007) | (0.197) | (0.868) | (0.003) |
| Randomization Inference tail mass | {0.410} | {0.316} | {0.006} | {0.384} | {0.366} | {0.096} |
| District-Clustered b.s. 95% CI | [-0.484 0.335] | [-0.158 0.159] | [0.106 0.674] | [-0.110 0.081] | [-0.052 0.117] | [0.037 0.346] |
| Observations | 5,604 | 5,604 | | 6,109 | 6,109 | |
| R-squared | 0.751 | 0.752 | | 0.618 | 0.619 | |
| Districts | 35 | 35 | | 35 | 35 | |
| MA Members per capita | 0.139 | 0.103 | $0.339^a$ | −0.082 | $-0.093^c$ | $0.148^a$ |
| District-Clustered b.s. p | (0.116) | (0.390) | (0.004) | (0.107) | (0.083) | (0.009) |
| Randomization Inference tail mass | {0.022} | {0.054} | {0.004} | {0.026} | {0.024} | {0.102} |
| District-Clustered b.s. 95% CI | [-0.040 0.423] | [-0.517 0.363] | [0.090 0.570] | [-0.246 0.102] | [-0.306 0.100] | [0.043 0.370] |
| Observations | 5,604 | 5,604 | | 6,109 | 6,109 | |
| R-squared | 0.751 | 0.752 | | 0.618 | 0.619 | |
| Districts | 35 | 35 | | 35 | 35 | |
| MA Assets per capita | −0.028 | −0.062 | $0.359^a$ | −0.055 | $-0.067^c$ | $0.153^b$ |
| District-Clustered b.s. p | (0.572) | (0.212) | (0.004) | (0.146) | (0.063) | (0.015) |
| Randomization Inference tail mass | {0.334} | {0.182} | {0.004} | {0.056} | {0.042} | {0.096} |
| District-Clustered b.s. 95% CI | [-0.132 0.104] | [-0.181 0.032] | [0.107 0.567] | [-0.112 0.066] | [-0.127 0.012] | [0.047 0.366] |
| Observations | 5,604 | 5,604 | | 6,109 | 6,109 | |
| R-squared | 0.751 | 0.752 | | 0.618 | 0.619 | |
| Districts | 35 | 35 | | 35 | 35 | |
| Network | 0.059 | 0.077 | $0.352^a$ | −0.009 | 0.013 | $0.122^b$ |
| District-Clustered b.s. p | (0.732) | (0.536) | (0.003) | (0.883) | (0.712) | (0.038) |
| Randomization Inference tail mass | {0.346} | {0.258} | {0.006} | {0.486} | {0.408} | {0.124} |
| District-Clustered b.s. 95% CI | [-0.216 0.212] | [-0.185 0.217] | [0.121 0.605] | [-0.160 0.145] | [-0.072 0.078] | [0.012 0.376] |
| Observations | 5,604 | 5,604 | | 6,109 | 6,109 | |
| R-squared | 0.753 | 0.754 | | 0.623 | 0.623 | |
| Districts | 35 | 35 | | 35 | 35 | |

*Significance levels, district-clustered b.s.:* [a] p<0.01, [b] p<0.05, [c] p<0.1

*Notes*: This table tests for heterogeneity on a variety of dimensions. Sample includes all municipalities in the regions of Sicily and Calabria. Columns (1) and (4) test for heterogeneity with respect to the listed variable, each of which is standardized to have mean zero and standard deviation one. Columns (2) and (3), and (5) and (6) are from regressions looking for heterogeneity on both the basis of the listed variable and the share of employment in agricultural day labor, presenting the results for both. P-values from a wild bootstrap at the district level in parentheses. Tail mass values from a randomization inference exercise in curly braces. Square brackets show 95-percent district-clustered confidence intervals from the wild bootstrap.

Table B.11: Additional difference-in-differences results with heterogeneity, alternative variables

| | Ellis Island | | | Official | | |
|---|---|---|---|---|---|---|
| | No Ag. Lab. | With Ag. Lab. | | No Ag. Lab. | With Ag. Lab. | |
| | (1) Inter. | (2) Inter. | (3) Ag. Lab. Inter. | (4) Inter. | (5) Inter. | (6) Ag. Lab. Inter. |
| Rail present (1908) | −0.065 | −0.005 | 0.353$^a$ | −0.068 | −0.041 | 0.129$^b$ |
| District-Clustered b.s. p | (0.613) | (0.956) | (0.000) | (0.120) | (0.368) | (0.017) |
| Randomization Inference tail mass | {0.114} | {0.386} | {0.004} | {0.070} | {0.156} | {0.112} |
| District-Clustered b.s. 95% CI | [-0.313 0.186] | [-0.199 0.207] | [0.130 0.598] | [-0.132 0.032] | [-0.108 0.080] | [0.020 0.343] |
| Observations | 5,604 | 5,604 | | 6,109 | 6,109 | |
| R-squared | 0.751 | 0.752 | | 0.619 | 0.619 | |
| Districts | 35 | 35 | | 35 | 35 | |
| Horsepower per capita (1911) | 0.032 | −0.029 | 0.363$^a$ | 0.011 | 0.006 | 0.143$^b$ |
| District-Clustered b.s. p | (0.625) | (0.647) | (0.002) | (0.677) | (0.775) | (0.015) |
| Randomization Inference tail mass | {0.298} | {0.402} | {0.002} | {0.382} | {0.432} | {0.108} |
| District-Clustered b.s. 95% CI | [-0.141 0.145] | [-0.207 0.079] | [0.136 0.542] | [-0.049 0.086] | [-0.050 0.075] | [0.038 0.369] |
| Observations | 5,604 | 5,604 | | 6,109 | 6,109 | |
| R-squared | 0.751 | 0.752 | | 0.618 | 0.619 | |
| Districts | 35 | 35 | | 35 | 35 | |
| Population density (1901) | −0.108 | −0.075 | 0.336$^a$ | −0.037 | −0.041 | 0.143$^b$ |
| District-Clustered b.s. p | (0.253) | (0.489) | (0.003) | (0.214) | (0.170) | (0.019) |
| Randomization Inference tail mass | {0.144} | {0.280} | {0.004} | {0.162} | {0.170} | {0.104} |
| District-Clustered b.s. 95% CI | [-0.330 0.205] | [-0.315 0.222] | [0.112 0.587] | [-0.173 0.027] | [-0.172 0.032] | [0.033 0.379] |
| Observations | 5,604 | 5,604 | | 6,109 | 6,109 | |
| R-squared | 0.751 | 0.752 | | 0.618 | 0.619 | |
| Districts | 35 | 35 | | 35 | 35 | |
| Post office present 1901 | 0.025 | 0.008 | 0.352$^a$ | 0.022 | 0.016 | 0.143$^b$ |
| District-Clustered b.s. p | (0.476) | (0.827) | (0.007) | (0.573) | (0.673) | (0.017) |
| Randomization Inference tail mass | {0.344} | {0.400} | {0.004} | {0.206} | {0.234} | {0.104} |
| District-Clustered b.s. 95% CI | [-0.052 0.101] | [-0.070 0.089] | [0.094 0.612] | [-0.050 0.123] | [-0.050 0.109] | [0.033 0.367] |
| Observations | 5,597 | 5,597 | | 6,101 | 6,101 | |
| R-squared | 0.750 | 0.752 | | 0.619 | 0.620 | |
| Districts | 35 | 35 | | 35 | 35 | |
| Telegraph present (1901) | −0.067$^b$ | −0.066 | 0.351$^a$ | 0.019 | 0.016 | 0.140$^b$ |
| District-Clustered b.s. p | (0.046) | (0.108) | (0.004) | (0.692) | (0.728) | (0.016) |
| Randomization Inference tail mass | {0.104} | {0.090} | {0.004} | {0.258} | {0.262} | {0.108} |
| District-Clustered b.s. 95% CI | [-0.145 -0.001] | [-0.163 0.028] | [0.096 0.567] | [-0.088 0.175] | [-0.085 0.163] | [0.034 0.362] |
| Observations | 5,597 | 5,597 | | 6,101 | 6,101 | |
| R-squared | 0.750 | 0.752 | | 0.619 | 0.620 | |
| Districts | 35 | 35 | | 35 | 35 | |
| Police station present (1901) | −0.052 | −0.061 | 0.354$^a$ | 0.011 | 0.006 | 0.143$^b$ |
| District-Clustered b.s. p | (0.493) | (0.358) | (0.003) | (0.736) | (0.847) | (0.015) |
| Randomization Inference tail mass | {0.344} | {0.264} | {0.004} | {0.372} | {0.420} | {0.102} |
| District-Clustered b.s. 95% CI | [-0.299 0.041] | [-0.313 0.038] | [0.102 0.605] | [-0.061 0.125] | [-0.066 0.118] | [0.048 0.362] |
| Observations | 5,597 | 5,597 | | 6,101 | 6,101 | |
| R-squared | 0.750 | 0.752 | | 0.619 | 0.620 | |
| Districts | 35 | 35 | | 35 | 35 | |
| District adult male literacy (1901) | −0.254$^b$ | −0.058 | 0.312$^b$ | −0.195$^b$ | −0.158 | 0.054 |
| District-Clustered b.s. p | (0.020) | (0.634) | (0.023) | (0.047) | (0.359) | (0.651) |
| Randomization Inference tail mass | {0.018} | {0.168} | {0.022} | {0.014} | {0.084} | {0.396} |
| District-Clustered b.s. 95% CI | [-0.660 -0.016] | [-0.663 0.165] | [0.062 0.681] | [-0.401 -0.002] | [-0.332 0.130] | [-0.185 0.305] |
| Observations | 5,604 | 5,604 | | 6,109 | 6,109 | |
| R-squared | 0.752 | 0.752 | | 0.623 | 0.626 | |
| Districts | 35 | 35 | | 35 | 35 | |
| District fraction under age 15 (1901) | −0.290 | −0.017 | 0.342$^a$ | 0.031 | 0.216 | 0.239$^a$ |
| District-Clustered b.s. p | (0.392) | (0.918) | (0.010) | (0.861) | (0.133) | (0.002) |
| Randomization Inference tail mass | {0.058} | {0.320} | {0.004} | {0.374} | {0.070} | {0.070} |
| District-Clustered b.s. 95% CI | [-1.079 0.226] | [-0.485 0.616] | [0.105 0.675] | [-0.278 0.569] | [-0.080 0.608] | [0.089 0.486] |
| Observations | 5,604 | 5,604 | | 6,109 | 6,109 | |
| R-squared | 0.751 | 0.752 | | 0.619 | 0.621 | |
| Districts | 35 | 35 | | 35 | 35 | |
| Ag. Lab. Share (Sicily) | | | 0.406$^b$ | | | 0.098 |
| District-Clustered b.s. p | | | (0.039) | | | (0.222) |
| Randomization Inference tail mass | | | {0.008} | | | {0.096} |
| District-Clustered b.s. 95% CI | | | [0.009 0.741] | | | [-0.240 0.239] |
| Observations | | | 2,708 | | | 2,815 |
| R-squared | | | 0.770 | | | 0.673 |
| Districts | | | 24 | | | 24 |
| Share of cultivated area devoted to trees | 0.009 | 0.197 | 0.412$^c$ | −0.053 | −0.019 | 0.076 |
| District-Clustered b.s. p | (0.709) | (0.163) | (0.087) | (0.533) | (0.827) | (0.389) |
| Randomization Inference tail mass | {0.368} | {0.092} | {0.000} | {0.136} | {0.336} | {0.254} |
| District-Clustered b.s. 95% CI | [-0.940 0.970] | [-1.002 0.939] | [-0.152 0.654] | [-0.802 0.132] | [-1.187 0.688] | [-0.463 0.825] |
| Observations | 2,631 | 2,631 | | 2,737 | 2,737 | |
| R-squared | 0.768 | 0.770 | | 0.673 | 0.674 | |
| Districts | 24 | 24 | | 24 | 24 | |
| Share of cultivated area devoted to citrus | 0.230$^c$ | 0.240$^c$ | 0.362$^c$ | −0.048 | −0.045 | 0.072 |
| District-Clustered b.s. p | (0.095) | (0.055) | (0.080) | (0.518) | (0.574) | (0.339) |
| Randomization Inference tail mass | {0.032} | {0.038} | {0.004} | {0.164} | {0.166} | {0.122} |
| District-Clustered b.s. 95% CI | [-0.181 0.495] | [-0.037 0.649] | [-0.181 0.606] | [-0.293 0.080] | [-0.231 0.097] | [-0.288 0.412] |
| Observations | 2,631 | 2,631 | | 2,737 | 2,737 | |
| R-squared | 0.769 | 0.772 | | 0.673 | 0.673 | |
| Districts | 24 | 24 | | 24 | 24 | |
| Share of area cultivated | 0.271$^b$ | 0.218$^c$ | 0.365$^c$ | 0.068 | 0.061 | 0.073 |
| District-Clustered b.s. p | (0.046) | (0.098) | (0.089) | (0.623) | (0.636) | (0.243) |
| Randomization Inference tail mass | {0.024} | {0.046} | {0.000} | {0.140} | {0.172} | {0.300} |
| District-Clustered b.s. 95% CI | [0.010 0.412] | [-0.113 0.387] | [-0.142 0.671] | [-0.265 0.395] | [-0.337 0.426] | [-0.280 0.283] |
| Observations | 2,700 | 2,700 | | 2,738 | 2,738 | |
| R-squared | 0.768 | 0.771 | | 0.673 | 0.674 | |
| Districts | 24 | 24 | | 24 | 24 | |
| Cultivated Hectares per capita | 0.210 | 0.173 | 0.405$^b$ | 0.253 | 0.258 | 0.072 |
| District-Clustered b.s. p | (0.356) | (0.592) | (0.020) | (0.236) | (0.265) | (0.272) |
| Randomization Inference tail mass | {0.048} | {0.066} | {0.000} | {0.018} | {0.018} | {0.294} |
| District-Clustered b.s. 95% CI | [-12.403 6.037] | [-28.851 6.425] | [0.034 0.764] | [-7.420 16.100] | [-14.590 7.672] | [-0.271 0.449] |
| Observations | 2,708 | 2,708 | | 2,729 | 2,729 | |
| R-squared | 0.767 | 0.770 | | 0.673 | 0.673 | |
| Districts | 24 | 24 | | 24 | 24 | |
| Savings bank credit per capita (1875) | −0.001 | 0.038 | 0.355$^a$ | −0.009 | 0.009 | 0.147$^b$ |
| District-Clustered b.s. p | (0.979) | (0.362) | (0.002) | (0.914) | (0.899) | (0.035) |
| Randomization Inference tail mass | {0.458} | {0.300} | {0.004} | {0.358} | {0.472} | {0.102} |
| District-Clustered b.s. 95% CI | [-0.082 0.233] | [-0.037 0.292] | [0.102 0.615] | [-0.079 0.142] | [-0.077 0.170] | [0.012 0.410] |
| Observations | 5,604 | 5,604 | | 6,109 | 6,109 | |
| R-squared | 0.751 | 0.752 | | 0.618 | 0.619 | |
| Districts | 35 | 35 | | 35 | 35 | |
| Savings bank credit per capita (25km) | −0.220$^b$ | −0.122$^b$ | 0.220 | −0.094$^b$ | −0.020 | 0.189$^b$ |
| District-Clustered b.s. p | (0.015) | (0.046) | (0.102) | (0.040) | (0.775) | (0.025) |
| Randomization Inference tail mass | {0.018} | {0.090} | {0.054} | {0.050} | {0.386} | {0.108} |
| District-Clustered b.s. 95% CI | [-0.506 -0.037] | [-0.443 -0.021] | [-0.041 0.430] | [-0.230 -0.009] | [-0.279 0.310] | [0.024 0.369] |
| Observations | 5,604 | 5,604 | | 6,109 | 6,109 | |
| R-squared | 0.752 | 0.752 | | 0.620 | 0.622 | |
| Districts | 35 | 35 | | 35 | 35 | |
| Pledge bank assets per capita (1896) | −0.123 | −0.128 | 0.359$^a$ | −0.004 | −0.008 | 0.149$^a$ |
| District-Clustered b.s. p | (0.191) | (0.238) | (0.001) | (0.873) | (0.652) | (0.009) |
| Randomization Inference tail mass | {0.026} | {0.022} | {0.004} | {0.450} | {0.402} | {0.102} |
| District-Clustered b.s. 95% CI | [-0.281 0.055] | [-0.307 0.084] | [0.109 0.583] | [-0.058 0.070] | [-0.044 0.041] | [0.047 0.363] |
| Observations | 5,604 | 5,604 | | 6,109 | 6,109 | |
| R-squared | 0.751 | 0.752 | | 0.618 | 0.620 | |
| Districts | 35 | 35 | | 35 | 35 | |
| Pledge bank assets per capita (25km) | −0.077 | 0.002 | 0.349$^a$ | −0.033 | 0.003 | 0.143$^a$ |
| District-Clustered b.s. p | (0.675) | (0.998) | (0.005) | (0.814) | (0.976) | (0.010) |
| Randomization Inference tail mass | {0.190} | {0.388} | {0.004} | {0.104} | {0.276} | {0.112} |
| District-Clustered b.s. 95% CI | [-1.184 1.581] | [-0.769 1.170] | [0.108 0.565] | [-0.085 0.571] | [-0.455 0.526] | [0.038 0.354] |
| Observations | 5,604 | 5,604 | | 6,109 | 6,109 | |
| R-squared | 0.751 | 0.752 | | 0.618 | 0.619 | |
| Districts | 35 | 35 | | 35 | 35 | |

*Significance levels, district-clustered b.s.*: $^a$ p<0.01, $^b$ p<0.05, $^c$ p<0.1

*Notes*: This table tests for heterogeneity on a variety of alternative dimensions, each of which is standardized to have mean zero and standard deviation one. Sample includes all municipalities in the regions of Sicily and Calabria, except for the agricultural measures, which include only Sicily. Columns (1) and (4) test for heterogeneity with respect to the listed variable. Columns (2) and (3), and (5) and (6) are from regressions looking for heterogeneity on both the basis of the listed variable and the share of employment in agricultural day labor, presenting the results for both. P-values from a wild bootstrap at the district level in parentheses. Tail mass values from a randomization inference exercise in curly braces. Square brackets show 95-percent district-clustered confidence intervals from the wild bootstrap exercise.

# C Return migration responses to the earthquake

Another possible margin of adjustment in the wake of the earthquake is return migration, which had become an increasingly important phenomenon in Sicily and Calabria by the time of the earthquake. Although we consider return migration to be a distinct phenomenon from the one we are interested in, and it is generally not part of the response of the population that actually experienced the disaster, it is certainly related and may shed light on how the earthquake changed incentives on the ground. For instance, individuals intending to return to Italy may have elected not to do so due to the damage from the earthquake. Some of the mechanisms discussed in text for a potential for the earthquake to impact migration are less relevant in the case of return migration (e.g., liquidity constraints). But others, such as disruption to productive activities, or, on the other hand, greater exigencies, are equally relevant. For example the committee in charge of aid and reconstruction efforts received an application from a migrant who asked for assistance to return and rebuild his family's home after no surviving close relatives were left to do so.

Data on return migration are only available at the province level, and an analysis based on comparing severely damaged and non-severely damaged municipalities (or districts) in the affected area is therefore not possible. Instead, we must test for differential trends in the two provinces where severe damage was concentrated—Messina and Reggio Calabria—relative to other provinces in the area. Indeed, return migration to Messina fell from 23,374 in 1908 to 5,134 in 1909; but this coincided with a national decline of 43 percent during the same period following the unusual peak in the wake of the Panic of 1907. The trends in Sicilian and Calabrian provinces are depicted in panel (a) of Figure C.1, showing a larger decline in return migration after the earthquake, particularly in Messina. The difference-in-differences coefficient corresponding to the event study in panel (b) of Figure C.1 yields a point estimate of the decline in return migration to the provinces of Messina and Reggio Calabria of about 50 percent, or about 32 percent when 1905 is omitted due to the clear pre-trend in panel (b).[81]

With return migration during the years 1909–1912 at approximately 80 percent the magnitude of emigration according to official statistics, these estimates, taken at face value, imply that the effect of the earthquake on return migration may have more than offset the negative emigration response in its effect on net migration.

[81] Note that this figure is not comparable to our benchmark specification, which compares severely to non-severely damaged municipalities within Calabria and Sicily. The equivalent difference-in-differences estimator at the province level is -0.23, though an event study reveals pre-trends that make interpreting this coefficient problematic (results available on request).

(a) Trends (b) Event Study

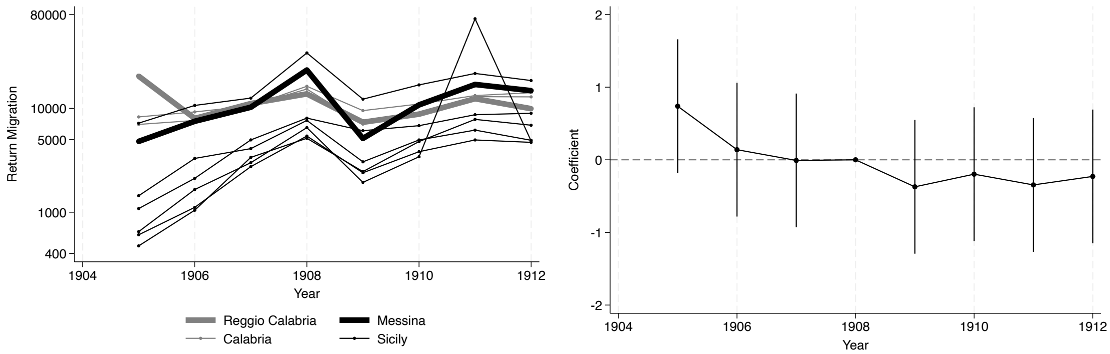



Figure C.1: Return migration patterns

*Note:* This figure studies return migration patterns for Sicily and Calabria over the period 1905–1912. The data are available at the level of the province. Panel (a) presents raw trends. Provinces in Sicily are in black and Messina is the bold line. Provinces in Calabria are in gray and Reggio Calabria is the bold line. Panel (b) performs an event study, with Messina and Reggio Calabria as the treated provinces and all others in Sicily and Calabria as controls. Standard errors in panel (b) are not adjusted in any way.

# D Adjustments for deaths and internal migration

This section addresses three main concerns. The first is that the response of the cities of Messina and Reggio Calabria to the earthquake may have been different from the response of other severely damaged municipalities. This is largely a minor concern, as these comprise only two of the many municipalities that we include in the analysis. Consistent with this expectation, the results in Tables D.1 and D.2 and Figures D.1 and D.2 are qualitatively unchanged relative to the main results. The second concern is that the use of the data on deaths, which may be inaccurate, might compromise our estimates. Figure D.3 supplements the measures that we have taken in the main text to address this concern by repeating the results of Panel A of Table 4 dropping municipalities in descending order of the official death rate until all municipalities reporting deaths are dropped. The qualitative results are largely unaffected by this exercise. Finally, there is a concern that internal migration into the major cities may have affected our estimates for the migration rates of other municipalities. To address this concern, we re-assign the internal migrants in these major cities alternatively to severely damaged or non-severely damaged municipalities according to their population shares. For instance, the 2,386 Calabria-born residents of Messina are assigned to either severely damaged or non-severely damaged municipalities of Calabria. The exception to this is that there were only two severely damaged municipalities in Sicily outside of the province of Messina, and so we assign all non-Messina Sicilians to non-severe municipalities. We do this only for post-earthquake years, effectively returning these migrants to the denominator of their municipality of origin. These results also exclude the two major cities from which the internal migrants are taken. This approach necessarily overstates the migration occurring after the earthquake as it assumes that all residents of the cities of Messina and Reggio Calabria who were not locally born immigrated after the earthquake. The results assigning individuals to severely damaged municipalities are in Tables D.3 and D.5 and Figures D.4 and D.6. The results assigning individuals to non-severely damaged municipalities are in Tables D.4 and D.6 and Figures D.5 and D.7. Although the magnitude of the estimates differs depending on this reassignment, the qualitative results are not affected.

Table D.1: Difference-in-differences results without main cities

| | Ellis Island | | | | Official Data | |
|---|---|---|---|---|---|---|
| | | | Prime-Age Only | | | |
| | (1) All | (2) Deaths | (3) All | (4) Deaths | (5) All | (6) Deaths |
| *Panel A: Sicily and Calabria* | −0.057 | −0.049 | −0.073 | −0.066 | −0.075 | −0.051 |
| District-Clustered b.s. p | (0.604) | (0.661) | (0.549) | (0.580) | (0.529) | (0.687) |
| Randomization Inference tail mass | {0.306} | {0.328} | {0.262} | {0.294} | {0.246} | {0.322} |
| District-Clustered b.s. 95% CI | [-0.419 0.390] | [-0.416 0.383] | [-0.442 0.367] | [-0.431 0.383] | [-0.306 0.127] | [-0.293 0.127] |
| Observations | 5,588 | 5,588 | 5,527 | 5,527 | 6,093 | 6,093 |
| R-squared | 0.751 | 0.751 | 0.741 | 0.741 | 0.617 | 0.616 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 |
| *Panel B: Italy* | −0.042 | −0.035 | −0.062 | −0.055 | −0.058 | −0.035 |
| District-Clustered p | (0.737) | (0.780) | (0.617) | (0.656) | (0.470) | (0.667) |
| Randomization Inference tail mass | {0.266} | {0.286} | {0.156} | {0.184} | {0.228} | {0.336} |
| District-Clustered 95% CI | [-0.290 0.205] | [-0.284 0.213] | [-0.307 0.182] | [-0.300 0.189] | [-0.215 0.100] | [-0.194 0.124] |
| Observations | 39,509 | 39,509 | 37,317 | 37,317 | 61,144 | 61,144 |
| R-squared | 0.799 | 0.799 | 0.795 | 0.795 | 0.788 | 0.788 |
| Districts | 284 | 284 | 284 | 284 | 284 | 284 |
| *Panel C: Sicily* | −0.100 | −0.099 | −0.103 | −0.102 | −0.045 | −0.030 |
| District-Clustered b.s. p | (0.602) | (0.620) | (0.635) | (0.605) | (0.752) | (0.820) |
| Randomization Inference tail mass | {0.198} | {0.198} | {0.196} | {0.196} | {0.330} | {0.372} |
| District-Clustered b.s. 95% CI | [-1.158 0.625] | [-0.758 0.563] | [-0.699 0.546] | [-0.729 0.548] | [-0.484 0.221] | [-0.429 0.225] |
| Observations | 2,700 | 2,700 | 2,670 | 2,670 | 2,807 | 2,807 |
| R-squared | 0.767 | 0.767 | 0.752 | 0.752 | 0.672 | 0.672 |
| Districts | 24 | 24 | 24 | 24 | 24 | 24 |
| *Panel D: Calabria* | −0.029 | −0.017 | −0.058 | −0.047 | −0.142 | −0.115 |
| District-Clustered b.s. p | (0.754) | (0.860) | (0.593) | (0.642) | (0.522) | (0.692) |
| Randomization Inference tail mass | {0.276} | {0.318} | {0.200} | {0.246} | {0.072} | {0.108} |
| District-Clustered b.s. 95% CI | [-0.203 0.197] | [-0.175 0.140] | [-0.220 0.131] | [-0.213 0.253] | [-0.612 0.634] | [-0.676 0.763] |
| Observations | 2,888 | 2,888 | 2,857 | 2,857 | 3,286 | 3,286 |
| R-squared | 0.733 | 0.733 | 0.728 | 0.728 | 0.514 | 0.511 |
| Districts | 11 | 11 | 11 | 11 | 11 | 11 |

*Significance levels, district-clustered b.s.*: $^a$ p<0.01, $^b$ p<0.05, $^c$ p<0.1

*Notes*: This table presents difference-in-differences coefficients for severe-times-post. P-values from a wild bootstrap clustering at the district level are in parentheses. Tail mass values from a randomization inference test are in curly braces. Ninety-five percent confidence intervals from the wild bootstrap clustered at the district level are in square braces. In panel B, the district-clustered p-values and confidence intervals are based on large-sample approximations, not the bootstrap. Columns (1), (3), and (5) use unadjusted base populations to compute migration rates; columns (2), (4), and (6) use populations adjusted for earthquake deaths. All regressions include municipality fixed effects, province-year indicators, and distance from the nearest emigration epicenter-times-post. Observations limited to 1905–1912.

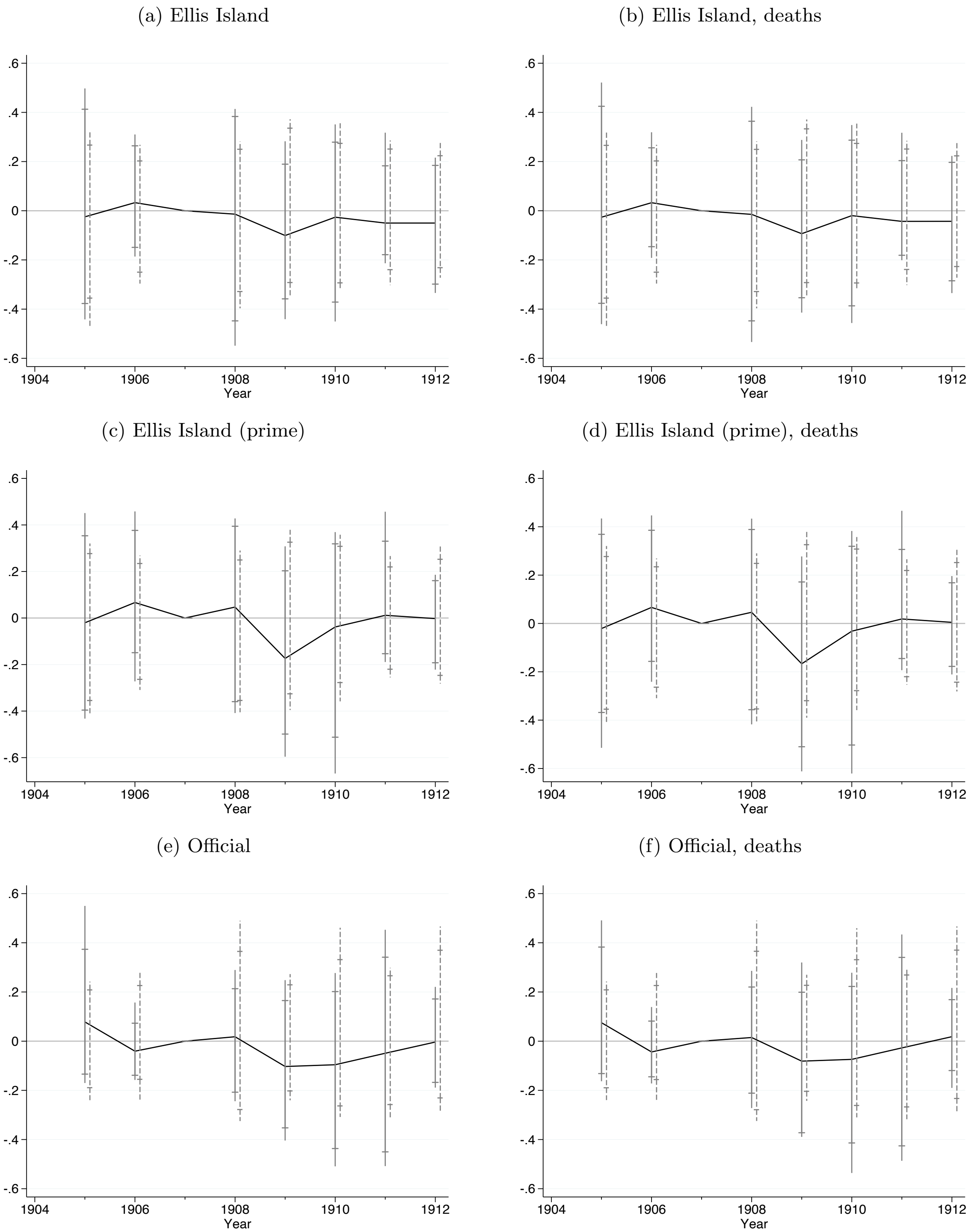



Figure D.1: Event studies for the effect of the earthquake on migration, without main cities

*Note:* Sample includes all municipalities in the regions of Sicily and Calabria except for Messina and Reggio Calabria. All event studies control for a year-specific function of distance from the nearest emigration epicenter and have 1907 as the base year. Solid bars indicate 95 percent confidence intervals from a wild bootstrap clustered on the district level. dashed bars are the middle 90 and 95 percent of results from the randomization inference exercise. The measure on the $y$-axis is the effect in logs.

## Table D.2: Difference-in-differences results without main cities

| | Ellis Island | | | | Official | | | |
|---|---|---|---|---|---|---|---|---|
| *Variables* | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Severe x Post | −0.057 | 0.009 | 0.005 | 0.013 | −0.075 | −0.036 | −0.095 | −0.074 |
| District-Clustered b.s. p | (0.593) | (0.940) | (0.981) | (0.930) | (0.503) | (0.533) | (0.443) | (0.332) |
| Randomization Inference tail mass | {0.306} | {0.496} | {0.418} | {0.406} | {0.246} | {0.412} | {0.312} | {0.284} |
| District-Clustered b.s. 95% CI | [-0.432 0.393] | [-0.175 0.313] | [-0.157 0.338] | [-0.217 0.312] | [-0.307 0.125] | [-0.192 0.088] | [-0.321 0.132] | [-0.193 0.104] |
| Severe x Post x Ag Lab Share | | $0.338^{a}$ | | $0.339^{b}$ | | $0.131^{b}$ | | 0.152 |
| District-Clustered b.s. p | | (0.009) | | (0.045) | | (0.013) | | (0.604) |
| Randomization Inference tail mass | | {0.004} | | {0.010} | | {0.112} | | {0.158} |
| District-Clustered b.s. 95% CI | | [0.080 0.585] | | [0.012 0.845] | | [0.023 0.387] | | [-0.295 0.564] |
| Severe x Post x Ag Share | | | 0.183 | −0.008 | | | $0.144^{c}$ | 0.057 |
| District-Clustered b.s. p | | | (0.200) | (0.943) | | | (0.077) | (0.676) |
| Randomization Inference tail mass | | | {0.116} | {0.376} | | | {0.148} | {0.186} |
| District-Clustered b.s. 95% CI | | | [-0.044 0.458] | [-0.335 0.372] | | | [-0.019 0.277] | [-0.455 0.309] |
| Observations | 5,588 | 5,588 | 5,588 | 5,588 | 6,093 | 6,093 | 6,093 | 6,093 |
| R-squared | 0.751 | 0.752 | 0.752 | 0.752 | 0.617 | 0.618 | 0.620 | 0.622 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 | 35 | 35 |

*Significance levels, district-clustered b.s.*: $^{a}$ p<0.01, $^{b}$ p<0.05, $^{c}$ p<0.1

*Notes*: Sample includes all municipalites in the regions of Sicily and Calabria except for Messina and Reggio Calabria. Dependent variable is unadjusted emigration rate relative to 1901 population from the source listed in the column header. P-values from a wild bootstrap clustering at the district level are in parentheses. Tail mass values from a randomization inference test are in curly braces. Ninety-five percent confidence intervals from the wild bootstrap clustered at the district level are in square braces. Ag Share is defined as the fraction of the 1901 male labor force in agriculture at the district level, and is standardized to have mean zero and standard deviation one. Ag Lab Share is the share of the 1901 male labor force in agricultural day labor (*giornalieri di campagna*) at the district level and is standardized to have mean zero and standard deviation one in the sample. All regressions include municipality fixed effects, province-year indicators, and distance from the nearest emigration epicenter-times-post. Observations limited to 1905–1912.

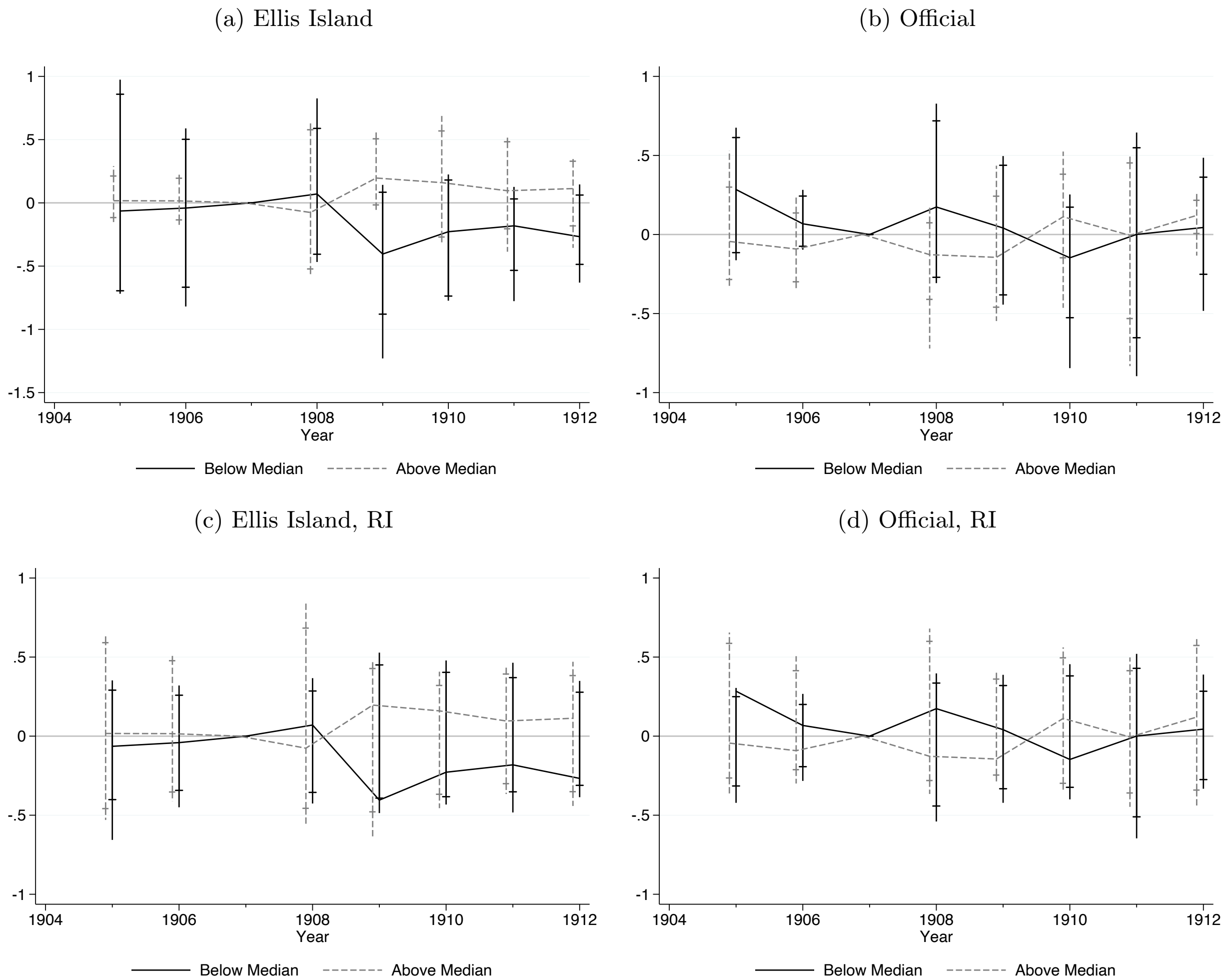



Figure D.2: Event studies divided by share of district employment in agricultural day labor without main cities

*Note:* Sample includes all municipalities in the regions of Sicily and Calabria except for Messina and Reggio Calabria. All event studies control for a year-specific function of distance from the nearest emigration epicenter and have 1907 as the base year. In panels (a) and (b), bars indicate 90- and 95-percent confidence intervals clustered on the district level, computed by a wild bootstrap. In panels (c) and (d), bars indicate 90- and 95-percent randomization inference bands. The division into "below median" and "above median" is based on the distribution of the share of employment in agricultural day labor in the sample. The measure on the $y$-axis is the effect in logs.

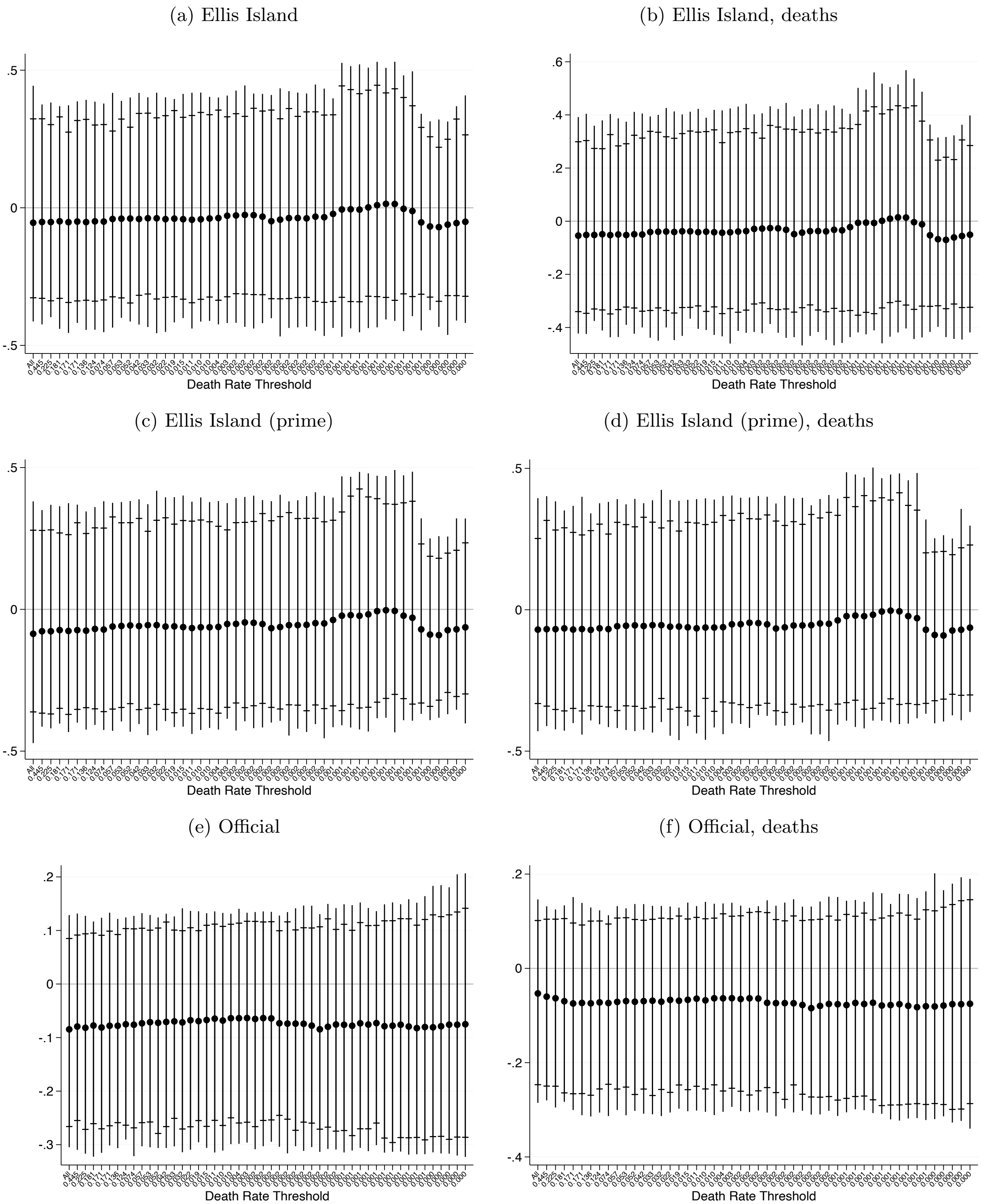



Figure D.3: Main difference-in-differences results dropping municipalities in order of death rate

*Note:* These Figures present the difference-in-differences coefficients of Panel A of Table 4, dropping municipalities in descending order of the official death toll per capita. Bars represent 90- and 95-percent confidence intervals from a wild bootstrap clustered at the district level.

Table D.3: Difference-in-differences results assigning potential internal migrants to severely damaged municipalities

| | Ellis Island | | | | Official Data | |
|---|---|---|---|---|---|---|
| | | | Prime-Age Only | | | |
| | (1) All | (2) Deaths | (3) All | (4) Deaths | (5) All | (6) Deaths |
| *Panel A: Sicily and Calabria* | −0.103 | −0.095 | −0.119 | −0.112 | −0.119 | −0.095 |
| District-Clustered b.s. p | (0.531) | (0.566) | (0.488) | (0.524) | (0.276) | (0.425) |
| Randomization Inference tail mass | {0.190} | {0.208} | {0.136} | {0.152} | {0.140} | {0.202} |
| District-Clustered b.s. 95% CI | [-0.487 0.327] | [-0.485 0.322] | [-0.509 0.303] | [-0.502 0.318] | [-0.335 0.072] | [-0.318 0.073] |
| Observations | 5,588 | 5,588 | 5,527 | 5,527 | 6,093 | 6,093 |
| R-squared | 0.751 | 0.750 | 0.741 | 0.741 | 0.618 | 0.617 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 |
| *Panel B: Italy* | −0.089 | −0.082 | −0.109 | −0.103 | −0.102 | −0.080 |
| District-Clustered p | (0.484) | (0.519) | (0.385) | (0.415) | (0.181) | (0.298) |
| Randomization Inference tail mass | {0.116} | {0.122} | {0.072} | {0.080} | {0.092} | {0.160} |
| District-Clustered 95% CI | [-0.340 0.161] | [-0.334 0.169] | [-0.356 0.138] | [-0.350 0.145] | [-0.252 0.048] | [-0.232 0.071] |
| Observations | 39,509 | 39,509 | 37,317 | 37,317 | 61,144 | 61,144 |
| R-squared | 0.799 | 0.799 | 0.795 | 0.795 | 0.788 | 0.788 |
| Districts | 284 | 284 | 284 | 284 | 284 | 284 |
| *Panel C: Sicily* | −0.177 | −0.177 | −0.180 | −0.180 | −0.122 | −0.108 |
| District-Clustered b.s. p | (0.556) | (0.572) | (0.578) | (0.560) | (0.495) | (0.498) |
| Randomization Inference tail mass | {0.088} | {0.088} | {0.074} | {0.074} | {0.156} | {0.190} |
| District-Clustered b.s. 95% CI | [-1.246 0.481] | [-0.818 0.482] | [-0.777 0.467] | [-0.785 0.464] | [-0.526 0.138] | [-0.476 0.142] |
| Observations | 2,700 | 2,700 | 2,670 | 2,670 | 2,807 | 2,807 |
| R-squared | 0.766 | 0.766 | 0.752 | 0.752 | 0.671 | 0.671 |
| Districts | 24 | 24 | 24 | 24 | 24 | 24 |
| *Panel D: Calabria* | −0.055 | −0.044 | −0.084 | −0.074 | −0.166 | −0.140 |
| District-Clustered b.s. p | (0.543) | (0.710) | (0.331) | (0.441) | (0.472) | (0.612) |
| Randomization Inference tail mass | {0.208} | {0.240} | {0.148} | {0.170} | {0.044} | {0.074} |
| District-Clustered b.s. 95% CI | [-0.226 0.162] | [-0.199 0.113] | [-0.245 0.102] | [-0.238 0.228] | [-0.636 0.607] | [-0.697 0.734] |
| Observations | 2,888 | 2,888 | 2,857 | 2,857 | 3,286 | 3,286 |
| R-squared | 0.733 | 0.733 | 0.728 | 0.728 | 0.515 | 0.512 |
| Districts | 11 | 11 | 11 | 11 | 11 | 11 |

*Significance levels, district-clustered b.s.*: [a] p<0.01, [b] p<0.05, [c] p<0.1

*Notes*: This table presents difference-in-differences coefficients for severe-times-post. P-values from a wild bootstrap clustering at the district level are in parentheses. Tail mass values from a randomization inference test are in curly braces. Ninety-five percent confidence intervals from the wild bootstrap clustered at the district level are in square braces. In panel B, district-clustered statistics are from large-sample approximations, not the bootstrap. Columns (1), (3), and (5) use unadjusted base populations to compute migration rates; columns (2), (4), and (6) use populations adjusted for earthquake deaths. All regressions include municipality fixed effects, province-year indicators, and distance from the nearest emigration epicenter-times-post. Observations limited to 1905–1912.

Table D.4: Difference-in-differences results assigning potential internal migrants to non-severely damaged municipalities

| | Ellis Island | | | | Official Data | |
|---|---|---|---|---|---|---|
| | | | Prime-Age Only | | | |
| | (1) All | (2) Deaths | (3) All | (4) Deaths | (5) All | (6) Deaths |
| *Panel A: Sicily and Calabria* | −0.029 | −0.021 | −0.045 | −0.038 | −0.047 | −0.023 |
| District-Clustered b.s. p | (0.762) | (0.841) | (0.667) | (0.711) | (0.701) | (0.828) |
| Randomization Inference tail mass | {0.392} | {0.418} | {0.346} | {0.364} | {0.330} | {0.406} |
| District-Clustered b.s. 95% CI | [-0.396 0.418] | [-0.399 0.411] | [-0.407 0.396] | [-0.403 0.409] | [-0.297 0.156] | [-0.288 0.156] |
| Observations | 5,588 | 5,588 | 5,527 | 5,527 | 6,093 | 6,093 |
| R-squared | 0.751 | 0.750 | 0.741 | 0.741 | 0.617 | 0.616 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 |
| *Panel B: Italy* | −0.015 | −0.008 | −0.034 | −0.027 | −0.030 | −0.007 |
| District-Clustered p | (0.909) | (0.953) | (0.786) | (0.828) | (0.720) | (0.931) |
| Randomization Inference tail mass | {0.394} | {0.440} | {0.270} | {0.300} | {0.360} | {0.496} |
| District-Clustered 95% CI | [-0.265 0.236] | [-0.259 0.244] | [-0.282 0.213] | [-0.276 0.221] | [-0.197 0.137] | [-0.177 0.162] |
| Observations | 39,509 | 39,509 | 37,317 | 37,317 | 61,144 | 61,144 |
| R-squared | 0.799 | 0.799 | 0.795 | 0.795 | 0.788 | 0.788 |
| Districts | 284 | 284 | 284 | 284 | 284 | 284 |
| *Panel C: Sicily* | −0.075 | −0.074 | −0.078 | −0.077 | −0.020 | −0.005 |
| District-Clustered b.s. p | (0.663) | (0.679) | (0.685) | (0.653) | (0.892) | (0.981) |
| Randomization Inference tail mass | {0.256} | {0.256} | {0.256} | {0.258} | {0.414} | {0.464} |
| District-Clustered b.s. 95% CI | [-1.128 0.649] | [-0.732 0.589] | [-0.674 0.574] | [-0.706 0.573] | [-0.475 0.248] | [-0.419 0.252] |
| Observations | 2,700 | 2,700 | 2,670 | 2,670 | 2,807 | 2,807 |
| R-squared | 0.767 | 0.767 | 0.752 | 0.752 | 0.672 | 0.672 |
| Districts | 24 | 24 | 24 | 24 | 24 | 24 |
| *Panel D: Calabria* | 0.001 | 0.012 | −0.028 | −0.017 | −0.113 | −0.086 |
| District-Clustered b.s. p | (0.995) | (0.918) | (0.791) | (0.832) | (0.668) | (0.694) |
| Randomization Inference tail mass | {0.402} | {0.454} | {0.310} | {0.358} | {0.112} | {0.172} |
| District-Clustered b.s. 95% CI | [-0.188 0.193] | [-0.173 0.188] | [-0.212 0.173] | [-0.206 0.299] | [-0.606 0.685] | [-0.685 0.821] |
| Observations | 2,888 | 2,888 | 2,857 | 2,857 | 3,286 | 3,286 |
| R-squared | 0.733 | 0.732 | 0.728 | 0.728 | 0.514 | 0.511 |
| Districts | 11 | 11 | 11 | 11 | 11 | 11 |

*Significance levels, district-clustered b.s.*: $^a$ p<0.01, $^b$ p<0.05, $^c$ p<0.1

*Notes*: This table presents difference-in-differences coefficients for severe-times-post. P-values from a wild bootstrap clustering at the district level are in parentheses. Tail mass values from a randomization inference test are in curly braces. Ninety-five percent confidence intervals from the wild bootstrap clustered at the district level are in square braces. In panel B, district-clustered statistics are from large-sample approximations, not the bootstrap. Columns (1), (3), and (5) use unadjusted base populations to compute migration rates; columns (2), (4), and (6) use populations adjusted for earthquake deaths. All regressions include municipality fixed effects, province-year indicators, and distance from the nearest emigration epicenter-times-post. Observations limited to 1905–1912.

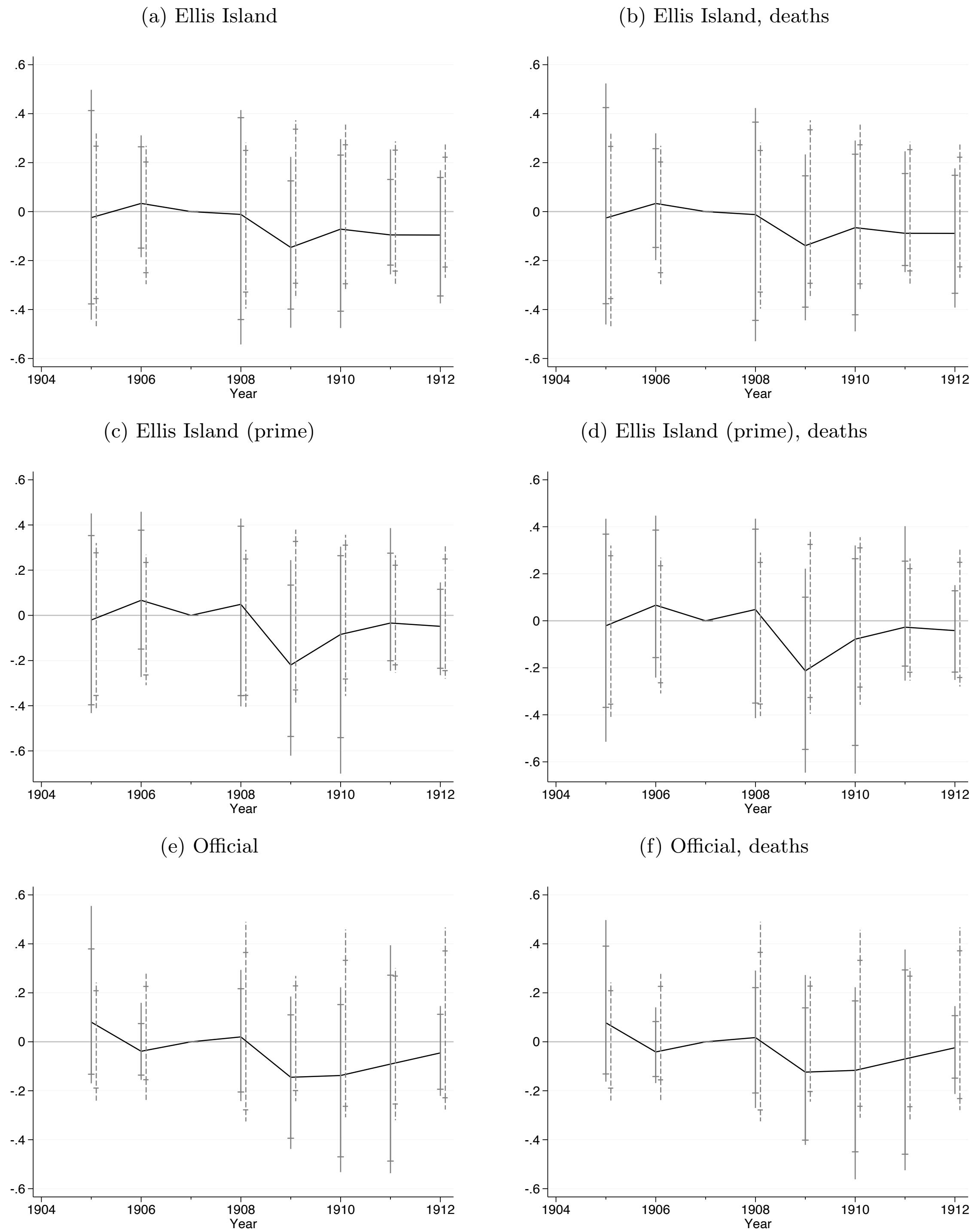
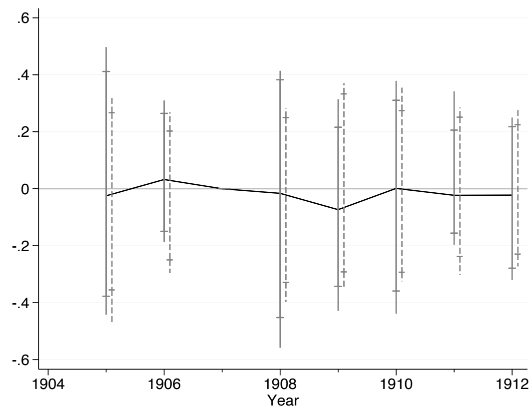


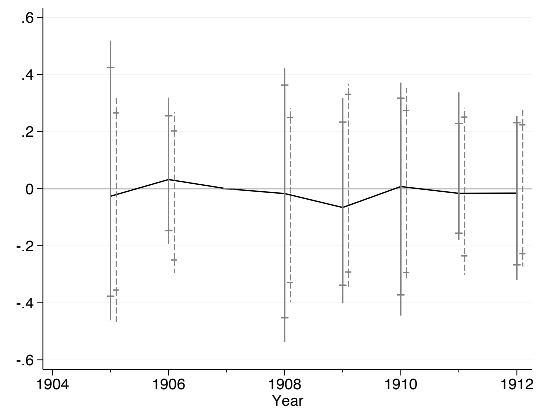

Figure D.4: Event studies for the effect of the earthquake on migration, assigning potential internal migrants to severely damaged municipalities

*Note:* Sample includes all municipalities in the regions of Sicily and Calabria. All event studies control for a year-specific function of distance from the nearest emigration epicenter and have 1907 as the base year. Solid bars indicate 95 percent confidence intervals from a wild bootstrap clustered on the district level. dashed bars are the middle 90 and 95 percent of results from the randomization inference exercise. The measure on the $y$-axis is the effect in logs.

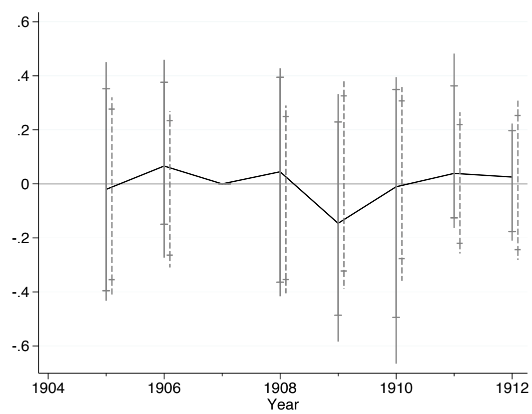
(a) Ellis Island



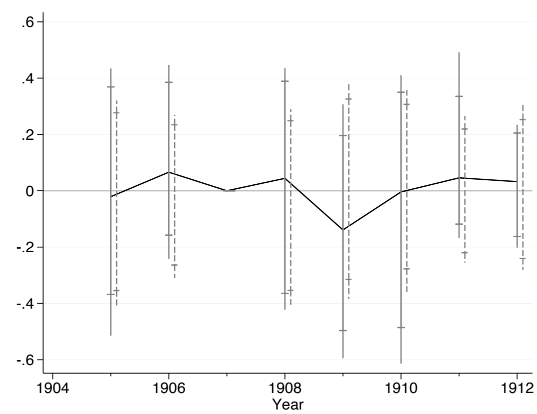

(b) Ellis Island, deaths



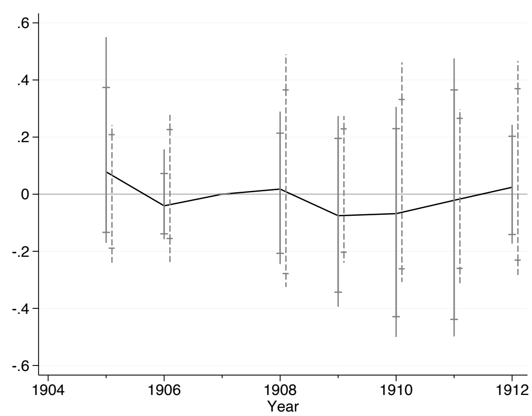

(c) Ellis Island (prime)




(d) Ellis Island (prime), deaths




(e) Official




(f) Official, deaths

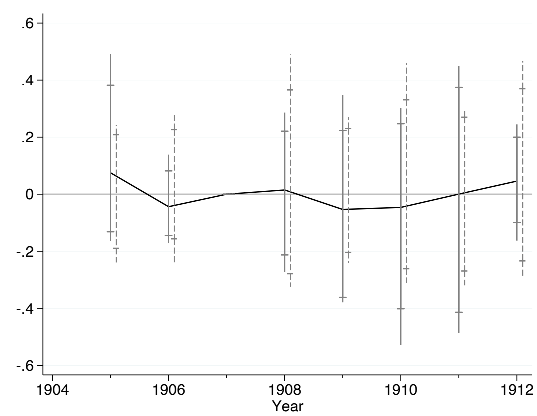



Figure D.5: Event studies for the effect of the earthquake on migration, assigning potential internal migrants to non-severely damaged municipalities

*Note:* Sample includes all municipalities in the regions of Sicily and Calabria. All event studies control for a year-specific function of distance from the nearest emigration epicenter and have 1907 as the base year. Solid bars indicate 95 percent confidence intervals from a wild bootstrap clustered on the district level. dashed bars are the middle 90 and 95 percent of results from the randomization inference exercise. The measure on the $y$-axis is the effect in logs.

Table D.5: Difference-in-differences results assigning potential internal migrants to severely damaged municipalities

| | Ellis Island | | | | Official | | | |
|---|---|---|---|---|---|---|---|---|
| *Variables* | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Severe x Post | −0.103 | −0.036 | −0.039 | −0.031 | −0.119 | −0.078 | −0.136 | −0.116 |
| District-Clustered b.s. p | (0.539) | (0.693) | (0.669) | (0.733) | (0.241) | (0.132) | (0.202) | (0.121) |
| Randomization Inference tail mass | {0.190} | {0.372} | {0.458} | {0.460} | {0.140} | {0.272} | {0.184} | {0.148} |
| District-Clustered b.s. 95% CI | [-0.504 0.331] | [-0.226 0.245] | [-0.192 0.256] | [-0.262 0.251] | [-0.333 0.069] | [-0.222 0.020] | [-0.347 0.066] | [-0.224 0.038] |
| Severe x Post x Ag Lab Share | | 0.345$^{a}$ | | 0.336$^{c}$ | | 0.138$^{a}$ | | 0.150 |
| District-Clustered b.s. p | | (0.009) | | (0.061) | | (0.007) | | (0.554) |
| Randomization Inference tail mass | | {0.004} | | {0.012} | | {0.106} | | {0.162} |
| District-Clustered b.s. 95% CI | | [0.065 0.581] | | [-0.068 0.886] | | [0.038 0.361] | | [-0.243 0.548] |
| Severe x Post x Ag Share | | | 0.190 | −0.000 | | | 0.150$^{b}$ | 0.064 |
| District-Clustered b.s. p | | | (0.213) | (0.998) | | | (0.048) | (0.634) |
| Randomization Inference tail mass | | | {0.114} | {0.358} | | | {0.144} | {0.174} |
| District-Clustered b.s. 95% CI | | | [-0.047 0.444] | [-0.365 0.436] | | | [0.002 0.271] | [-0.420 0.292] |
| Observations | 5,588 | 5,588 | 5,588 | 5,588 | 6,093 | 6,093 | 6,093 | 6,093 |
| R-squared | 0.751 | 0.752 | 0.752 | 0.752 | 0.618 | 0.618 | 0.621 | 0.622 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 | 35 | 35 |

*Significance levels, district-clustered b.s.*: $^{a}$ p<0.01, $^{b}$ p<0.05, $^{c}$ p<0.1
*Notes*: Sample includes all municipalities in the regions of Sicily and Calabria. Dependent variable is unadjusted emigration rate relative to 1901 population from the source listed in the column header. P-values from a wild bootstrap clustering at the district level are in parentheses. Tail mass values from a randomization inference test are in curly braces. Ninety-five percent confidence intervals from the wild bootstrap clustered at the district level are in square braces. Ag Share is defined as the fraction of the 1901 male labor force in agriculture at the district level, and is standardized to have mean zero and standard deviation one. Ag Lab Share is the share of the 1901 male labor force in agricultural day labor (*giornalieri di campagna*) at the district level and is standardized to have mean zero and standard deviation one in the sample. All regressions include commune fixed effects, province-year indicators, and log distance from the nearest emigration epicenter-times-post. Observations limited to 1905–1912.

Table D.6: Difference-in-differences results assigning potential internal migrants to non-severely damaged municipalities

| | Ellis Island | | | | Official | | | |
|---|---|---|---|---|---|---|---|---|
| *Variables* | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Severe x Post | −0.029 | 0.038 | 0.033 | 0.041 | −0.047 | −0.007 | −0.067 | −0.046 |
| District-Clustered b.s. p | (0.756) | (0.785) | (0.852) | (0.776) | (0.661) | (0.899) | (0.656) | (0.517) |
| Randomization Inference tail mass | {0.392} | {0.406} | {0.352} | {0.340} | {0.330} | {0.500} | {0.410} | {0.412} |
| District-Clustered b.s. 95% CI | [-0.408 0.420] | [-0.157 0.333] | [-0.146 0.363] | [-0.195 0.333] | [-0.299 0.151] | [-0.185 0.119] | [-0.315 0.156] | [-0.177 0.136] |
| Severe x Post x Ag Lab Share | | 0.343$^{a}$ | | 0.350$^{c}$ | | 0.136$^{b}$ | | 0.165 |
| District-Clustered b.s. p | | (0.008) | | (0.053) | | (0.014) | | (0.621) |
| Randomization Inference tail mass | | {0.004} | | {0.008} | | {0.112} | | {0.146} |
| District-Clustered b.s. 95% CI | | [0.089 0.586] | | [-0.023 0.848] | | [0.023 0.402] | | [-0.302 0.593] |
| Severe x Post x Ag Share | | | 0.185 | −0.013 | | | 0.145$^{c}$ | 0.050 |
| District-Clustered b.s. p | | | (0.178) | (0.897) | | | (0.085) | (0.715) |
| Randomization Inference tail mass | | | {0.116} | {0.400} | | | {0.142} | {0.186} |
| District-Clustered b.s. 95% CI | | | [-0.040 0.458] | [-0.333 0.373] | | | [-0.025 0.279] | [-0.482 0.306] |
| Observations | 5,588 | 5,588 | 5,588 | 5,588 | 6,093 | 6,093 | 6,093 | 6,093 |
| R-squared | 0.751 | 0.752 | 0.751 | 0.752 | 0.617 | 0.618 | 0.620 | 0.622 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 | 35 | 35 |

*Significance levels, district-clustered b.s.*: $^{a}$ p<0.01, $^{b}$ p<0.05, $^{c}$ p<0.1
*Notes*: Sample includes all municipalities in the regions of Sicily and Calabria. Dependent variable is unadjusted emigration rate relative to 1901 population from the source listed in the column header. P-values from a wild bootstrap clustering at the district level are in parentheses. Tail mass values from a randomization inference test are in curly braces. Ninety-five percent confidence intervals from the wild bootstrap clustered at the district level are in square braces. Ag Share is defined as the fraction of the 1901 male labor force in agriculture at the district level, and is standardized to have mean zero and standard deviation one. Ag Lab Share is the share of the 1901 male labor force in agricultural day labor (*giornalieri di campagna*) at the district level and is standardized to have mean zero and standard deviation one in the sample. All regressions include commune fixed effects, province-year indicators, and log distance from the nearest emigration epicenter-times-post. Observations limited to 1905–1912.

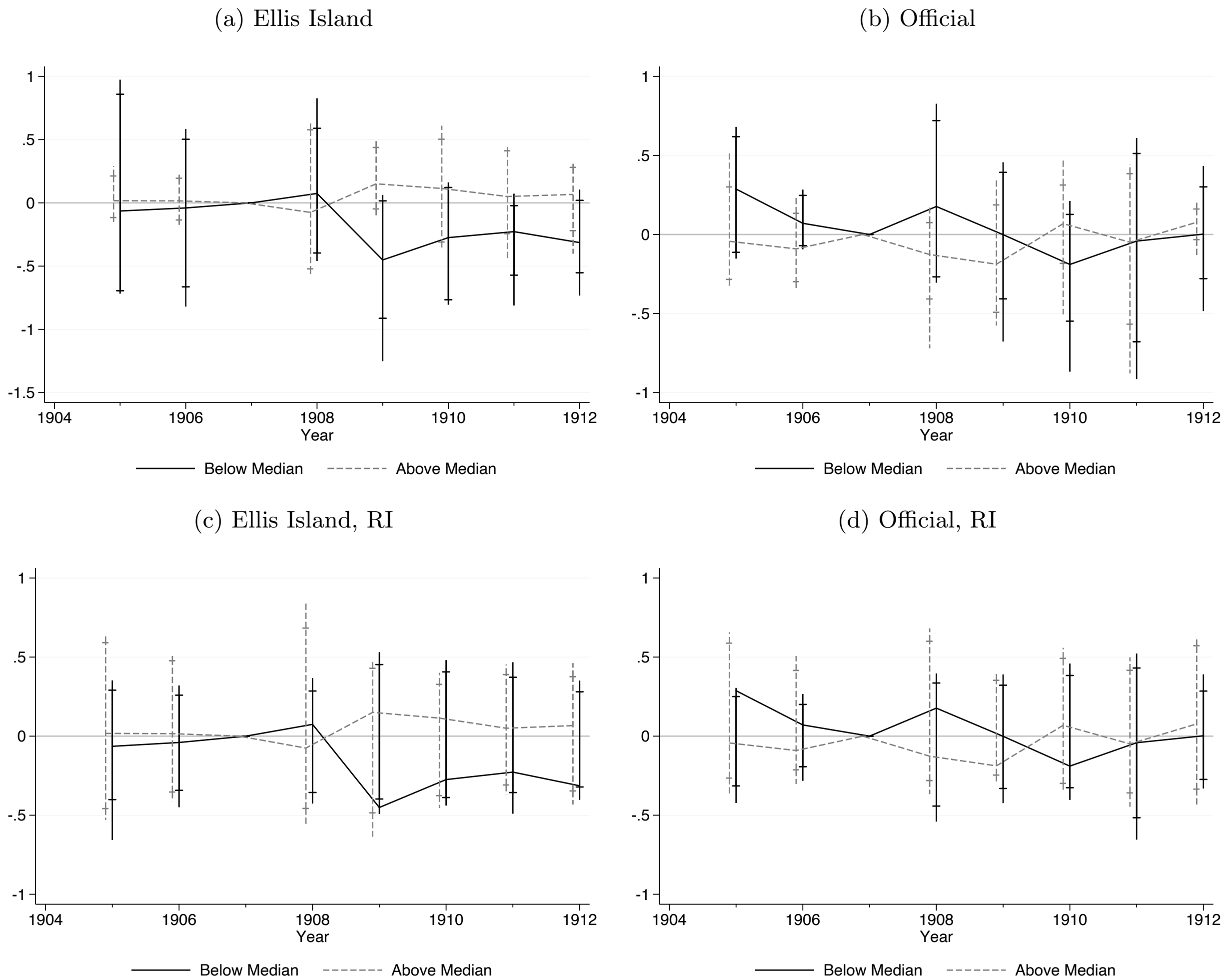



Figure D.6: Event studies divided by share of district employment in agricultural day labor, assigning potential internal migrants to severely damaged municipalities

*Note:* Sample includes all municipalities in the regions of Sicily and Calabria. All event studies control for a year-specific function of distance from the nearest emigration epicenter and have 1907 as the base year. In panels (a) and (b), bars indicate 90- and 95-percent confidence intervals clustered on the district level, computed by a wild bootstrap. In panels (c) and (d), bars indicate the middle 90 and 95 percent of estimates from the randomization inference exercise. The division into "below median" and "above median" is based on the distribution of the share of employment in agricultural day labor in the sample. The measure on the $y$-axis is the effect in logs.

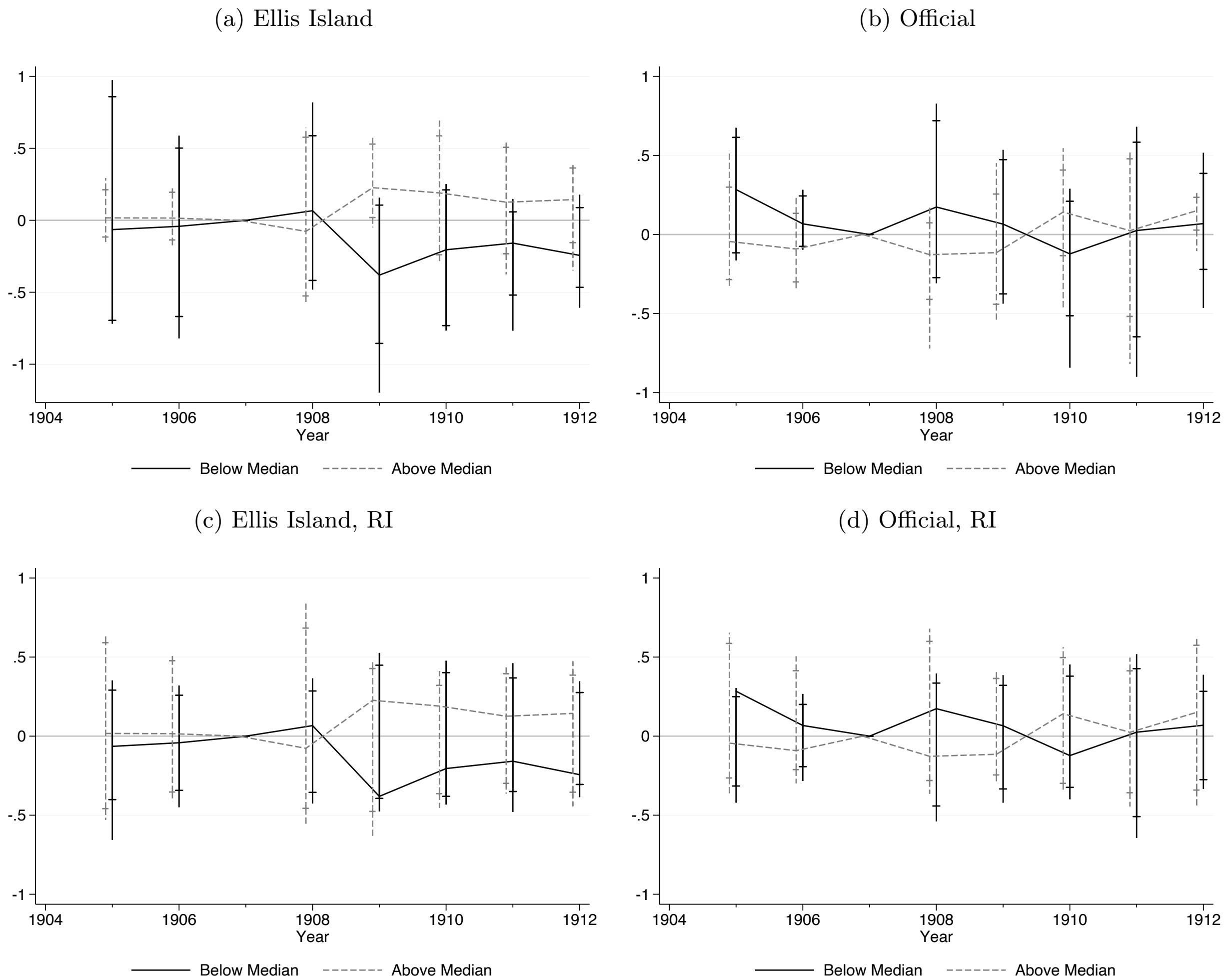



Figure D.7: Event studies divided by share of district employment in agricultural day labor, assigning potential internal migrants to non-severely damaged municipalities

*Note:* Sample includes all municipalities in the regions of Sicily and Calabria. All event studies control for a year-specific function of distance from the nearest emigration epicenter and have 1907 as the base year. In panels (a) and (b), bars indicate 90- and 95-percent confidence intervals clustered on the district level, computed by a wild bootstrap. In panels (c) and (d), bars indicate the middle 90 and 95 percent of estimates from the randomization inference exercise. The division into "below median" and "above median" is based on the distribution of the share of employment in agricultural day labor in the sample. The measure on the $y$-axis is the effect in logs.

# E Additional details of data cleaning

The data that we created by geo-locating Ellis Island passengers to modern municipalities required standardization before analysis. The main issue in this case was the fact that some individuals were assigned to municipalities that were created after 1901, when our population data were taken.[82] Without any correction, it would not be possible to compute emigration rates for these municipalities because of the lack of a denominator. To address this issue, we manually scoured a variety of sources, such as the *Comuni e Loro Poplazione ai Censimenti dal 1861 al 1951*, published by ISTAT in 1960, which lists border histories to 1951. (In cases where municipalities were founded after 1951, other sources were consulted, such as official websites of each municipality.) These sources enabled us to determine which historic municipality these new municipalities were a part of in the early 20th century. We can then simply use the population of that historic municipality as the denominator and the total number of emigrants from any component of the historic municipality as the numerator.

For the official statistics, the difficulty was in geo-locating municipalities that no longer exist because they were merged into modern municipalities. If the municipalities no longer exist, we do not have the ability to determine their severity, distance from the epicenter, or distance from the nearest emigration epicenter. To address this issue, we used the *Comuni e Loro Popolazione* and related sources to determine the municipality into which they were combined. We then used the centroid of the modern municipality (ISTAT 2018) as the location of all historic municipalities now making up the modern municipality. This approach is necessarily imperfect, but is the best approach available to us given the absence (to our knowledge) of a map of historic municipalities.

[82]This would occur if an individual listed a village within a municipality as his last place of residence and this village later separated to become a separate municipality.

# F Details of the randomization inference method

In addition to reporting $p$-values and confidence intervals from the Roodman et al. (2019) wild bootstrap in the main results, we also include results from a randomization inference exercise. We conduct this exercise as follows.

1. We simulated 500 random earthquakes in the study region for the regression in question. That is, we randomly choose 500 points to assign as epicenters for our counterfactual earthquakes. For the main results, we simulate these 500 random earthquakes in Sicily and Calabria.

2. For the true earthquake, we estimate a probit relationship relating distance to the earthquake epicenter to the probability of experiencing severe damage.

3. Using the estimates from step 2 and the distance from each municipality to each simulated earthquake, we determined the estimated probability that the municipality would experience severe damage. Based on these probabilities, we randomly assigned each municipality to experience severe or non-severe damage from each simulated earthquake.

4. We estimate the regression equation of interest for each simulated earthquake.

5. We determine the fraction of the distribution of counterfactual estimates in the (one-sided) tail beyond the true estimate. That is, if the true estimate is below the median of the counterfactual estimates, we report the share below, and vice versa. For the results of Panel A of Table 4, an example of this is presented in Figure F.1

When applying the randomization inference to an alternative measure of earthquake damage, we adjust step 2 accordingly. For instance, when we use the Mercalli severity score, we estimate a regression of the severity score on distance from the earthquake epicenter for the true earthquake and use these estimates, together with the mean-squared error of the regression, to randomly assign severity scores for each municipality and each counterfactual earthquake.[83] When applying this exercise to district-level analysis, we

[83] That is, we estimate

$$s_i = \alpha + \beta d_i + \varepsilon_i,$$

where $s_i$ is the Mercalli severity score and $d_i$ is the municipality's distance from the earthquake epicenter. We then assign the counterfactual Mercalli scores as

$$\hat{s}_{ik} = \hat{\alpha} + \hat{\beta} d_i^k + \hat{\sigma} u_{ik},$$

where $\hat{s}_{ik}$ is the estimated severity in municipality $i$ for counterfactual earthquake $k$, $d_i^k$ is the distance from municipality $i$ to the epicenter of counterfactual earthquake $k$, $\hat{\sigma}$ is the estimated root-mean-squared error of the regression, and $u_{ik}$ is a random draw from the standard normal distribution.

assign the 5 closest districts to each counterfactual earthquake as severely damaged rather than using random assignment.

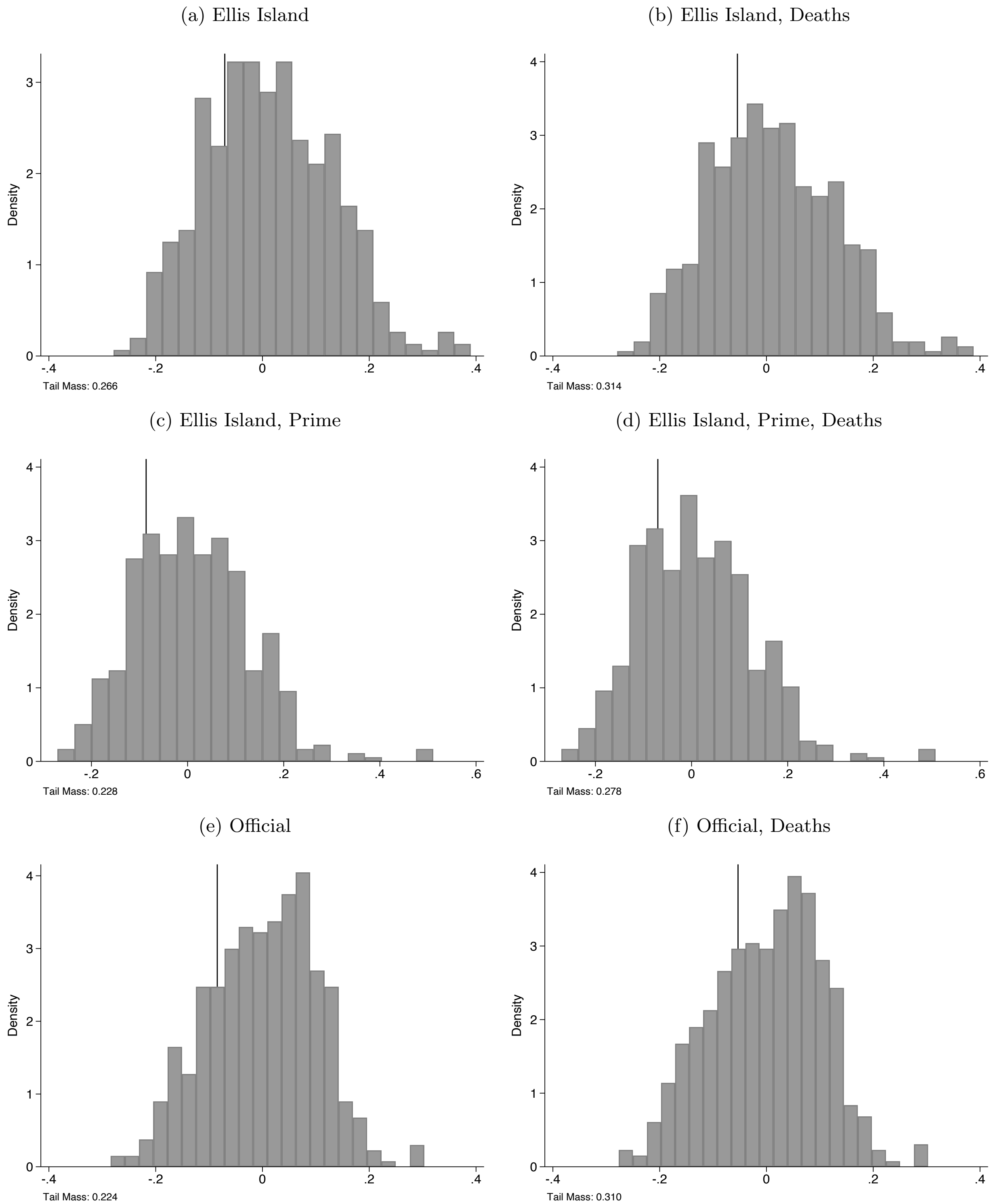



Figure F.1: These histograms present the distribution of results for the main difference-in-differences results of panel A of Table 4 for the randomization inference permutation. The actual estimate is marked by the vertical line. The tail mass is the mass more extreme (in this case, more negative) than the estimate.

# G Alternative severity definitions

This section repeats the main results with severely damaged municipalities defined as those experiencing a Mercalli score of at least VII (marked in Figure G.1. Municipalities with no severity data are assigned to the non-severe group. The results are largely similar to those of the main text, though in some ways they are stronger. Table G.1 determines the magnitude of the effect of the earthquake. The magnitudes of the estimates are slightly larger than those of Table 4 and in many cases the randomization inference results point to a significant result. But the magnitude of the coefficients remain small and the confidence intervals are tighter, allowing us to be more certain in this conclusion. Figure G.2 performs event study analyses. There are somewhat stronger pre-trends in these results than in Figure 5, but otherwise the main results are qualitatively unaffected, continuing to show a short-lived decline. Table G.2 estimates regressions with interactions with the agricultural labor share. Although the coefficients are not as strongly statistically significant as in the main results, the pattern of results is largely unchanged. Finally, Figure G.3 presents event studies dividing the sample by above- and below-median districts in agricultural labor, finding even stronger evidence of heterogeneous responses in this dimension than in Figure 7.

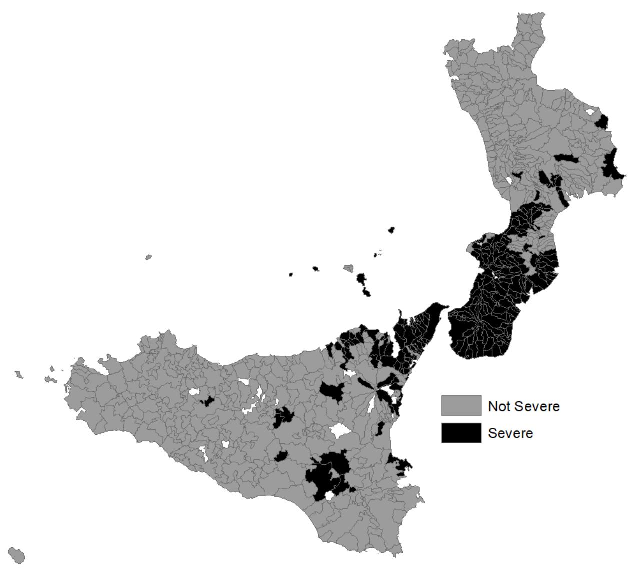



Figure G.1: Municipality-level severity indicator with a Mercalli score cutoff of VII

*Note:* See text

Table G.1: Difference-in-differences results with severity cutoff VII

| | Ellis Island | | | | Official Data | |
|---|---|---|---|---|---|---|
| | | | Prime-Age Only | | | |
| | (1) All | (2) Deaths | (3) All | (4) Deaths | (5) All | (6) Deaths |
| *Panel A: Sicily and Calabria* | −0.100 | −0.097 | −0.150 | −0.147 | −0.124 | −0.118 |
| District-Clustered b.s. p | (0.384) | (0.415) | (0.246) | (0.249) | (0.428) | (0.436) |
| Randomization Inference tail mass | {0.098} | {0.100} | {0.032} | {0.032} | {0.092} | {0.100} |
| District-Clustered b.s. 95% CI | [-0.350 0.149] | [-0.351 0.162] | [-0.398 0.108] | [-0.401 0.105] | [-0.311 0.113] | [-0.313 0.116] |
| Observations | 5,604 | 5,604 | 5,543 | 5,543 | 6,109 | 6,109 |
| R-squared | 0.751 | 0.751 | 0.742 | 0.742 | 0.619 | 0.618 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 |
| *Panel B: Italy* | −0.081 | −0.078 | −0.132 | −0.129 | −0.100 | −0.094 |
| District-Clustered p | (0.410) | (0.422) | (0.180) | (0.186) | (0.266) | (0.298) |
| Randomization Inference tail mass | {0.094} | {0.102} | {0.028} | {0.028} | {0.054} | {0.070} |
| District-Clustered 95% CI | [-0.273 0.112] | [-0.267 0.112] | [-0.326 0.062] | [-0.321 0.062] | [-0.276 0.077] | [-0.272 0.084] |
| Observations | 39,525 | 39,525 | 37,333 | 37,333 | 61,160 | 61,160 |
| R-squared | 0.799 | 0.799 | 0.795 | 0.795 | 0.788 | 0.788 |
| Districts | 284 | 284 | 284 | 284 | 284 | 284 |
| *Panel C: Sicily* | −0.165 | −0.161 | −0.238 | −0.235 | −0.057 | −0.051 |
| District-Clustered b.s. p | (0.367) | (0.358) | (0.157) | (0.144) | (0.555) | (0.610) |
| Randomization Inference tail mass | {0.050} | {0.050} | {0.008} | {0.010} | {0.204} | {0.244} |
| District-Clustered b.s. 95% CI | [-0.592 0.238] | [-0.568 0.210] | [-0.639 0.114] | [-0.647 0.111] | [-0.290 0.184] | [-0.273 0.203] |
| Observations | 2,708 | 2,708 | 2,678 | 2,678 | 2,815 | 2,815 |
| R-squared | 0.767 | 0.767 | 0.753 | 0.753 | 0.672 | 0.672 |
| Districts | 24 | 24 | 24 | 24 | 24 | 24 |
| *Panel D: Calabria* | −0.032 | −0.031 | −0.061 | −0.060 | −0.216 | −0.213 |
| District-Clustered b.s. p | (0.921) | (0.928) | (0.876) | (0.899) | (0.519) | (0.502) |
| Randomization Inference tail mass | {0.132} | {0.134} | {0.062} | {0.064} | {0.002} | {0.002} |
| District-Clustered b.s. 95% CI | [-0.866 5.175] | [-0.864 4.685] | [-1.162 11.802] | [-1.006 11.369] | [-1.027 0.249] | [-1.174 0.252] |
| Observations | 2,896 | 2,896 | 2,865 | 2,865 | 3,294 | 3,294 |
| R-squared | 0.733 | 0.733 | 0.728 | 0.728 | 0.517 | 0.514 |
| Districts | 11 | 11 | 11 | 11 | 11 | 11 |

*Significance levels, district-clustered b.s.*: $^a$ p<0.01, $^b$ p<0.05, $^c$ p<0.1

*Notes*: This table presents difference-in-differences coefficients for severe-times-post. P-values from a wild bootstrap clustering at the district level are in parentheses. Tail mass values from a randomization inference test are in curly braces. Ninety-five percent confidence intervals from the wild bootstrap clustered at the district level are in square braces. In panel B, district-clustered statistics are based on large-sample approximations, not the bootstrap. Columns (1), (3), and (5) use unadjusted base populations to compute migration rates; columns (2), (4), and (6) use populations adjusted for earthquake deaths. All regressions include municipality fixed effects, province-year indicators, and distance from the nearest emigration epicenter-times-post. Observations limited to 1905–1912.

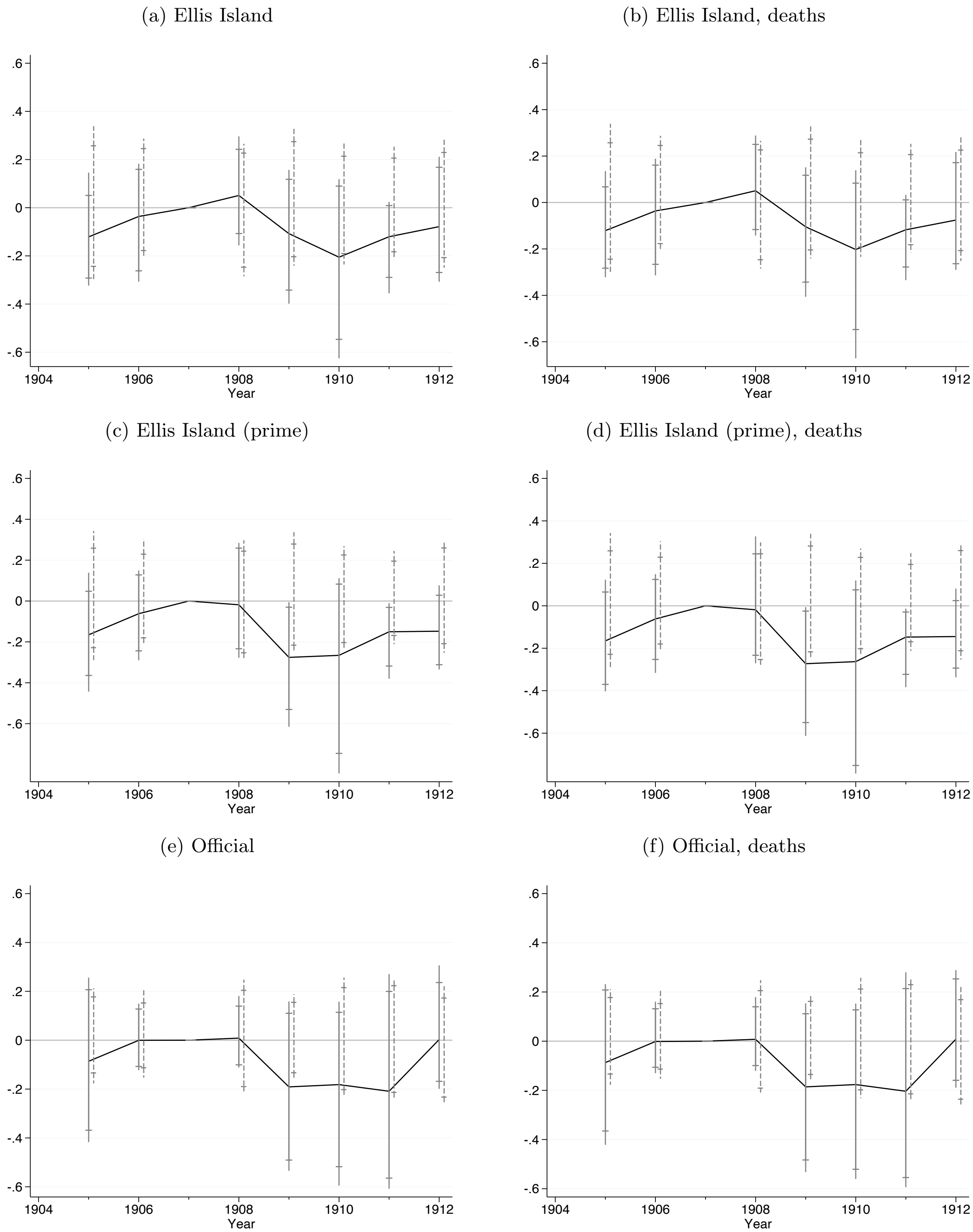



Figure G.2: Event studies for the effect of the earthquake on migration with severity cutoff VII

*Note:* Sample includes all municipalities in the regions of Sicily and Calabria. All event studies control for a year-specific function of distance from the nearest emigration epicenter and have 1907 as the base year. Solid bars are 90- and 95-percent confidence intervals from a wild bootstrap clustered on the district level. dashed bars are the middle 90 and 95 percent of randomization inference replications. The measure on the $y$-axis is the effect in logs.

Table G.2: Heterogeneous responses to the earthquake with severity cutoff VII

| | Ellis Island | | | | Official | | | |
|---|---|---|---|---|---|---|---|---|
| *Variables* | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Severe x Post | −0.100 | −0.109 | −0.072 | −0.096 | −0.124 | −0.120 | −0.125 | −0.122 |
| District-Clustered b.s. p | (0.397) | (0.239) | (0.519) | (0.314) | (0.434) | (0.211) | (0.445) | (0.184) |
| Randomization Inference tail mass | {0.198} | {0.176} | {0.382} | {0.306} | {0.120} | {0.144} | {0.222} | {0.136} |
| District-Clustered b.s. 95% CI | [-0.336 0.149] | [-0.331 0.081] | [-0.303 0.149] | [-0.305 0.109] | [-0.311 0.106] | [-0.240 0.083] | [-0.323 0.090] | [-0.227 0.065] |
| Severe x Post x Ag Lab Share | | $0.242^c$ | | 0.183 | | $0.131^b$ | | 0.156 |
| District-Clustered b.s. p | | (0.065) | | (0.140) | | (0.025) | | (0.384) |
| Randomization Inference tail mass | | {0.038} | | {0.170} | | {0.114} | | {0.156} |
| District-Clustered b.s. 95% CI | | [-0.024 0.469] | | [-0.098 0.635] | | [0.013 0.405] | | [-0.165 0.497] |
| Severe x Post x Ag Share | | | 0.176 | 0.070 | | | $0.137^b$ | 0.045 |
| District-Clustered b.s. p | | | (0.174) | (0.473) | | | (0.016) | (0.688) |
| Randomization Inference tail mass | | | {0.122} | {0.200} | | | {0.164} | {0.198} |
| District-Clustered b.s. 95% CI | | | [-0.046 0.334] | [-0.137 0.304] | | | [0.032 0.230] | [-0.145 0.251] |
| Observations | 5,604 | 5,604 | 5,604 | 5,604 | 6,109 | 6,109 | 6,109 | 6,109 |
| R-squared | 0.751 | 0.752 | 0.752 | 0.752 | 0.619 | 0.620 | 0.622 | 0.624 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 | 35 | 35 |

*Significance levels, district-clustered b.s.*: $^a$ p<0.01, $^b$ p<0.05, $^c$ p<0.1
*Notes*: Sample includes all municipalities in the regions of Sicily and Calabria. Dependent variable is unadjusted emigration rate relative to 1901 population from the source listed in the column header. P-values from a wild bootstrap clustering at the district level are in parentheses. Tail mass values from a randomization inference test are in curly braces. Ninety-five percent confidence intervals from the wild bootstrap clustered at the district level are in square braces. Ag Share is defined as the fraction of the 1901 male labor force in agriculture at the district level, and is standardized to have mean zero and standard deviation one. Ag Lab Share is the share of the 1901 male labor force in agricultural day labor (*giornalieri di campagna*) at the district level and is standardized to have mean zero and standard deviation one in the sample. All regressions include commune fixed effects, province-year indicators, and log distance from the nearest emigration epicenter-times-post. Observations limited to 1905–1912.

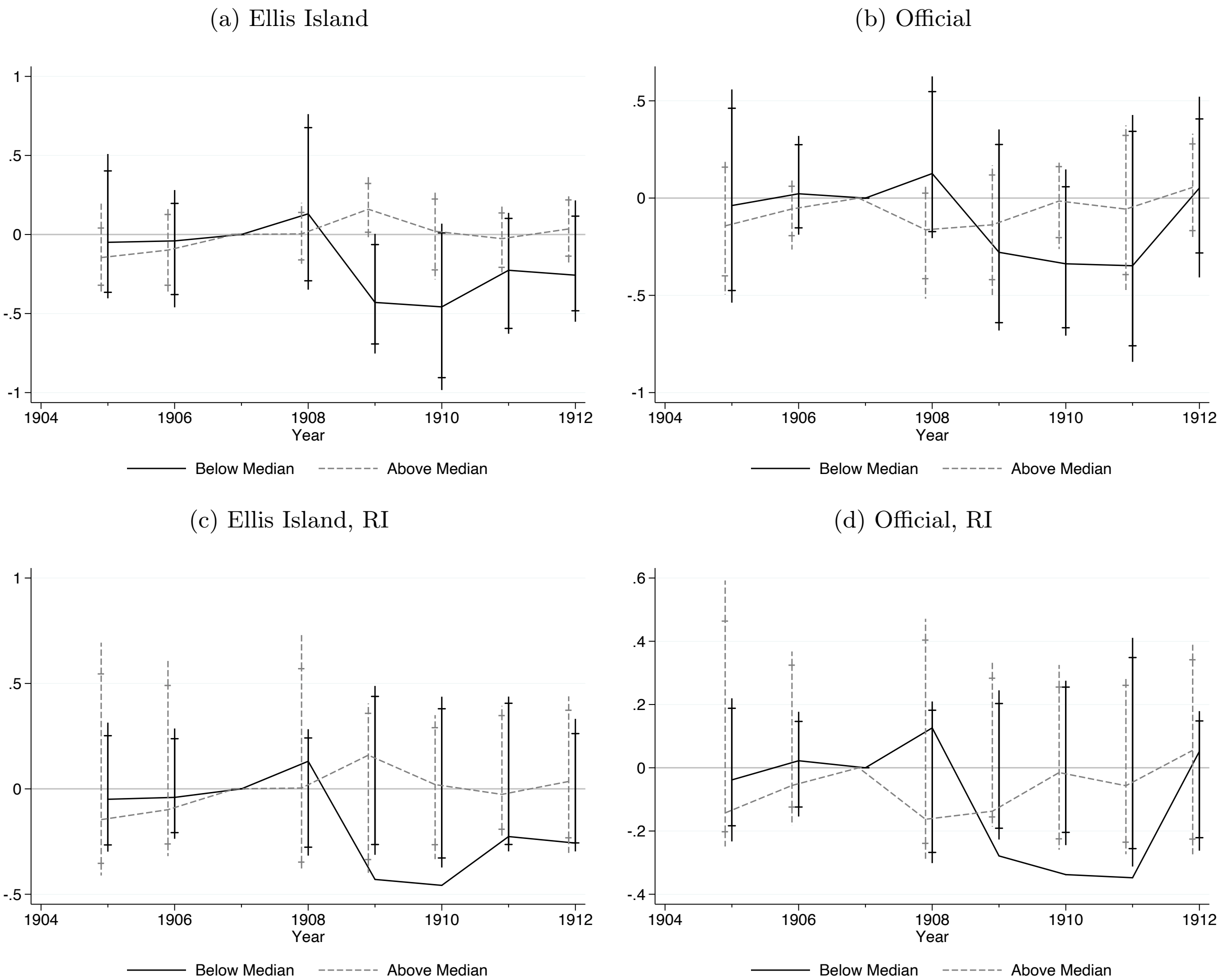



Figure G.3: Event studies divided by share of district employment in agricultural day labor with severity cutoff VII

*Note:* Sample includes all municipalities in the regions of Sicily and Calabria. All event studies control for a year-specific function of distance from the nearest emigration epicenter and have 1907 as the base year. In panels (a) and (b), bars indicate 90- and 95-percent confidence intervals clustered on the district level, computed by a wild bootstrap. In panels (c) and (d), bars indicate 90- and 95-percent randomization inference bands. The division into "below median" and "above median" is based on the distribution of the share of employment in agricultural day labor in the sample. The measure on the $y$-axis is the effect in logs.

# H Comparison to effects of other shocks

In this appendix, we expand on the comparison of the magnitude of our estimates to those of Mahajan and Yang (2020), and make some additional comparisons to other historical and modern shocks.

## H.1 Comparison to Effects of Hurricanes

The benchmark estimate that Mahajan and Yang (2020) provide (Table 3) is that a so-called "one-standard-deviation" hurricane increased migration by 11.8 percent of the mean. Mahajan and Yang (2020, Tables 3 and 5) also find that the effect of a hurricane on migration is increasing in the size of the existing migrant stock from the country of origin in the United States. We estimate that in 1908, the Italian migrant stock in the United States was about 3.8 percent of the Italian population. In 1900, the stock of Italian-born individuals in the United States was 484,027; by 1910, this figure was 1,343,125 (Haines 2006). To estimate the stock in 1908, we use information from Barde et al. (2006) on Italian arrivals in the United States. Based on the fact that 1,830,340 arrivals in 1901–1909 corresponded to an increase in population of 859,098 from 1900 to 1910, we assume that the 1,647,122 arrivals from 1901–1908 corresponded to a (proportional) population increase of 773,102, making our estimated stock of Italians in the United States after 1908 1,257,129, which is 3.8 percent of the 1901 Italian population of 32,965,504. This estimate provides a lower bound on the previous migrant stock in the United States of the earthquake-affected region, where emigration rates to the United States were considerably higher: in the period 1900–1914, the emigration rate to North America from Sicily and Calabria was about 2.4 times that of Italy as a whole, implying that a previous migrant stock of as much as 9 percent in the United States is not unreasonable. A lack of information on within-country place of origin in US censuses prevents us from performing a similar calculation regarding the stock of prior migrants for the earthquake-affected region.

According to the estimates of Mahajan and Yang (2020, Tables 3 and 5), a previous migrant stock of 3.8 percent corresponds to an estimated effect of a "one-standard-deviation" hurricane of about 10 percent of the mean, though if we accept an estimate of a prior migrant stock of 9 percent, the implied estimate is of an effect of 27.9 percent of the mean. Arguably, the most appropriate comparison given the open borders of the period we study, however, is Mahajan and Yang's (2020) estimate of the effect of a hurricane conditional on the stock of US citizens (rather than simply prior migrants) from the country of origin. This is because these individuals are able to support migration of individuals from their country of origin with the least interference of policy, as would have been the case for Italians in the United States in the open border period of the Age of Mass Migration. Applying the estimates of Mahajan and Yang (2020, Table 5) to a citizen stock of 3.8 percent of the origin population yields an estimated effect of a shock of 43.9 percent of the sample mean and a citizen stock of 9 percent of the population yields an estimated effect of 106 percent of the sample mean.

## H.2 Other Historical Shocks

The small impact of the earthquake relative to that of the business cycle is similar to Boustan's (2007) finding regarding the effect of pogroms on the Russian Jewish emigration to the United States. Similarly, Spitzer (2021) estimates the effect of a pogrom on the emigration of Russian Jews from another "migration-saturated" region—the Pale of Settlement in the Russian Empire around 1905. He finds that exposure to a pogrom led to an increase in emigration of about 20 percent, which is near the upper extreme of the earthquake's 90-percent interval. The Great Irish Famine of 1846–1850 is another useful comparison, in the sense that it was a disaster that had had an immense toll in lives and has been linked to subsequent mass migration (Ó Gráda 2019; Ó Gráda and O'Rourke 1997). The large impact of the shock is evident in the fact that the number of Irish immigrants to the United States more than doubled between 1846 and 1847 (Barde et al. 2006) when the famine intensified, even as the base population declined dramatically. We can easily rule out changes of this magnitude in our context.

## H.3 Comparison to Estimated Migration Declines in Response to Modern Earthquakes

Our estimated effect of the Messina-Reggio Calabria earthquake on migration is also precise enough to conclude that any negative aggregate impact of the earthquake was small relative to estimated migration declines in response to modern earthquakes.

The first such estimate is provided by Halliday (2006), who studies the effects of earthquakes in El Salvador in 2001. Halliday's (2006) main estimate is that a one-standard deviation increase in earthquake damage was associated with a decline in emigration probability of about 37 percent. The most comparable estimate in our case is of the effect of a one-standard deviation increase in our measure of damage—the Mercalli score. The estimates of Online Appendix Table B.5 imply that an increase in the Mercalli score of 1.43 (one standard deviation) was associated with a decline in emigration of at most 17.6 log points (according to the 95-percent confidence interval), meaning that we can rule our responses as large as those documented by Halliday (2006).

Another point of comparison is given by Shakya et al. (2022), who estimate that the 2015 Nepal earthquake caused a decline in emigration of about 38 percent in severely affected areas. While this figure is within the 95-percent confidence interval of our main estimates, it is outside the 90-percent confidence interval, implying that our estimates are likely smaller in magnitude. Moreover, Shakya et al. (2022) estimate that severely affected areas experienced a 49-percent decline in male emigration. This estimate is most comparable to our estimates for prime-aged males in column (3) of Table 4, where the lower end of our 95-percent confidence interval for the effect of the earthquake is a decline of 46.3 log points.

On the basis of both of these comparison, we can conclude that our estimates are precise enough both to rule out any large positive impact of the earthquake on emigration, as well as to show that any negative effect of the earthquake was also relatively small, thus supporting a conclusion that any aggregate effect of the earthquake was small.

# I Results with Controls

This section repeats the main results, but includes differential trends on the basis of a variety of controls in order to address potential differential pre-trends due to imbalance in the treatment. For the difference-in-differences specifications, each regression includes each of the controls interacted with the "post" indicator. For the event study specifications each regression includes each of the controls interacted with time. The controls included are all of those in Tables 3 and B.2 except that the event studies for heterogeneity with regards to the agricultural labor share do not include agricultural labor as a control.[84]

Table I.1: Difference-in-differences results with controls

| | Ellis Island | | | | Official Data | |
|---|---|---|---|---|---|---|
| | | | Prime-Age Only | | | |
| | (1) All | (2) Deaths | (3) All | (4) Deaths | (5) All | (6) Deaths |
| *Panel A: Sicily and Calabria* | 0.045 | 0.054 | 0.029 | 0.038 | −0.070 | −0.056 |
| District-Clustered b.s. p | (0.842) | (0.816) | (0.890) | (0.846) | (0.384) | (0.471) |
| Randomization Inference tail mass | {0.334} | {0.318} | {0.384} | {0.362} | {0.182} | {0.254} |
| District-Clustered b.s. 95% CI | [-0.196 0.411] | [-0.169 0.407] | [-0.227 0.399] | [-0.230 0.425] | [-0.258 0.110] | [-0.234 0.107] |
| Observations | 5,597 | 5,597 | 5,536 | 5,536 | 6,101 | 6,101 |
| R-squared | 0.754 | 0.754 | 0.745 | 0.745 | 0.631 | 0.630 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 |
| *Panel B: Sicily* | 0.194 | 0.210 | 0.151 | 0.166 | −0.053 | −0.027 |
| District-Clustered b.s. p | (0.672) | (0.607) | (0.710) | (0.705) | (0.768) | (0.911) |
| Randomization Inference tail mass | {0.070} | {0.052} | {0.112} | {0.098} | {0.212} | {0.318} |
| District-Clustered b.s. 95% CI | [-3.118 0.654] | [-2.488 0.651] | [-0.381 0.594] | [-0.493 0.585] | [-0.841 2.143] | [-0.544 0.205] |
| Observations | 2,708 | 2,708 | 2,678 | 2,678 | 2,815 | 2,815 |
| R-squared | 0.776 | 0.776 | 0.761 | 0.761 | 0.688 | 0.688 |
| Districts | 24 | 24 | 24 | 24 | 24 | 24 |
| *Panel C: Calabria* | −0.018 | −0.015 | −0.019 | −0.015 | −0.005 | 0.000 |
| District-Clustered b.s. p | (0.789) | (0.821) | (0.734) | (0.760) | (0.921) | (0.992) |
| Randomization Inference tail mass | {0.392} | {0.406} | {0.398} | {0.426} | {0.458} | {0.492} |
| District-Clustered b.s. 95% CI | [-0.328 0.337] | [-0.324 0.335] | [-0.166 0.179] | [-0.164 0.195] | [-0.148 0.200] | [-0.142 0.273] |
| Observations | 2,889 | 2,889 | 2,858 | 2,858 | 3,286 | 3,286 |
| R-squared | 0.736 | 0.736 | 0.732 | 0.731 | 0.541 | 0.537 |
| Districts | 11 | 11 | 11 | 11 | 11 | 11 |

*Significance levels, district-clustered b.s.*: $^{a}$ p<0.01, $^{b}$ p<0.05, $^{c}$ p<0.1
*Notes*: This table presents difference-in-differences coefficients for severe-times-post. P-values from a wild bootstrap clustering at the district level are in parentheses. Tail mass values from a randomization inference test are in curly braces. Ninety-five percent confidence intervals from the wild bootstrap clustered at the district level are in square braces. Columns (1), (3), and (5) use unadjusted base populations to compute migration rates; columns (2), (4), and (6) use populations adjusted for earthquake deaths. All regressions include municipality fixed effects, province-year indicators, distance from the nearest emigration epicenter-times-post, and various controls-times-post. Observations limited to 1905–1912.

[84] Recall that these regressions include an interaction of severe with an indicator for being above or below median in terms of the agricultural labor share.

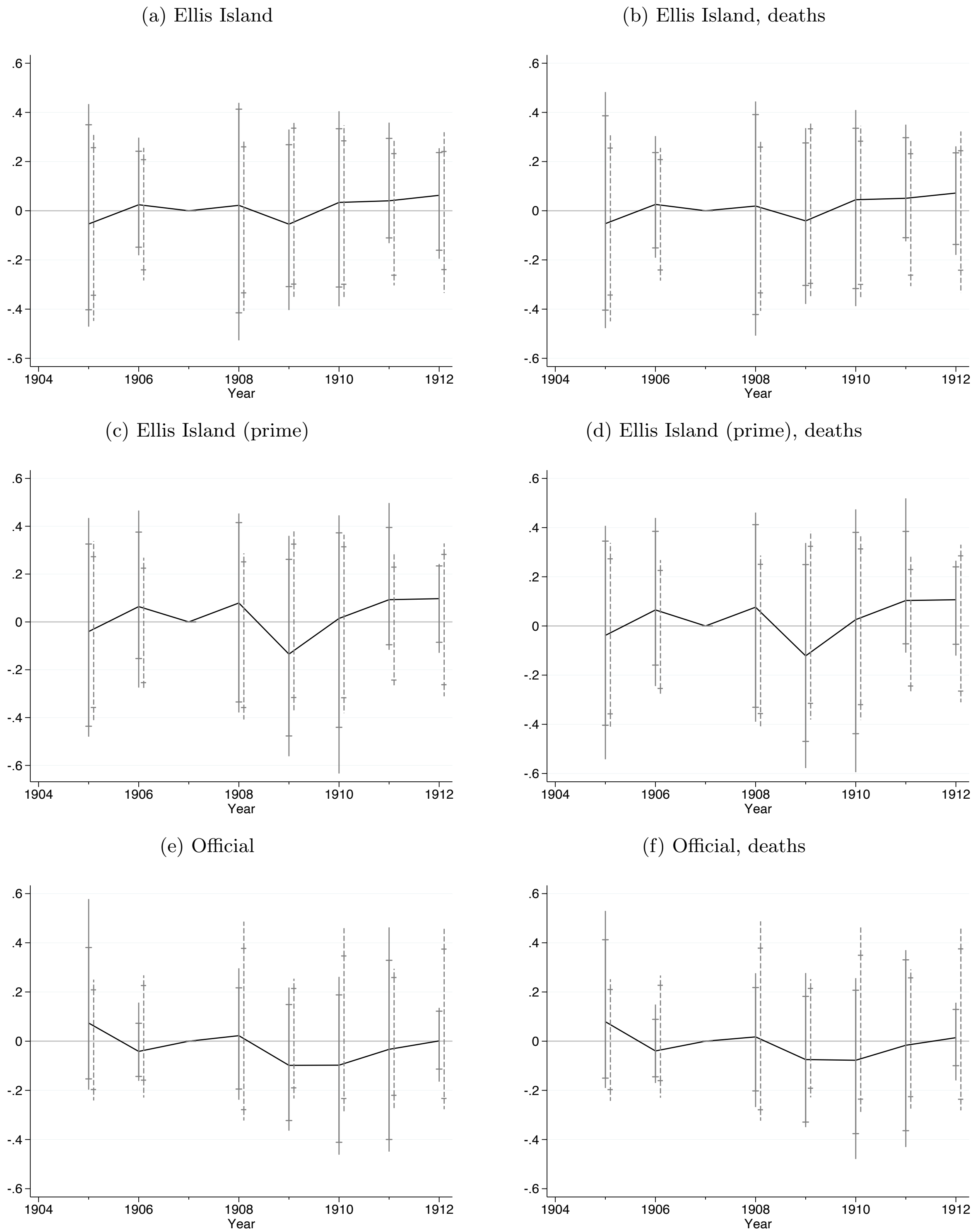



Figure I.1: Event studies for the effect of the earthquake on migration with controls

*Note:* Sample includes all municipalities in the regions of Sicily and Calabria. All event studies control for a year-specific function of distance from the nearest emigration epicenter and have 1907 as the base year. Solid bars indicate 95 percent confidence intervals from a wild bootstrap clustered on the district level. dashed bars are the middle 90 and 95 percent of results from the randomization inference exercise. The measure on the $y$-axis is the effect in logs.

Table I.2: Heterogeneous responses to the earthquake with controls

| | Ellis Island | | | | Official | | | |
|---|---|---|---|---|---|---|---|---|
| *Variables* | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Severe x Post | 0.045 | 0.056 | 0.044 | 0.057 | −0.070 | −0.054 | −0.072 | −0.057 |
| District-Clustered b.s. p | (0.847) | (0.739) | (0.817) | (0.754) | (0.384) | (0.285) | (0.332) | (0.333) |
| Randomization Inference tail mass | {0.334} | {0.304} | {0.368} | {0.318} | {0.182} | {0.210} | {0.270} | {0.332} |
| District-Clustered b.s. 95% CI | [-0.213 0.405] | [-0.160 0.344] | [-0.148 0.387] | [-0.160 0.383] | [-0.258 0.110] | [-0.214 0.072] | [-0.218 0.094] | [-0.203 0.082] |
| Severe x Post x Ag Lab Share | | $0.227^{b}$ | | $0.256^{c}$ | | $0.214^{a}$ | | 0.175 |
| District-Clustered b.s. p | | (0.044) | | (0.056) | | (0.007) | | (0.315) |
| Randomization Inference tail mass | | {0.024} | | {0.062} | | {0.070} | | {0.142} |
| District-Clustered b.s. 95% CI | | [0.003 0.495] | | [-0.008 1.146] | | [0.065 0.511] | | [-0.110 0.737] |
| Severe x Post x Ag Share | | | 0.125 | −0.031 | | | $0.147^{c}$ | 0.040 |
| District-Clustered b.s. p | | | (0.259) | (0.728) | | | (0.097) | (0.650) |
| Randomization Inference tail mass | | | {0.140} | {0.440} | | | {0.102} | {0.158} |
| District-Clustered b.s. 95% CI | | | [-0.087 0.681] | [-0.395 0.341] | | | [-0.335 0.779] | [-0.308 0.406] |
| Observations | 5,597 | 5,597 | 5,597 | 5,597 | 6,101 | 6,101 | 6,101 | 6,101 |
| R-squared | 0.754 | 0.754 | 0.754 | 0.754 | 0.631 | 0.632 | 0.632 | 0.632 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 | 35 | 35 |

*Significance levels, district-clustered b.s.*: $^{a}$ p<0.01, $^{b}$ p<0.05, $^{c}$ p<0.1

*Notes*: Sample includes all municipalities in the regions of Sicily and Calabria. Dependent variable is unadjusted emigration rate relative to 1901 population from the source listed in the column header. P-values from a wild bootstrap clustering at the district level are in parentheses. Tail mass values from a randomization inference test are in curly braces. Ninety-five percent confidence intervals from the wild bootstrap clustered at the district level are in square braces. Ag Share is defined as the fraction of the 1901 male labor force in agriculture at the district level, and is standardized to have mean zero and standard deviation one. Ag Lab Share is the share of the 1901 male labor force in agricultural day labor (*giornalieri di campagna*) at the district level and is standardized to have mean zero and standard deviation one in the sample. All regressions include municipality fixed effects, province-year indicators, distance from the nearest emigration epicenter-times-post, and various controls-times-post. Observations limited to 1905–1912.

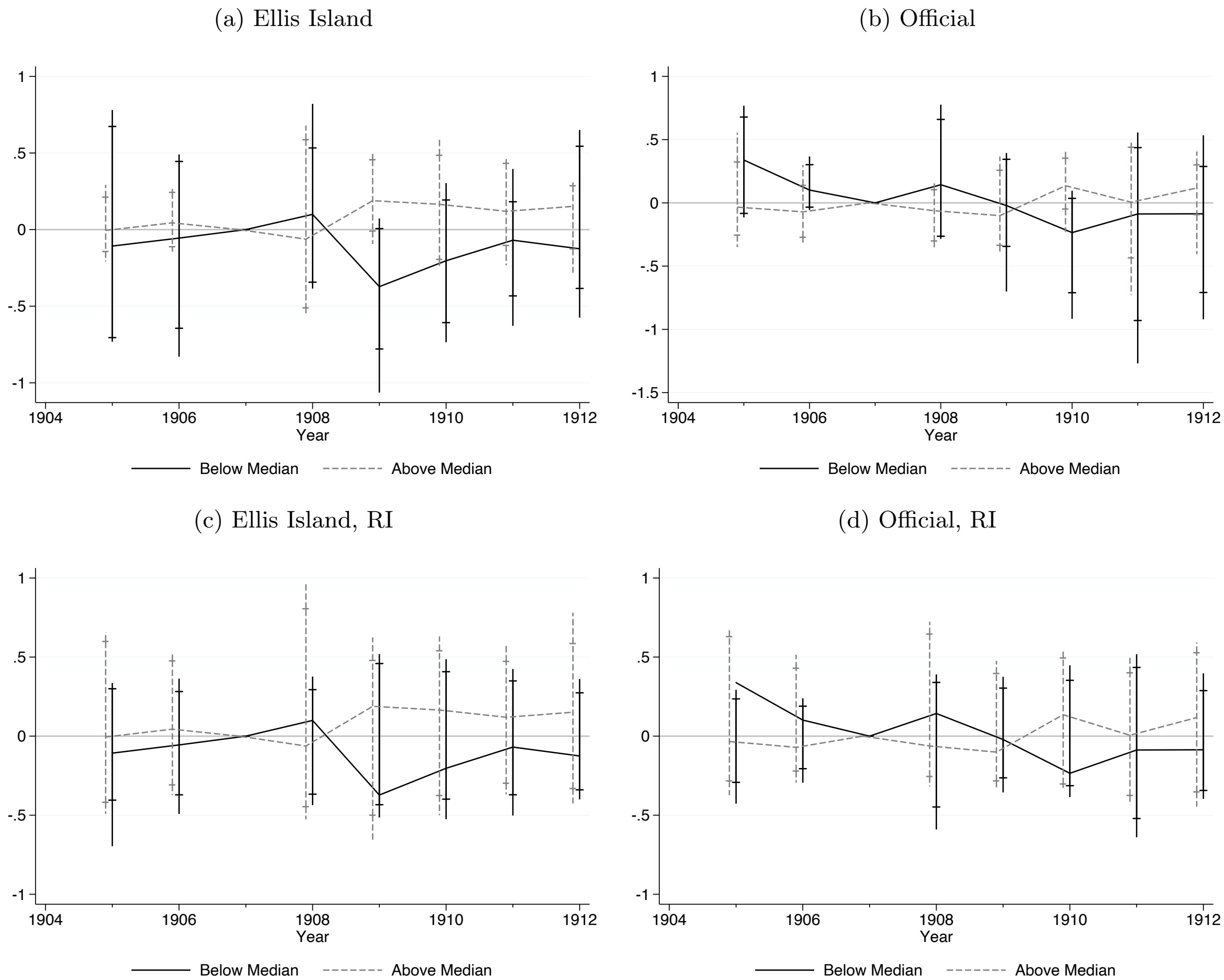



Figure I.2: Event studies divided by share of district employment in agricultural day labor with controls

*Note:* Sample includes all municipalities in the regions of Sicily and Calabria. All event studies control for a year-specific function of distance from the nearest emigration epicenter and have 1907 as the base year. In panels (a) and (b), bars indicate 90- and 95-percent confidence intervals clustered on the district level, computed by a wild bootstrap. In panels (c) and (d), bars indicate 90- and 95-percent randomization inference bands. The division into "below median" and "above median" is based on the distribution of the share of employment in agricultural day labor in the sample. The measure on the $y$-axis is the effect in logs.

## J Spillovers

The main analysis focused on the municipality as the unit of analysis. The disadvantage of focusing on such a high resolution is that the no-spillovers assumption inherent in the difference-in-differences and event study analyses is more likely to be violated. Severe damage in one municipality could have changed the migration incentives for individuals in adjacent municipalities that were not severely damaged. If these spillovers operated in the same direction, this would bias our estimates toward zero, limiting our ability to rule out a large effect. We take three approaches to addressing this concern. First, we point out that the analysis of Online Appendix Table B.6, using distance from the earthquake epicenter as the treatment, treats severely damaged and non-severely damaged municipalities equidistant from the earthquake epicenter equally, obviating concerns about spillovers in this regard. Second, Table J.1 repeats the results of Table 4 but omits any municipality that was not severely damaged but was contiguous to a severely damaged municipality.[85] There were 105 such municipalities, and omitting them leads to a more negative (albeit less precise) estimate, inconsistent with a spillover mechanism of this type being responsible for our findings.

Finally, we expand the unit of analysis to the district. Specifically, we estimate versions of equations (1) and (2) in which the unit of observation is the district-year and the province-year fixed effects are replaced with region-year fixed effects (there were 16 regions, or *compartimenti*). An important caveat in this analysis is that focusing on a larger administrative unit considerably reduces the number of observations. For this reason we focus on a specification that includes all Italian districts in the sample. We classify a district as treated if at least 40 percent of its population lived in a municipality that was severely damaged by the earthquake.[86] In practice, this results in the five closest districts to the epicenter (Castroreale and Messina in Sicily and Gerace Marina, Palmi, and Reggio Calabria in Calabria) being treated. The small number of treated units in this analysis raises issues of inference (MacKinnon and Webb 2020). To address this concern, we adjust our main randomization inference approach to the new unit of observation, randomly choosing 500 points in Italy, assigning the five closest districts as treated, and estimating the adapted version of equation (2).

We focus these results, which are presented in Figure J.2, on the unadjusted official statistics because they

[85] Figure J.1 presents analogous event study results.

[86] We chose this cutoff because it is a natural break in the distribution. The marginal district classified as severely damaged is Gerace Marina in the province of Reggio Calabria, where, 43.5 percent of the population lived in severely damaged municipalities. The next district with greater damage is Castroreale in the province of Messina, where 47.1 percent of the population lived in severely damaged municipalities. The marginal district classified as non-severely damaged is Monteleone di Calabria in the province of Catanzaro, where 23.8 percent of the population lived in a severely damaged municipality. Alternative results reclassifying the marginal district, Monteleone di Calabria, as severely damaged are presented in Online Appendix Figure J.4 with similar results.

provide more complete coverage of areas outside of southern Italy, which are included in these specifications. In these areas, migration to destinations other than the United States was considerably more important,[87] and therefore the loss of information arising from focusing only on migration to the United States—as the Ellis Island data entail—may be more severe.[88] The point estimate of -0.086 is remarkably similar to our estimates from Table 4, and like those estimates this one does not stand out from the placebo replications. In the event study, unlike the municipality-level results, there is some evidence of differential pre-trends between affected and unaffected districts, which complicates interpretation of the difference-in-differences coefficient, but at least from 1906 these are moderated. The point estimate for 1909 is consistent with a transient effect, if any, of the earthquake on emigration, this time a positive one. Any difference between affected and unaffected districts vanishes by 1910. None of these coefficients are statistically significant according to the randomization inference. In sum, we do not find evidence that the municipality-level analysis obscures aggregate district-level effects by neglecting to account for spillovers.

[87] In the northern regions of Piedmont, Lombardy, Veneto, and Emilia Romagna, over 77 percent of migrants in the period 1900–1914 traveled to other European countries.

[88] Results for alternative dependent variables are presented in Figure J.3.

Table J.1: Difference-in-differences results, spillovers

| | Ellis Island | | | | Official Data | |
|---|---|---|---|---|---|---|
| | | | Prime-Age Only | | | |
| | (1) All | (2) Deaths | (3) All | (4) Deaths | (5) All | (6) Deaths |
| *Panel A: Sicily and Calabria* | −0.093 | −0.080 | −0.103 | −0.090 | −0.100 | −0.074 |
| District-Clustered b.s. p | (0.613) | (0.606) | (0.599) | (0.667) | (0.652) | (0.747) |
| Randomization Inference tail mass | {0.220} | {0.238} | {0.196} | {0.234} | {0.190} | {0.260} |
| District-Clustered b.s. 95% CI | [-0.566 0.556] | [-0.513 0.543] | [-0.447 0.514] | [-0.440 0.541] | [-0.389 0.220] | [-0.379 0.232] |
| Observations | 4,844 | 4,844 | 4,788 | 4,788 | 5,281 | 5,281 |
| R-squared | 0.750 | 0.750 | 0.741 | 0.741 | 0.625 | 0.624 |
| Districts | 35 | 35 | 35 | 35 | 35 | 35 |
| *Panel B: Italy* | −0.061 | −0.047 | −0.079 | −0.065 | −0.061 | −0.037 |
| District-Clustered p | (0.716) | (0.773) | (0.634) | (0.687) | (0.624) | (0.775) |
| Randomization Inference tail mass | {0.198} | {0.248} | {0.114} | {0.152} | {0.226} | {0.318} |
| District-Clustered 95% CI | [-0.391 0.269] | [-0.368 0.274] | [-0.406 0.248] | [-0.384 0.253] | [-0.307 0.185] | [-0.290 0.217] |
| Observations | 38,765 | 38,765 | 36,578 | 36,578 | 60,332 | 60,332 |
| R-squared | 0.799 | 0.799 | 0.794 | 0.794 | 0.789 | 0.789 |
| Districts | 284 | 284 | 284 | 284 | 284 | 284 |
| *Panel C: Sicily* | −0.093 | −0.076 | −0.058 | −0.041 | −0.116 | −0.087 |
| District-Clustered b.s. p | (0.771) | (0.794) | (0.861) | (0.902) | (0.548) | (0.630) |
| Randomization Inference tail mass | {0.218} | {0.266} | {0.320} | {0.348} | {0.190} | {0.260} |
| District-Clustered b.s. 95% CI | [-0.846 0.758] | [-0.840 0.759] | [-0.781 0.776] | [-0.764 0.786] | [-0.527 0.260] | [-0.463 0.270] |
| Observations | 2,368 | 2,368 | 2,339 | 2,339 | 2,466 | 2,466 |
| R-squared | 0.775 | 0.775 | 0.758 | 0.759 | 0.681 | 0.681 |
| Districts | 24 | 24 | 24 | 24 | 24 | 24 |
| *Panel D: Calabria* | −0.084 | −0.075 | −0.138 | −0.130 | −0.164 | −0.147 |
| District-Clustered b.s. p | (0.637) | (0.786) | (0.295) | (0.334) | (0.683) | (0.669) |
| Randomization Inference tail mass | {0.138} | {0.150} | {0.068} | {0.078} | {0.050} | {0.058} |
| District-Clustered b.s. 95% CI | [-0.301 0.568] | [-1.624 0.781] | [-0.314 0.364] | [-1.270 0.803] | [-0.706 0.285] | [-7.202 0.558] |
| Observations | 2,476 | 2,476 | 2,449 | 2,449 | 2,815 | 2,815 |
| R-squared | 0.720 | 0.720 | 0.719 | 0.719 | 0.514 | 0.509 |
| Districts | 11 | 11 | 11 | 11 | 11 | 11 |

*Significance levels, district-clustered b.s.*: [a] p<0.01, [b] p<0.05, [c] p<0.1

*Notes*: This table presents difference-in-differences coefficients for severe-times-post, where the sample omits municipalities that are not severely damaged by the earthquake, but which are spatially contiguous to a severely damaged municipality. Tail mass-values from a wild bootstrap clustering at the district level are in parentheses. P-values from a randomization inference test are in curly braces. Ninety-five percent confidence intervals from the wild bootstrap clustered at the district level are in square braces. In panel B, the district-clustered p-values and confidence intervals are based on large-sample approximations, not the bootstrap. Columns (1), (3), and (5) use unadjusted base populations to compute migration rates; columns (2), (4), and (6) use populations adjusted for earthquake deaths. All regressions include commune fixed effects, province-year indicators, and distance from the nearest emigration epicenter-times-post. Observations limited to 1905–1912.

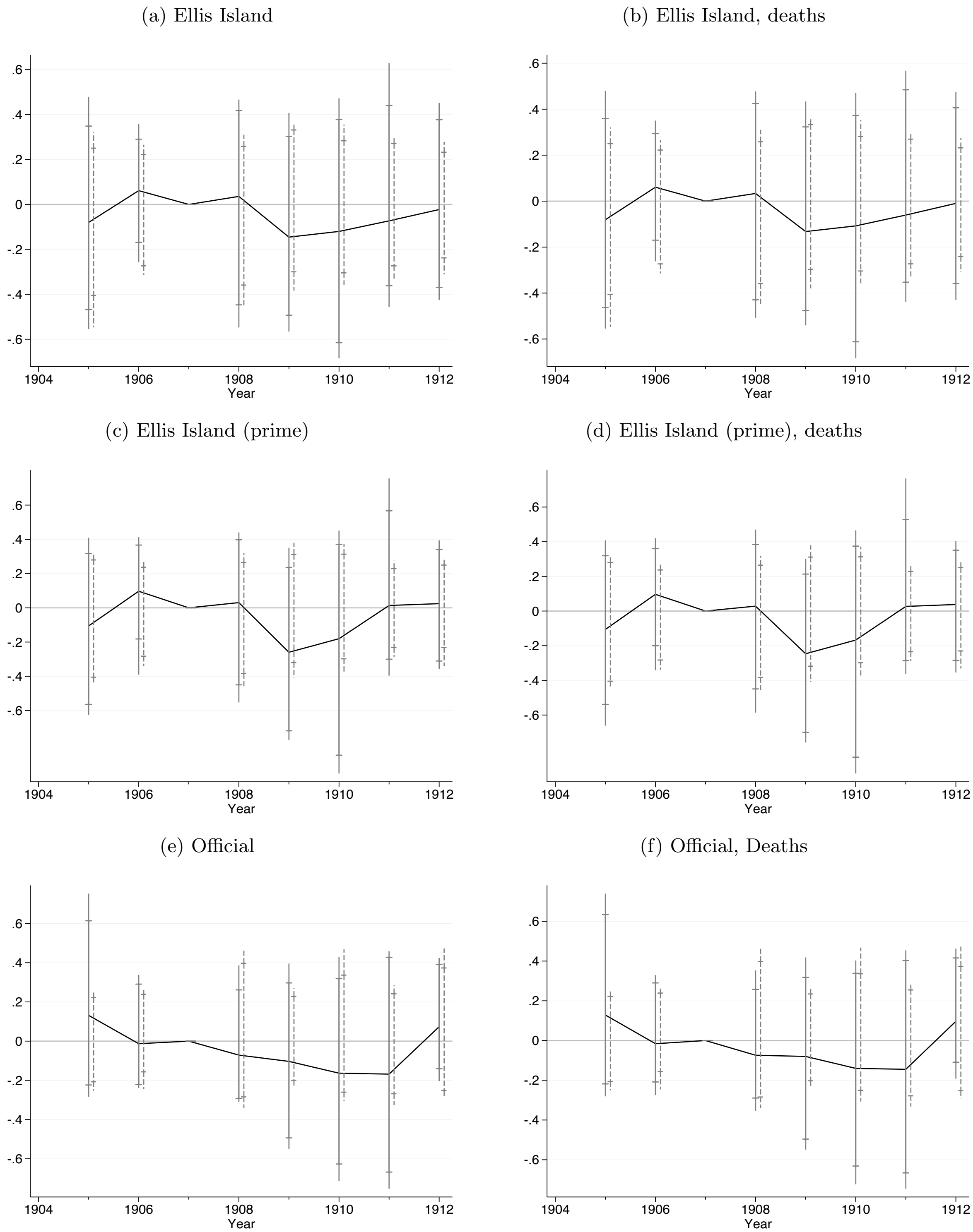


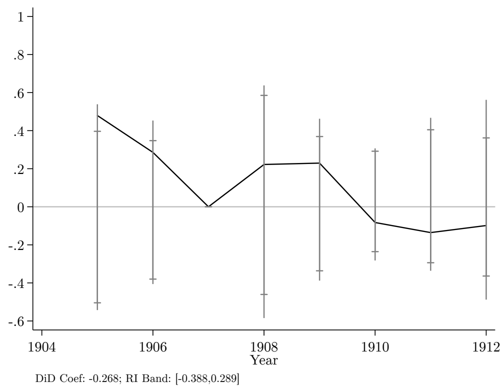

Figure J.1: Event studies for the effect of the earthquake on migration, spillovers

*Note:* Sample includes all municipalities in the regions of Sicily and Calabria, but excludes non-severely damaged municipalities contiguous to severely damaged municipalities. All event studies control for a year-specific function of distance from the nearest emigration epicenter and have 1907 as the base year. Solid bars are 90- and 95-percent confidence intervals from a wild bootstrap clustered on the district level. dashed bars are the middle 90 and 95 percent of randomization inference replications. The measure on the $y$-axis is the effect in logs.

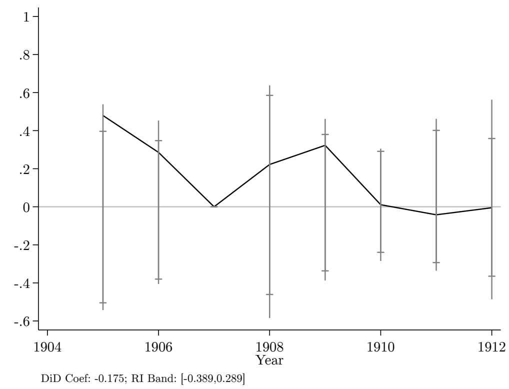

Figure J.2: Event studies at the district level

*Note:* Sample includes all districts in Italy. Bars represent 90- and 95-percent ranges of estimates from the randomization inference exercise. The measure on the $y$-axis is the year-specific effect in terms of log migration rates.

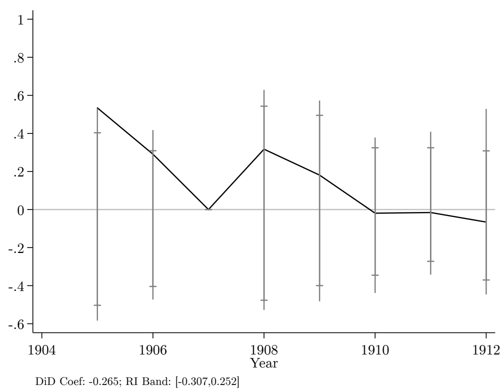
(a) Ellis Island



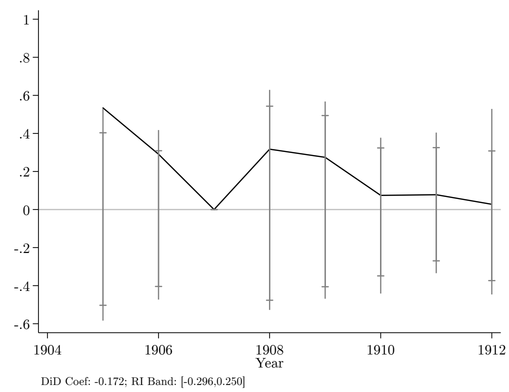

(b) Ellis Island, deaths




(c) Ellis Island (prime)




(d) Ellis Island (prime), deaths




(e) Official, deaths


Figure J.3: Event studies at the district level

*Note:* Sample includes all districts in Italy. Bars represent the middle 90 and 95 percent of estimates from the randomization inference exercise. The measure on the $y$-axis is the effect in logs.

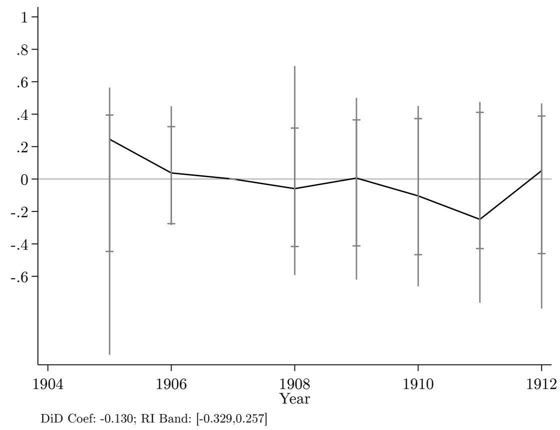



Figure J.4: Event studies at the district level, alternative severity definition

*Note:* These results incorporate the marginal non-severe district from Figure J.2 into the severe category. Sample includes all districts in Italy. Bars represent the middle 90 and 95 percent ranges of estimates from the randomization inference exercise. The measure on the $y$-axis is the effect in logs.

# References

Barde, Robert, Susan B. Carter, and Richard Sutch (2006). "Table Ad106–120: Immigrants, by country of last residence—Europe, 1820–1997." In *Historical Statistics of the United States*. Susan B. Carter, Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright (ed.). Cambridge: Cambridge University Press, pp. 1.560–1.563.

Boustan, Leah Platt (2007). "Were Jews Political Refugees or Economic Migrants? Assessing the Persecution Theory of Jewish Emigration, 1881–1914." In *The New Comparative Economic History: Essays in Honor of Jeffrey G. Williamson*. Kevin H. O'Rourke, Alan M. Taylor, Jeffrey G. Williamson, and Timothy J. Hatton (ed.). Cambridge: MIT Press, pp. 267–290.

Haines, Michael R. (2006). "Table Ad354–443: Foreign-born population, by country of birth: 1850–1990." In *Historical Statistics of the United States*. Susan B. Carter, Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright (ed.). Cambridge: Cambridge University Press, pp. 1.601–1.606.

Halliday, Timothy (2006). "Migration, Risk, and Liquidity Constraints in El Salvador." *Economic Development and Cultural Change* 54:4, pp. 893–925.

ISTAT (2018). "Confini delle unità amministrative a fini statistici al 1 gennaio 2018." [machine-readable database].

MacKinnon, James G. and Matthew D. Webb (2020). "Randomization Inference for Difference-in-Differences with Few Treated Clusters." *Journal of Econometrics* 218:2, pp. 435–450.

Mahajan, Parag and Dean Yang (2020). "Taken By Storm: Hurricanes, Migrant Networks, and US Immigration." *American Economic Journal: Applied Economics* 12:2, pp. 250–277.

Ó Gráda, Cormac (2019). "The Next World and the New World: Relief, Migration, and the Great Irish Famine." *Journal of Economic History* 79:2, pp. 319–355.

Ó Gráda, Cormac and Kevin H. O'Rourke (1997). "Migration as Disaster Relief: Lessons from the Great Irish Famine." *European Review of Economic History* 1:3–25.

Roodman, David, Morten Ørregaard Nielsen, James G. MacKinnon, and Matthew D. Webb (2019). "Fast and Wild: Bootstrap Inference in Stata using boottest." *Stata Journal* 19:1, pp. 4–60.

Shakya, Shishir, Subuna Basnet, and Jayash Paudel (2022). "Natural Disasters and Labor Migration: Evidence from Nepal's Earthquake." *World Development* 151, 105748.

Spitzer, Yannay (2021). "Pogroms, Networks, and Migration: The Jewish Migration from the Russian Empire to the United States 1881–1914." Mimeo., Hebrew University of Jerusalem.

US Geological Survey (2019). "Magnitude/Intensity Comparison." URL: https://earthquake.usgs.gov/learn/topics/mag_vs_int.php.