***Appendix***

*Table 1.1:* Political Instability Task Force onset cases used in this analysis

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Indonesia | 1965 | Sudan | 1983 | Guinea-Bissau | 1998 |
| Thailand | 1965 | South Africa | 1984 | Yugoslavia | 1998 |
| Guatemala | 1966 | Turkey | 1984 | Ethiopia | 1999 |
| Uganda | 1966 | Liberia | 1985 | Russia | 1999 |
| India | 1967 | Israel | 1987 | Guinea | 2000 |
| Jordan | 1970 | Burundi | 1988 | Liberia | 2000 |
| Oman | 1970 | China | 1988 | Rwanda | 2001 |
| Pakistan | 1971 | Somalia | 1988 | Ivory Coast | 2002 |
| United Kingdom | 1971 | Liberia | 1989 | Iraq | 2003 |
| Zimbabwe | 1972 | Papua New Guinea | 1989 | Pakistan | 2004 |
| Chile | 1973 | Romania | 1989 | Thailand | 2004 |
| Colombia | 1975 | Mali | 1990 | Turkey | 2004 |
| Indonesia | 1975 | Rwanda | 1990 | Yemen | 2004 |
| Lebanon | 1975 | Algeria | 1991 | Chad | 2005 |
| Morocco | 1975 | Djibouti | 1991 | Nigeria | 2006 |
| Argentina | 1976 | Kenya | 1991 | Ethiopia | 2007 |
| Congo Kinshasa | 1977 | Sierra Leone | 1991 | Mexico | 2007 |
| Iran | 1977 | Yugoslavia | 1991 | China | 2009 |
| Afghanistan | 1978 | Congo Kinshasa | 1992 | Egypt | 2011 |
| Nicaragua | 1978 | Egypt | 1992 | Libya | 2011 |
| Syria | 1979 | Senegal | 1992 | Sudan | 2011 |
| Iraq | 1980 | Russia | 1994 | Syria | 2011 |
| Nigeria | 1980 | Yemen | 1994 | Mali | 2012 |
| Peru | 1982 | Nepal | 1996 | Ukraine | 2014 |
| India | 1983 | Congo Brazzaville | 1997 |  |  |
| Pakistan | 1983 | Indonesia | 1997 |  |  |

*Table 1.2:* Peace Research Institute Oslo onset cases used in this analysis

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sudan | 1963 | Peru | 1982 | Israel | 2000 |
| Colombia | 1964 | India | 1983 | Liberia | 2000 |
| Guatemala | 1965 | Sudan | 1983 | United States | 2001 |
| Bolivia | 1967 | Cameroon | 1984 | Congo Brazzaville | 2002 |
| Nigeria | 1967 | Turkey | 1984 | Thailand | 2003 |
| Pakistan | 1971 | Somalia | 1986 | Iraq | 2004 |
| United Kingdom | 1971 | Laos | 1989 | Pakistan | 2004 |
| Iraq | 1973 | Paraguay | 1989 | Azerbaijan | 2005 |
| Zimbabwe | 1973 | Rwanda | 1990 | Congo Kinshasa | 2005 |
| Argentina | 1974 | Algeria | 1991 | Iran | 2005 |
| Pakistan | 1974 | Sierra Leone | 1991 | Peru | 2007 |
| Thailand | 1974 | Yugoslavia | 1991 | Azerbaijan | 2008 |
| Bangladesh | 1975 | Indonesia | 1992 | Rwanda | 2009 |
| Indonesia | 1975 | Burundi | 1994 | Yemen | 2009 |
| Lebanon | 1975 | Russia | 1994 | Tajikistan | 2010 |
| Mauritania | 1975 | Yemen | 1994 | Congo Kinshasa | 2011 |
| Morocco | 1975 | Iran | 1996 | Libya | 2011 |
| Chad | 1976 | Nepal | 1996 | Nigeria | 2011 |
| Congo Kinshasa | 1977 | Azerbaijan | 1997 | Senegal | 2011 |
| Mozambique | 1977 | Chad | 1997 | Syria | 2011 |
| Nicaragua | 1977 | Congo Brazzaville | 1997 | Azerbaijan | 2012 |
| Afghanistan | 1978 | Indonesia | 1997 | Mali | 2012 |
| Cambodia | 1978 | Senegal | 1997 | Mozambique | 2013 |
| El Salvador | 1979 | Angola | 1998 | Burundi | 2014 |
| Iran | 1979 | United Kingdom | 1998 | Ukraine | 2014 |
| Syria | 1979 | Yugoslavia | 1998 |  |  |
| Uganda | 1979 | Russia | 1999 |  |  |

*Table 2*: Description of independent variables

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Unit of Measurement** | **Temporal Coverage** | **Definition** | **Data Source** |
| Discrimination (sum) | Index | 1950-2018 | Discriminatory index value per country year (sum of political and economic discrimination) | Minorities at Risk (Marshall and Asal 2007) |
| GDP/capita | 2011 PPP$ | 1960-2015 | Gross Domestic Product per Capita | IMF World Economic Outlook Data (2022) |
| Mean-centered, Annualized Infant Mortality | Number of Deaths/Thousand from Mean | 1960-2015 | Number of deaths per thousand live births from the mean value (0) by year | WDI, WHO, World Bank, UNICEF (2021) |
| Polity (squared) | Index (0-100) | 1956-2015 | Polity value of a country squared to measure level of anocracy | Polity IV- Center for Systemic Peace (Marshall, Gurr, and Jaggers 2016) |
| Liberal Democracy Index | Index (0-1) | 1956-2015 | Index measuring the level of rights of individuals and minorities protecting them from the state | Variety of Democracy (Coppedge et al. 2021) |
| Ratio of Total Borders at War | Ratio (0-1) | 1946-2012 | Percent of total of countries which share a border and are experiencing any type of conflict | Center for Systemic Peace (Marshall, Gurr, and Harff 2021) |
| Youth Bulge (15-29/15+) | Percentage | 1956-2015 | Percentage of the population ages 15-29 over population 15 years and older | UN Procurement Division (Department of Economic and Social Affairs 2021) |

***Table 3****:* Descriptive statistics of independent variables (Base model, PITF cases, unnormalized)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Predictor** | **Associated Theory of State Fragility** | **N** | **Mean** | **Std. Dev.** | **Min** | **Max** |
| Regime Type (squared) | Governance | 76 | 40.49 | 26.02 | 0 | 100 |
| Youth bulge, 15+ (%) | Demographics | 76 | 44.63 | 6.88 | 24.31 | 53.24 |
| GDP/capita (ln) | Economic Development | 76 | 8.10 | 1.20 | 2.37 | 10.25 |
| Borders in conflict (Ratio) | Neighborhood Effects | 76 | 0.37 | 0.32 | 0 | 1 |
| Infant mortality rate (%, annual, mean centered) | Human Development | 76 | 22.51 | 40.15 | -72.98 | 109.17 |
| State-led discrimination (sum political & economic) | State-led Discrimination | 76 | 12.43 | 10.74 | 0 | 40 |

***Table 4***: Summary statistics of clusters (Base model, PITF cases, unnormalized)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Cluster** | **Summary statistic** | **Polity IV (squared)** | **Youth bulge, 15+ (%)** | **GDP per capita (ln)** | **Borders in conflict (Ratio)** | **Infant mortality rate (%, annual, mean centered)** | **State-led discrimination (sum political & economic)** |
| Anocratic - Younger - Less Developed | Mean | 31.03 | 48.18 | 7.61 | 0.15 | 47.01 | 5.5 |
| Median | 25 | 48.16 | 7.70 | 0 | 51.78 | 6 |
| Older - Wealthier - Moderate Discrimination | Mean | 46.77 | 31.5 | 9.44 | 0.18 | -27.51 | 15.38 |
| Median | 49 | 29.7 | 9.46 | 0.20 | -25.22 | 16 |
| Higher Discrimination - Worse Neighborhood - Younger | Mean | 46.61 | 46.57 | 8.01 | 0.66 | 19.93 | 17.58 |
| Median | 49 | 46.30 | 7.78 | 0.6 | 21.23 | 14 |

***Table 5****:*Model and Data Specification for Sensitivity Analysis

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Clustering algorithm** | | **Distance metric** | | **Sample** | | **Variable substitution** | |
| **Agnes** | **PAM** | **Euclidean** | **Gower** | **PITF** | **PRIO** | **Polity** | **Vdem** |
| x |  | x |  | x |  | x |  |
|  | x | x |  | x |  | x |  |
| x |  |  | x | x |  | x |  |
|  | x |  | x | x |  | x |  |
| x |  | x |  |  | x | x |  |
|  | x | x |  |  | x | x |  |
| x |  |  | x |  | x | x |  |
|  | x |  | x |  | x | x |  |
| x |  | x |  | x |  |  | x |
|  | x | x |  | x |  |  | x |
| x |  |  | x | x |  |  | x |
|  | x |  | x | x |  |  | x |
| x |  | x |  |  | x |  | x |
|  | x | x |  |  | x |  | x |
| x |  |  | x |  | x |  | x |
|  | x |  | x |  | x |  | x |

*Table 6:* Mean values of cluster-by-cluster specifications (median values in parentheses).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Method** | **Cluster** | **N** | **Polity IV (squared) *OR* Vdem Index** | **Youth bulge, 15+ (%)** | **GDP per capita (ln)** | **Borders in conflict (Ratio)** | **Infant mortality rate (%, annual, mean centered)** | **State-led discrimination (sum political & economic)** |
| EUCL+PITF+AGNES (Base Model) | Anocratic - Younger - Less Developed | 30 | 31.03 (25.0) | 48.18 (48.16) | 7.61 (7.7) | 0.15 (0.0) | 47.01 (51.78) | 5.5 (6.0) |
| EUCL+PITF+PAM | Anocratic - Younger - Less Developed | 28 | 19.61 (16.0) | 47.85 (48.06) | 7.4 (7.4) | 0.34 (0.33) | 51.05 (55.31) | 6.29 (6.0) |
| EUCL+PITF+VDEM+AGNES | Anocratic - Younger - Less Developed | 30 | 0.14 (0.13) | 48.3 (48.16) | 7.49 (7.55) | 0.16 (0.0) | 51.34 (55.31) | 6.33 (6.0) |
| EUCL+PITF+VDEM+PAM | Anocratic - Younger - Less Developed | 22 | 0.14 (0.13) | 47.93 (47.77) | 7.56 (7.7) | 0.06 (0.0) | 51.47 (52.91) | 7.23 (6.5) |
| EUCL+PRIO+AGNES | Anocratic - Younger - Less Developed | 45 | 37.49 (49.0) | 47.42 (48.17) | 7.83 (8.05) | 0.21 (0.25) | 34.19 (27.13) | 6.4 (6.0) |
| EUCL+PRIO+PAM | Anocratic - Younger - Less Developed | 25 | 30.16 (25.0) | 46.66 (48.17) | 7.81 (7.92) | 0.1 (0.0) | 37.74 (35.86) | 8.36 (7.0) |
| EUCL+PRIO+VDEM+AGNES | Anocratic - Younger - Less Developed | 46 | 0.14 (0.09) | 47.56 (48.2) | 7.87 (8.06) | 0.22 (0.25) | 32.82 (26.62) | 6.96 (6.0) |
| EUCL+PRIO+VDEM+PAM | Anocratic - Younger - Less Developed | 24 | 0.14 (0.09) | 47.04 (47.19) | 7.91 (7.99) | 0.05 (0.0) | 37.18 (37.56) | 10.54 (8.0) |
| GOWER+PITF+AGNES | Anocratic - Younger - Less Developed | 32 | 32.16 (30.5) | 47.76 (48.08) | 7.76 (7.77) | 0.15 (0.0) | 43.32 (48.26) | 5.56 (6.0) |
| GOWER+PITF+PAM | Anocratic - Younger - Less Developed | 22 | 40.64 (49.0) | 48.02 (47.77) | 7.53 (7.6) | 0.07 (0.0) | 53.34 (58.05) | 7.0 (6.5) |
| GOWER+PITF+VDEM+AGNES | Anocratic - Younger - Less Developed | 30 | 0.15 (0.13) | 48.41 (48.24) | 7.46 (7.54) | 0.18 (0.0) | 51.6 (55.31) | 5.63 (6.0) |
| GOWER+PITF+VDEM+PAM | Anocratic - Younger - Less Developed | 20 | 0.15 (0.13) | 47.65 (47.48) | 7.94 (7.8) | 0.02 (0.0) | 46.85 (48.26) | 7.45 (7.0) |
| GOWER+PRIO+AGNES | Anocratic - Younger - Less Developed | 39 | 39.79 (49.0) | 47.39 (48.23) | 7.78 (7.92) | 0.22 (0.25) | 31.24 (25.7) | 5.13 (5.0) |
| GOWER+PRIO+PAM | Anocratic - Younger - Less Developed | 27 | 22.11 (16.0) | 48.46 (48.46) | 7.22 (7.3) | 0.31 (0.25) | 51.63 (53.62) | 8.26 (6.0) |
| GOWER+PRIO+VDEM+AGNES | Anocratic - Younger - Less Developed | 46 | 0.13 (0.09) | 47.28 (48.2) | 7.88 (8.06) | 0.21 (0.22) | 31.88 (26.62) | 6.87 (6.0) |
| GOWER+PRIO+VDEM+PAM | Anocratic - Younger - Less Developed | 49 | 0.14 (0.09) | 47.52 (48.17) | 7.87 (8.05) | 0.28 (0.25) | 31.31 (26.11) | 6.55 (6.0) |
| EUCL+PITF+AGNES (Base Model) | Older - Wealthier - Moderate Discrimination | 13 | 46.77 (49.0) | 31.5 (29.7) | 9.44 (9.46) | 0.18 (0.2) | -27.51 (-25.22) | 15.38 (16.0) |
| EUCL+PITF+PAM | Older - Wealthier - Moderate Discrimination | 12 | 46.58 (42.5) | 30.53 (29.14) | 9.37 (9.45) | 0.17 (0.17) | -28.67 (-25.65) | 16.67 (16.5) |
| EUCL+PITF+VDEM+AGNES | Older - Wealthier - Moderate Discrimination | 14 | 0.36 (0.32) | 32.12 (30.17) | 9.36 (9.45) | 0.18 (0.17) | -27.14 (-24.88) | 15.71 (16.5) |
| EUCL+PITF+VDEM+PAM | Older - Wealthier - Moderate Discrimination | 14 | 0.36 (0.32) | 32.12 (30.17) | 9.36 (9.45) | 0.18 (0.17) | -27.14 (-24.88) | 15.71 (16.5) |
| EUCL+PRIO+AGNES | Older - Wealthier - Moderate Discrimination | 12 | 60.25 (65.0) | 29.9 (27.87) | 9.57 (9.47) | 0.11 (0.14) | -30.77 (-28.33) | 15.67 (14.5) |
| EUCL+PRIO+PAM | Older - Wealthier - Moderate Discrimination | 14 | 67.14 (81.0) | 33.86 (30.87) | 9.37 (9.37) | 0.09 (0.0) | -23.61 (-26.7) | 13.0 (10.5) |
| EUCL+PRIO+VDEM+AGNES | Older - Wealthier - Moderate Discrimination | 12 | 0.45 (0.37) | 29.9 (27.87) | 9.57 (9.47) | 0.11 (0.14) | -30.77 (-28.33) | 15.67 (14.5) |
| EUCL+PRIO+VDEM+PAM | Older - Wealthier - Moderate Discrimination | 12 | 0.45 (0.37) | 29.9 (27.87) | 9.57 (9.47) | 0.11 (0.14) | -30.77 (-28.33) | 15.67 (14.5) |
| GOWER+PITF+AGNES | Older - Wealthier - Moderate Discrimination | 12 | 47.83 (42.5) | 30.54 (29.14) | 9.39 (9.45) | 0.2 (0.21) | -29.11 (-25.65) | 17.08 (17.5) |
| GOWER+PITF+PAM | Older - Wealthier - Moderate Discrimination | 17 | 44.06 (36.0) | 34.26 (32.5) | 9.39 (9.45) | 0.25 (0.21) | -25.47 (-24.54) | 16.41 (17.0) |
| GOWER+PITF+VDEM+AGNES | Older - Wealthier - Moderate Discrimination | 14 | 0.36 (0.32) | 32.12 (30.17) | 9.36 (9.45) | 0.18 (0.17) | -27.14 (-24.88) | 15.71 (16.5) |
| GOWER+PITF+VDEM+PAM | Older - Wealthier - Moderate Discrimination | 13 | 0.35 (0.31) | 31.25 (29.7) | 9.4 (9.46) | 0.19 (0.2) | -27.68 (-25.22) | 16.77 (17.0) |
| GOWER+PRIO+AGNES | Older - Wealthier - Moderate Discrimination | 12 | 60.25 (65.0) | 29.9 (27.87) | 9.57 (9.47) | 0.11 (0.14) | -30.77 (-28.33) | 15.67 (14.5) |
| GOWER+PRIO+PAM | Older - Wealthier - Moderate Discrimination | 16 | 64.38 (81.0) | 34.12 (30.87) | 9.23 (9.37) | 0.08 (0.0) | -19.25 (-25.65) | 12.62 (10.5) |
| GOWER+PRIO+VDEM+AGNES | Older - Wealthier - Moderate Discrimination | 11 | 0.48 (0.43) | 29.42 (27.37) | 9.58 (9.45) | 0.12 (0.14) | -31.08 (-29.35) | 16.18 (16.0) |
| GOWER+PRIO+VDEM+PAM | Older - Wealthier - Moderate Discrimination | 12 | 0.45 (0.37) | 29.9 (27.87) | 9.57 (9.47) | 0.11 (0.14) | -30.77 (-28.33) | 15.67 (14.5) |
| EUCL+PITF+AGNES (Base Model) | Higher Discrimination - Worse Neighborhood - Younger | 33 | 46.61 (49.0) | 46.57 (46.3) | 8.01 (7.78) | 0.66 (0.6) | 19.93 (21.23) | 17.58 (14.0) |
| EUCL+PITF+PAM | Higher Discrimination - Worse Neighborhood - Younger | 36 | 54.69 (49.0) | 46.83 (46.87) | 8.22 (8.13) | 0.47 (0.5) | 17.36 (19.77) | 15.81 (13.5) |
| EUCL+PITF+VDEM+AGNES | Higher Discrimination - Worse Neighborhood - Younger | 32 | 0.14 (0.09) | 46.66 (46.27) | 8.12 (7.98) | 0.66 (0.6) | 17.2 (17.08) | 16.72 (13.5) |
| EUCL+PITF+VDEM+PAM | Higher Discrimination - Worse Neighborhood - Younger | 40 | 0.15 (0.1) | 47.19 (47.37) | 7.95 (7.77) | 0.62 (0.56) | 23.95 (23.98) | 14.15 (9.5) |
| EUCL+PRIO+AGNES | Higher Discrimination - Worse Neighborhood - Younger | 22 | 42.36 (49.0) | 46.63 (46.59) | 8.02 (7.97) | 0.67 (0.6) | 26.28 (26.62) | 22.32 (23.0) |
| EUCL+PRIO+PAM | Higher Discrimination - Worse Neighborhood - Younger | 40 | 41.2 (49.0) | 46.95 (47.46) | 7.93 (8.1) | 0.55 (0.5) | 28.36 (25.84) | 14.4 (11.5) |
| EUCL+PRIO+VDEM+AGNES | Higher Discrimination - Worse Neighborhood - Younger | 24 | 0.13 (0.09) | 46.69 (46.73) | 8.1 (8.26) | 0.7 (0.61) | 22.21 (18.82) | 22.21 (20.0) |
| EUCL+PRIO+VDEM+PAM | Higher Discrimination - Worse Neighborhood - Younger | 46 | 0.13 (0.09) | 47.37 (48.01) | 7.97 (8.14) | 0.55 (0.5) | 25.01 (23.88) | 13.04 (8.5) |
| GOWER+PITF+AGNES | Higher Discrimination - Worse Neighborhood - Younger | 32 | 46.06 (49.0) | 46.78 (46.49) | 7.96 (7.77) | 0.66 (0.6) | 21.05 (22.32) | 17.56 (14.0) |
| GOWER+PITF+PAM | Higher Discrimination - Worse Neighborhood - Younger | 37 | 38.76 (49.0) | 47.38 (47.69) | 7.84 (7.75) | 0.61 (0.57) | 26.22 (24.54) | 13.84 (9.0) |
| GOWER+PITF+VDEM+AGNES | Higher Discrimination - Worse Neighborhood - Younger | 32 | 0.14 (0.08) | 46.56 (46.27) | 8.15 (7.98) | 0.64 (0.6) | 16.95 (17.08) | 17.38 (14.0) |
| GOWER+PITF+VDEM+PAM | Higher Discrimination - Worse Neighborhood - Younger | 43 | 0.15 (0.11) | 47.27 (47.63) | 7.78 (7.75) | 0.59 (0.5) | 26.36 (24.54) | 13.44 (8.0) |
| GOWER+PRIO+AGNES | Higher Discrimination - Worse Neighborhood - Younger | 28 | 38.11 (36.0) | 46.84 (47.07) | 8.04 (8.24) | 0.56 (0.54) | 32.08 (42.9) | 20.68 (19.0) |
| GOWER+PRIO+PAM | Higher Discrimination - Worse Neighborhood - Younger | 36 | 47.64 (49.0) | 46.23 (46.75) | 8.36 (8.36) | 0.44 (0.41) | 18.37 (15.83) | 15.06 (13.5) |
| GOWER+PRIO+VDEM+AGNES | Higher Discrimination - Worse Neighborhood - Younger | 25 | 0.13 (0.09) | 46.74 (46.9) | 8.13 (8.39) | 0.69 (0.6) | 21.95 (18.32) | 21.88 (20.0) |
| GOWER+PRIO+VDEM+PAM | Higher Discrimination - Worse Neighborhood - Younger | 21 | 0.12 (0.08) | 46.64 (46.68) | 8.14 (8.39) | 0.63 (0.6) | 24.21 (19.32) | 25.33 (26.0) |

*Table 7:* Summary results of cluster analysis sensitivity results, normalized (standard deviation measures the deviation around the mean across alternate clustering techniques, with higher values indicating greater volatility).

|  |  |  |
| --- | --- | --- |
| **Cluster - Variable** | **Standard Deviation** | **Mean value per cluster** |
| **Anocratic - Younger - Less Developed** | | |
| Polity squared ***OR*** Vdem Index | 0.096 | 0.239 |
| Youth bulge (%) | 0.040 | 0.783 |
| GDP per capita (ln) | 0.026 | 0.657 |
| Borders in conflict (ratio) | 0.094 | 0.171 |
| Infant mortality rate | 0.040 | 0.644 |
| State-led discrimination | 0.027 | 0.16 |
| Average Standard Deviation | 0.054 |  |
| **Older - Wealthier - Moderate Discrimination** | | |
| Polity squared ***OR*** Vdem Index | 0.073 | 0.528 |
| Youth bulge (%) | 0.056 | 0.236 |
| GDP per capita (ln) | 0.022 | 0.873 |
| Borders in conflict (ratio) | 0.049 | 0.148 |
| Infant mortality rate | 0.017 | 0.251 |
| State-led discrimination | 0.050 | 0.361 |
| Average Standard Deviation | 0.045 |  |
| **Higher Discrimination - Worse Neighborhood - Younger** | | |
| Polity squared ***OR*** Vdem Index | 0.155 | 0.3 |
| Youth bulge (%) | 0.031 | 0.754 |
| GDP per capita (ln) | 0.022 | 0.701 |
| Borders in conflict (ratio) | 0.075 | 0.606 |
| Infant mortality rate | 0.027 | 0.535 |
| State-led discrimination | 0.072 | 0.404 |
| Average Standard Deviation | 0.064 |  |

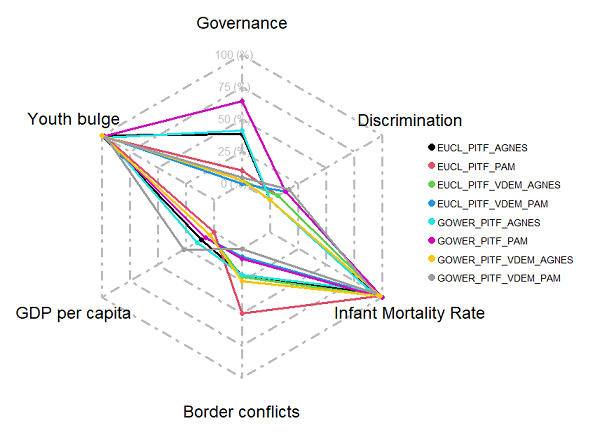
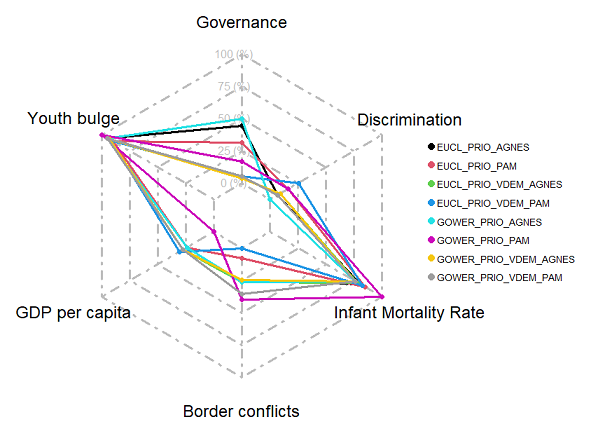
*Table 8:* PITF Cases categorized as having “broad disagreement” (bold text indicates the category under which our base model clustered each observation case). Values in cluster columns reflect the number out of eight possible clustering sensitivity cases. For example, Chile 1973 was clustered in *anocratic-younger-less developed* in three out of eight clustering exercises.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Country** | **Year** | **Anocratic - Younger - Less Developed** | **Older - Wealthier - Moderate Discrimination** | **Higher Discrimination - Worse Neighborhood - Younger** |
| Chile | 1973 | 3 | 4 | 1 |
| Libya | 2011 | 1 | 2 | 5 |
| Congo Kinshasa | 1977 | 4 | 0 | 4 |
| Rwanda | 1990 | 2 | 0 | 6 |
| Algeria | 1991 | 3 | 0 | 5 |
| Senegal | 1992 | 6 | 0 | 2 |
| Congo Brazzaville | 1997 | 5 | 0 | 3 |
| Guinea-Bissau | 1998 | 5 | 0 | 3 |
| Ethiopia | 1999 | 5 | 0 | 3 |
| Guinea | 2000 | 5 | 0 | 3 |
| Rwanda | 2001 | 2 | 0 | 6 |
| Ivory Coast | 2002 | 5 | 0 | 3 |
| Chad | 2005 | 6 | 0 | 2 |
| Ethiopia | 2007 | 5 | 0 | 3 |
| Mexico | 2007 | 0 | 6 | 2 |

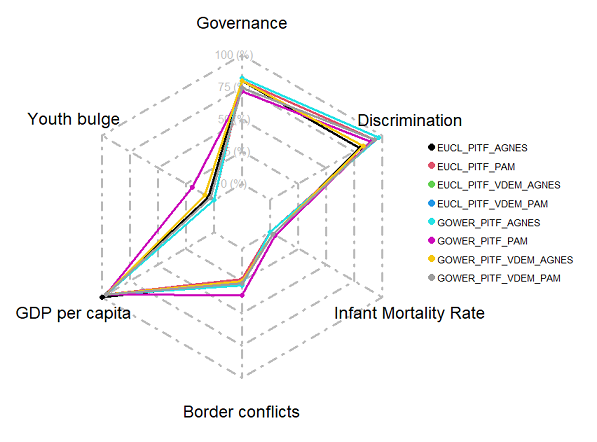
*Table 8.2:* PRIO Cases categorized as having “broad disagreement” (bold text indicates the category under which our base model clustered each observation case).

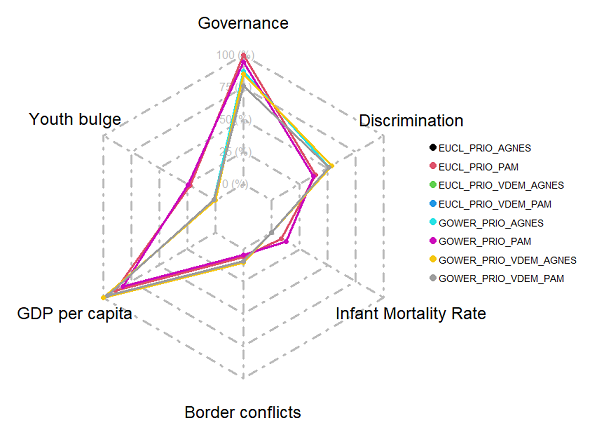
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Country** | **Year** | **Anocratic - Younger - Less Developed** | **Older - Wealthier - Moderate Discrimination** | **Higher Discrimination - Worse Neighborhood - Younger** |
| Sudan | 1963 | 6 | 0 | 2 |
| Morocco | 1975 | 6 | 2 | 0 |
| Congo Kinshasa | 1977 | 3 | 0 | 5 |
| Nicaragua | 1977 | 5 | 0 | 3 |
| Peru | 1982 | 3 | 0 | 5 |
| Sudan | 1983 | 5 | 0 | 3 |
| Cameroon | 1984 | 5 | 0 | 3 |
| Somalia | 1986 | 6 | 0 | 2 |
| Laos | 1989 | 2 | 0 | 6 |
| Paraguay | 1989 | 6 | 2 | 0 |
| Rwanda | 1990 | 5 | 0 | 3 |
| Algeria | 1991 | 6 | 0 | 2 |
| Burundi | 1994 | 6 | 0 | 2 |
| Nepal | 1996 | 2 | 0 | 6 |
| Azerbaijan | 1997 | 5 | 0 | 3 |
| Chad | 1997 | 6 | 0 | 2 |
| Congo Brazzaville | 1997 | 6 | 0 | 2 |
| Congo Brazzaville | 2002 | 5 | 0 | 3 |
| Iraq | 2004 | 6 | 2 | 0 |
| Azerbaijan | 2008 | 5 | 0 | 3 |
| Yemen | 2009 | 6 | 0 | 2 |
| Congo Kinshasa | 2011 | 6 | 0 | 2 |
| Libya | 2011 | 5 | 0 | 3 |
| Azerbaijan | 2012 | 5 | 0 | 3 |
| Burundi | 2014 | 6 | 0 | 2 |

*Figure 1*. Sensitivity radials for “Anocratic - Younger - Less Developed” (Top PITF cases; Bottom PRIO cases)

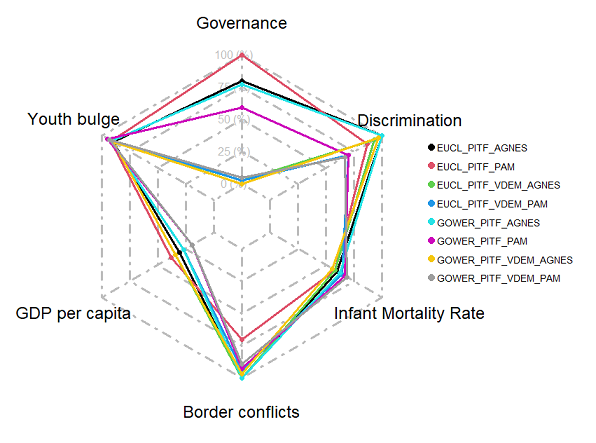
 

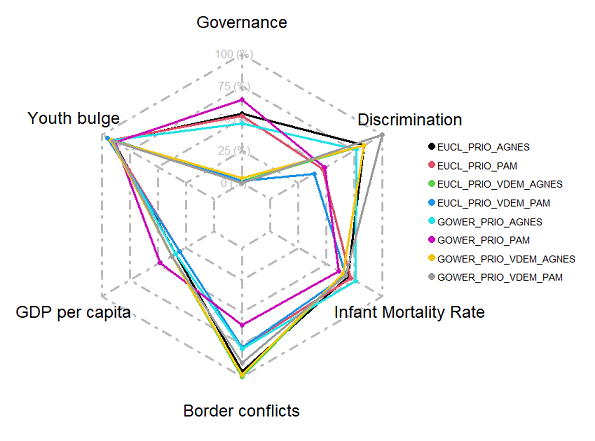
*Figure 2*. Sensitivity radials for “Older - Wealthier - Moderate Discrimination” (Top PITF cases; Bottom PRIO cases)





*Figure 3*. Sensitivity radials for “Higher Discrimination - Worse Neighborhood - Younger” (Top PITF cases; Bottom PRIO cases)





**Explanation of process for generating a conceptually comparable UCDP/PRIO onset sample**

The UCDP/PRIO Armed Conflict Dataset classifies events of state-based armed conflict worldwide from 1946-2020. Armed conflicts in their data are defined as “a contested incompatibility that concerns government and/or territory where the use of armed force between two parties, of which at least one is the government of a state, results in at least 25 battle-related deaths in a calendar year.” Within this set, the creators include four types of armed conflict: 1) extrasystemic; 2) interstate; 3) intrastate; and 4) internationalized intrastate. These conflicts are tracked by the ongoing year and are given descriptive variables (side A, side B, location, etc.) as well as cumulative scale totals of fatalities by ongoing year and cumulatively.

The Political Instability Task Force’s State Failure Problem Set classifies events of state failure worldwide between 1955-2018. Their data are organized at the case-year and inclusion is based on an event being categorized under one (or more) of four distinct types of state failure: 1) revolutionary wars; 2) ethnic wars; 3) adverse regime changes; and 4) genocides and politicides. Events classified as meeting the definitional criteria of at least two failure types are considered “complex” events and counted for both. In general, civil conflict types (excluding adverse regime changes) are qualified if there were at least 1000 people mobilized on each side of the conflict and the event resulted in at least 1000 direct conflict deaths in the span of its operations, of which at least one year must have a value exceeding 100 deaths.

In order to assess the robustness of our cluster findings, we applied our clustering models to civil war onset events for one sample drawing on PITF data and another drawing on UCDP/PRIO data. However, these data sets are not precisely equivalent, so several modeling choices had to be made to produce conceptually similar data samples that could be tested in comparable ways. As our PITF data sample included ethnic and revolutionary wars, some of which were internationalized, we elected to include both intrastate and internationalized intrastate event onsets coded in the UCDP/PRIO set.

The scaled magnitude of the UCDP/PRIO sets did not allow us to draw out precisely the equivalent levels of deaths per year, so we proceeded with sampling rules that we believe met the closest conceptual and quantitative approximation of a comparative sample as could be achieved between the UCDP/PRIO and PITF failure sets. For the UCDP/PRIO Armed Conflict Dataset 22.1, we first removed armed conflict instances that were not civil conflicts, as is included in our original PITF set. We thus removed *type\_of\_conflict* events coded as “1 = extrasystemic” and “2 = intrastate”, leaving only event-years that are “3 = intrastate” and “4 = internationalized intrastate”.

We took a conservative modeling choice for UCDP/PRIO comparison for a numerical death qualifier to determine the date of onset. The PITF dataset requires for inclusion that an event have at least 100 deaths in a single year and 1000 total overall deaths for the cumulative event. Matching UCDP/PRIO events to this was a challenge, because of varying scales for measuring battle deaths. However, we used the *cumulative\_intensity* variable to identify start and end date spells between two parties that eventually resulted in 1000 deaths, similar to PITF. If a spell included at least one year where the *cumulative\_intensity* = 1 (meaning 1000 deaths had been met), then we coded the *start\_date2* as the onset year for the UCDP/PRIO event sample, as this variable marked the first year that at least 25-999 battle deaths were achieved, a scale that fits (as close as can be approximated) the 100 death marker used by PITF. Comparable to our sampling choice for PITF, if two events meeting these two qualifiers had end years and start years that were within two calendar years of one another, we excluded the later onset on the understanding that it was likely a continuation of the previous conflict and not enough time had passed to consider it a conceptually unique onset.

We also note that there are many events that are included in both data sets, some have differing onset years due to these death magnitude qualifiers, both in the original data and in our test samples. Further, there are events coded in one onset sample that do not appear in the other sample, due to inclusion decisions made by the data creators and based on their assessments of which events navigate their decision routes toward coding inclusion. In short, the PITF and UCDP/PRIO samples, even if made as conceptually comparable as possible for our civil war focus, will still differ in observations

**Sensitivity Analysis adding Population and Peace Years**

We estimated model results using two additional variables associated with civil war onset, population size and years since peace. These additional variables were included into the base model using Euclidean distance and Agnes clustering. We maintained three clusters for this analysis, as suggested by our index-based technique.

This analysis confirms our previous work, which shows that young, poorly developed and anocratic states are the primary cluster that also have low levels of peace years and smaller populations. The second cluster also is maintained, with moderately wealthy countries with consolidated regimes, higher populations and a very large number of peace years. The third cluster is also maintained, though with moderate levels of peace years and population size.

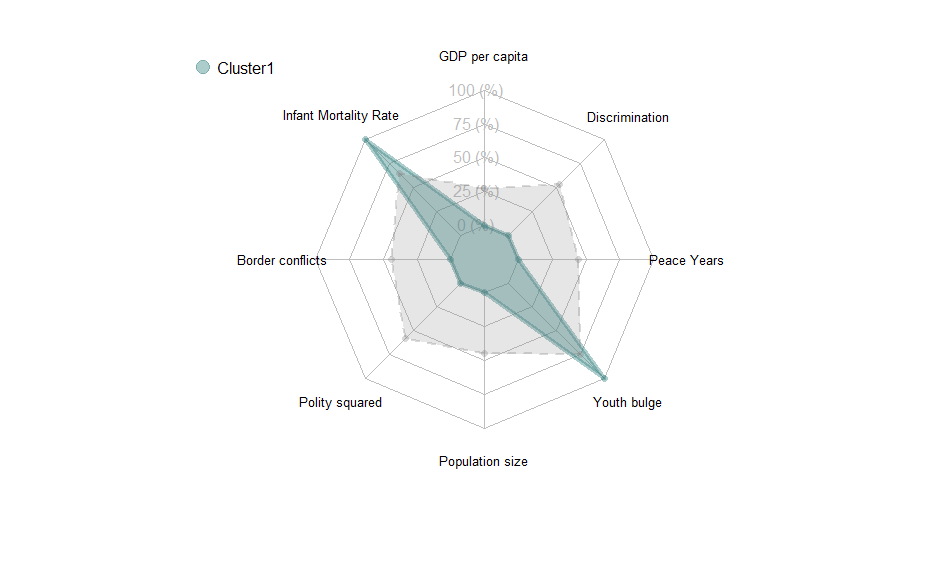
A Summary of these clusters is as follows, with number, central case, and percentage of onsets that were grouped into each:

-Cluster1, Guatemala\_1966, 29 cases, 39.2%.

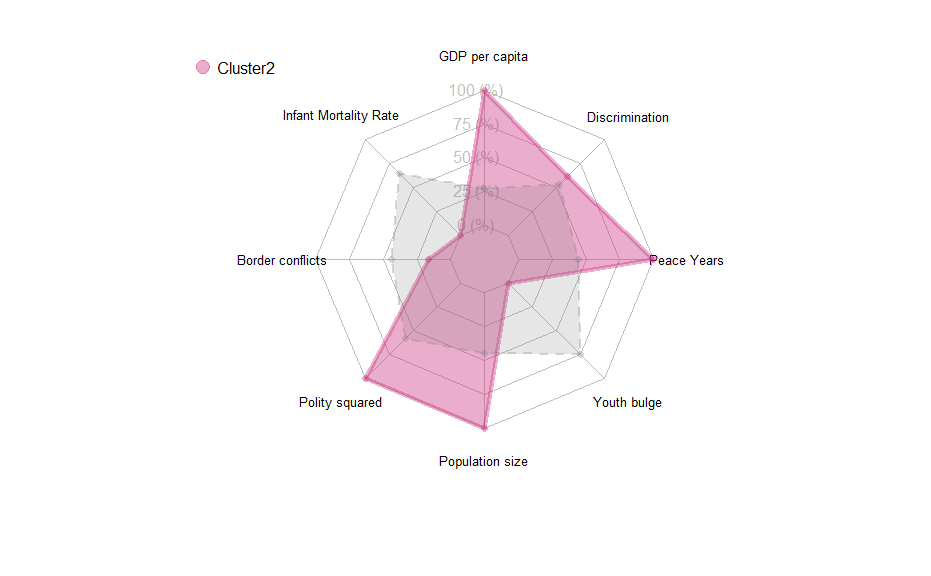
-Cluster2, Argentina\_1976, 15 cases, 20.3%.

-Cluster3, Sudan\_1983, 30, 40.5%

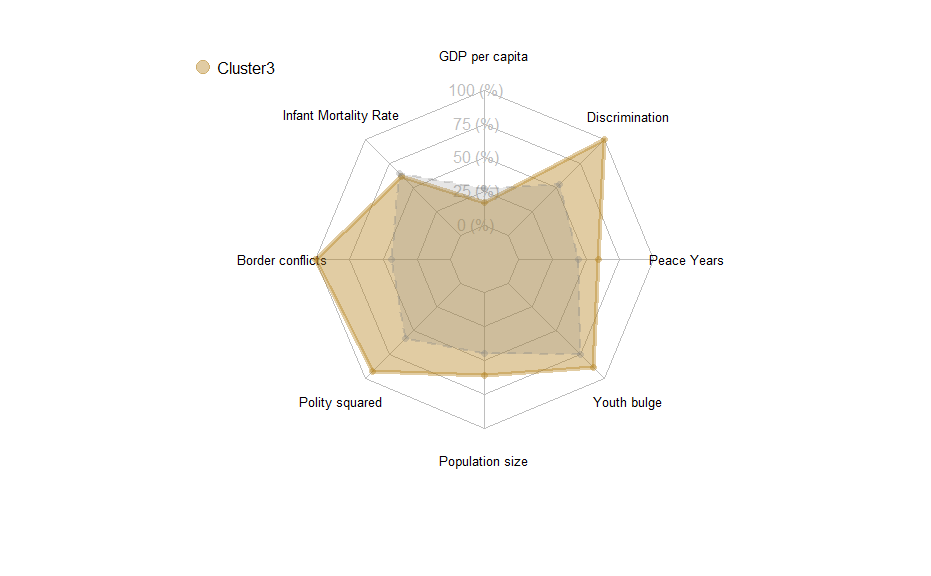
*Figure 4. Cluster 1 for sensitivity analysis including population and peace years.*



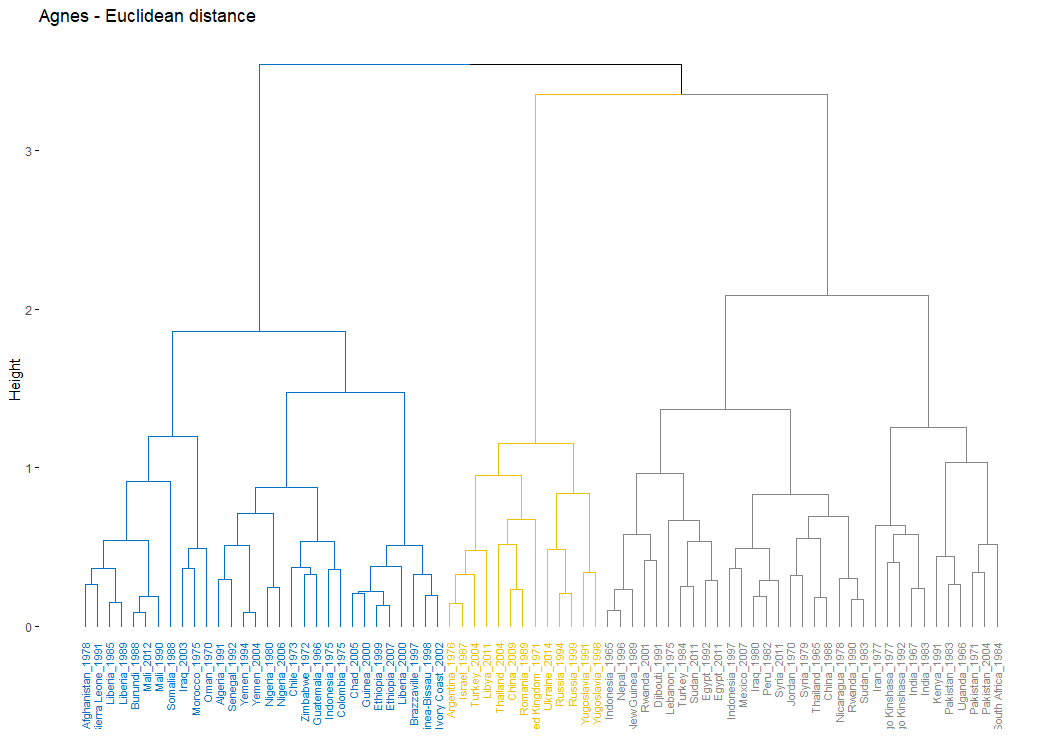
*Figure 5. Cluster 2 for sensitivity analysis including population and peace years.*



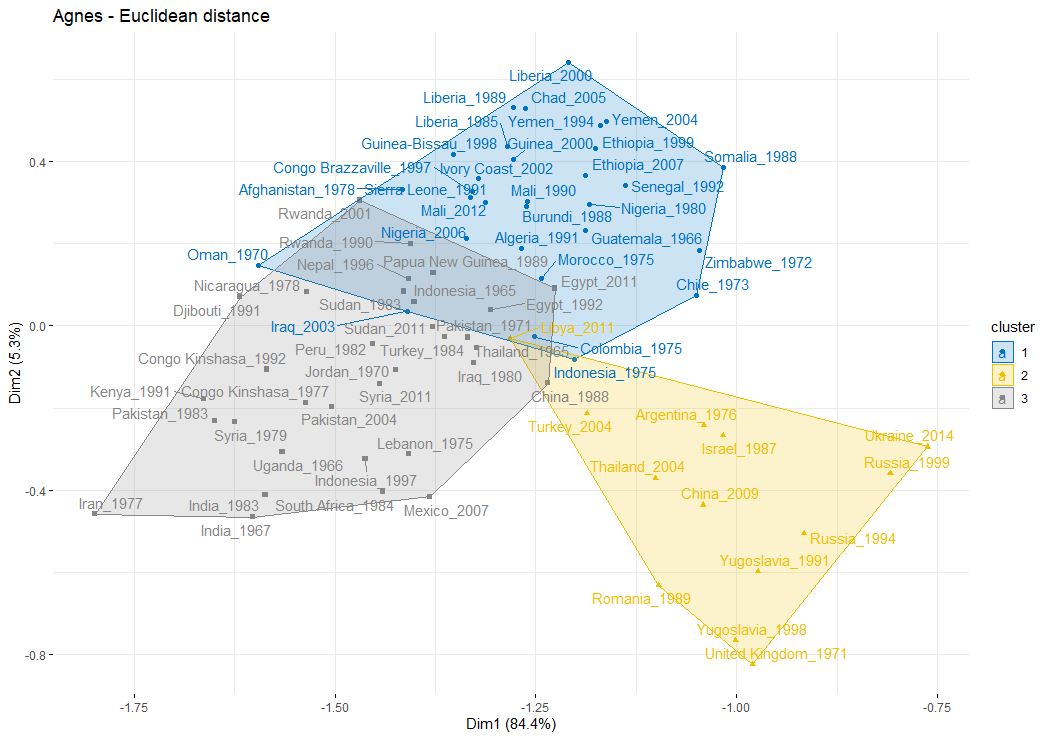
*Figure 6. Cluster 3 for sensitivity analysis including population and peace years.*



***Figure 7.***Example of Dendrogram used in Agnes clustering.



***Figure 8.*** Example of scatter plot based on the first two principal components



**References**

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, nazifa Alizada, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Lisa Gastaldi, Haakon Gerløw, Adam Glynn, Allen Hicken, Garry Hindle, Nina Ilchenko, Joshua Krusell, Anna Luhrmann, Seraphine F. Maerz, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Juraj Medzihorsky, Pamela Paxon, Daniel Pemstein, Josefine Pernes, Johannes von Römer, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Aksel Sundström, Ei-tan Tzelgov, Yi-ting Wang, Tore Wig, Steven Wilson, and Daniel Ziblatt. 2021. *V-Dem [Country-Year/Country-Date] Dataset V11.1*. Gothenberg, Sweden: Varieties of Democracy Institute.

Department of Economic and Social Affairs. 2021. *World Population Prospects*. New York: United Nations.

IMF. 2022. *World Economic Outlook*. Washington, D.C.: International Monetary Fund.

Marshall, Monty G., and Victor Asal. 2007. *Political and Economic Discrimination Against Ethnic Minorities, 1946-2007*. College Park, MD: Minorities at Risk Project, CIDCM, University of Maryland, College Park. 2018 update (provided by Monty G. Marshall).

Marshall, Monty G., Ted Robert Gurr, and Barbara Harff. 2021. “PITF - State Failure Problem Set: Internal Wars and Failure of Governance, 1955-2018.” https://www.systemicpeace.org/inscrdata.html.

Marshall, Monty G., Ted Robert Gurr, and Keith Jaggers. 2016. “Polity IV Project: Dataset Users’ Manual.” http://www.nber.org/ens/feldstein/ENSA\_Sources/CSP/Polity%20Score/p4manualv2015.pdf (June 18, 2019).

World Bank. 2021. “World Development Indicators (WDI): Data Catalog.” https://datacatalog.worldbank.org/dataset/world-development-indicators.