ViEWS Independent Variables

June 5, 2018

Contents

1 Country-level predictors

1.1 GDP per capita (lnGDPpc200)

The GDP per capita in a given country-year. This variable uses historical data stitched together in order to gather information from 1900 until today, and continuing with projections into the future. The variable is constructed with data from Maddison (2010), World Bank (2017) and Samir and Lutz (2008). In order to be able to put the data together into a series we needed to create divisors in order to convert them all into purchasing power parity adjusted 2005 US dollars. First, we harmonize all candidate data to the WDI2005PPP (NY.GDP.MKTP.PP.KD) series (we want our measurement in 2005PPP units). We here assume that the ratio between the candidate data and WDI2005PPP series is constant over time within a country. We then assemble a time-series based on the following preference: WDI2005PPP ¿ WDI2005 Constant US ¿ Maddison WP-4. This preference is based on looking at the data and finding the series that fit best with the WDI data. Third, the SSP projection (second scenario 'Middle of the Road') is added for the future projections. Lastly, we interpolate the log(gdp/cap) data in order to remove missing values where this is possible. This variable is included in the natural logarithm form of the original variable.

1.2 GDP per capita, oilrents only (fvp_lngdpcap_oilrent)

The GDP per capita in a given country-year. This variable captures the percentage of GDP per capita that comes from oil rents in a given country-year. It is calculated using the World Bank Development indicators named Oil rents (NY.GDP.PETR.RT.ZS) (World Bank, 2017) and the population data noted below. For the projections into the future we calculate the average percentage of GDP per capita constituted by oil rents for each country over the last five years of data and assume

that the same rate will continue into the future. This variable is included in the natural logarithm form of the original variable.

1.3 GDP per capita, excluding oilrents (fvp_lngdpcap_nonoilrent)

The GDP per capita in a given country-year. This variable uses historical data stitched together in order to gather information from 1900 until today, and continuing with projections into the future. The variable is constructed with data from Maddison (2010), World Bank (2017) and Samir and Lutz (2008). In order to be able to put the data together into a series we needed to create divisors in order to convert them all into purchasing power parity adjusted 2005 US dollars. First, we harmonize all candidate data to the WDI2005PPP (NY.GDP.MKTP.PP.KD) series (we want our measurement in 2005PPP units). We here assume that the ratio between the candidate data and WDI2005PPP series is constant over time within a country. We then assemble a time-series based on the following preference: WDI2005PPP ¿ WDI2005 Constant US ¿ Maddison WP-4. This preference is based on looking at the data and finding the series that fit best with the WDI data. Third, the SSP projection (second scenario 'Middle of the Road') is added for the future projections. Lastly, we interpolate the log(gdp/cap) data in order to remove missing values where this is possible. Then, the value of fvp_lngdpcap_oilrent is extracted from this value. For the projections into the future we calculate the average percentage of GDP per capita constituted by oil rents for each country over the last five years of data and assume that the same rate will continue into the future. This variable is included in the natural logarithm form of the original variable.

1.4 Growth in GDP per capita, oilrents only (fvp_grlngdpcap_oilrent)

The GDP per capita in a given country-year. This variable captures the growth in GDP per capita that comes from oil rents in a given country-year. It is calculated by taking the difference between the current level of GDP per capita coming from oil rents and the value from the year before. This variable is included in the natural logarithm form of the original variable.

1.5 GDP per capita, excluding oil rents (fvp_grlngdpcap_nonoilrent)

The GDP per capita in a given country-year. This variable captures the growth in GDP per capita that does not come from oil rents in a given country-year. It is calculated by taking the difference between the current level of non-oil GDP per capita and the value from the year before. This variable is included in the natural logarithm form of the original variable.

1.6 Population size (ln_fvp_population200)

The total population of a given country. This variable uses historical data stitched together in order to gather information from 1900 until today, and continuing with projections into the future. The variable is constructed with data from Maddison (2010), World Bank (2017) and Samir and Lutz (2008). First, we harmonize candidate data to the WDI series (we want our measurement in thousands). We then assemble a time-series based on the following preference: WDI ¿ Maddison WP-4. Third, the SSP projection (second scenario 'Middle of the Road') is added for the future projections. Lastly, we interpolate the logged data in order to remove missing values where this is possible. In order to have a gradual transition from one series to another we calculate a gradual exchange between the WDI value and the SSP value from 2007 until 2017. This variable is included in the natural logarithm form of the original variable.

1.7 Democracy index (fvp_democracy)

This index uses three dimensions from V-Dem 7.1 Coppedge et al. (2015) to classify the level of democracy in each country. These dimensions are centered on: free and fair elections (v2x_polyarchy), democratic participation (v2x_partip) and constraints on the executive (v2x_liberal). The higher the performance on any or each of these leads to a higher score on the democracy index, and vice versa. The variable ranges from 0 to one, where one is the highest possible level of democracy.

1.8 Democracy (fvp_demo)

This variable is a dummy variable, indicating whether or not this country is to be regarded a democracy or not. If it does fall into the definition of democracy it is coded one, if not, it is coded 0. This classification is constructed on the basis of **??**. To construct the categories we rely on a cube produced by plotting the three dimensions against each other. We calculate the Euclidean distance from the democracy and autocracy corners and standardize the distance to fall between 0 and 1. We then define as a democracy any country-year that is no further from the democracy corner than 0.40, and as autocratic any country that is less than 0.40 from the autocracy corner. The semi-democracies, thus, are those that are in between.

1.9 Autocracy (fvp_auto)

This variable is a dummy variable, indicating whether or not this country is to be regarded an autocracy or not. If it does fall into the definition of autocracy it is coded one, if not, it is coded 0. This classification is constructed on the basis of **??**. To construct the categories we rely on a cube produced by plotting the three dimensions against each other. We calculate the Euclidean distance from the democracy and autocracy corners and standardize the distance to fall between 0 and 1. We then define as a democracy any country-year that is no further from the democracy corner than 0.40, and as autocratic any country that is less than 0.40 from the autocracy corner. The semi-democracies, thus, are those that are in between.

1.10 Semi-democracy (fvp_semi)

This variable is a dummy variable, indicating whether or not this country is to be regarded a semi-democracy or not. If it does fall into the definition of semi-democracy it is coded one, if not, it is coded 0. This classification is constructed on the basis of **??**. To construct the categories we rely on a cube produced by plotting the three dimensions against each other. We calculate the Euclidean distance from the democracy and autocracy corners and standardize the distance to fall between 0 and 1. We then define as a democracy any country-year that is no further from the democracy corner than 0.40, and as autocratic any country that is less than 0.40 from the autocracy corner. The semi-democracies, thus, are those that are in between.

1.11 Number of months in peace (ln_cw0_SSP1_conflict)

This variable is created in dynasim, calculating the total number of months (or years, depending on the unit of analysis) since the country experienced conflict. Included in natural logarithm form.

1.12 Time since independence (ln_fvp_timeindep)

The total number of years since the country became an internationally recognized sovereign state. This is based on the entrance dates set by Gleditsch and Ward (1999). Included in natural logarithm

form.

1.13 Time since pre-independence war (ln_fvp_timesincepreindepwar)

The total number of years since the country experienced a pre-independence war. This is based on the entrance dates set by Gleditsch and Ward (1999). Included in natural logarithm form.

1.14 Time since regime change (ln_fvp_timesinceregimechange)

The total number of years since the country experienced regime change. Defined by changes between the dummy variables fvp_demo, fvp_semi and fvp_auto. Included in natural logarithm form.

1.15 Proportion of population not in working age (SSP1_Non_WorkAgePopProp)

The proportion of the total population that is below 15 or above 60, implying that they are outside the general age range for working. This variable is constructed using Samir and Lutz (2008) historical data from IIASA.

1.16 Proportion of population between 15 and 24 with at least lower secondary education (ssp2_edu_sec_15_24_prop)

The proportion of the population between 15 and 24 that has completed at least lower secondary schooling implies those that have completed lower or upper secondary school. Those that have attained tertiary education are included in this number. This variable is constructed using Samir and Lutz (2008) historical data from IIASA.

1.17 Proportion of population living in urban areas (ssp2_urban_share_iiasa)

The proportion of the total population living in an urban area. This variable is taken from Samir and Lutz (2008) IIASA data, using historical data collected by the UN.

2 PRIO-GRID-level predictors, in use

2.1 Distance to neighboring country (lnbdist1).

The spherical distance in kilometer from the cell centroid to the border of the nearest land-contiguous neighboring country, based on country border data using cShapes v.0.4-2. (Weidmann, Kuse, and Gleditsch, 2010). Included in natural logarithm form of the original variable.

2.2 Travel time (lnttime).

Log-transformed estimate of the travel time to the nearest major city, derived from a global high-resolution raster map of accessibility developed for the EU (Uchida, 2009). Collected from PRIO-GRID. Included in natural logarithm form of the original variable.

2.3 Distance to capital city (lncapdist).

The spherical distance in kilometers from the cell centroid to the national capital city in the corresponding country, based on coordinate pairs of capital cities derived from the cShapes dataset

v.0.4-2. It captures changes over time wherever relevant (Weidmann, Kuse, and Gleditsch, 2010). Included in natural logarithm form of the original variable.

2.4 Population size (lnpop).

Population size for each populated cell in the grid, taken from the History Database of the Global Environment (HYDE) version 3.1. Population estimates are available for 1950, 1960, 1970, 1980, 1990, 2000, and 2005. The original pixel value is number of persons. Included in natural logarithm form of the original variable. Collected from PRIO-GRID, based on Klein Goldewijk et al. (2011).

2.5 Distance to nearest secondary diamonds resource (ln_dist_diamsec).

Captures the distance from the grid cell to the nearest secondary diamonds resource (static data). The distance is measured in logged WGS86 units (decimal degrees), a unit closely resembling the design choices for the overall grid. The original variable is named diamsec_s and is collected from PRIO-GRID, based on Klein Goldewijk et al. (2011).

2.6 Distance to nearest petroleum resource (ln_dist_petroleum).

Captures the distance from the grid cell to the nearest petroleum resource (static data, and only onshore production). The distance is measured in logged WGS86 units (decimal degrees), a unit closely resembling the design choices for the overall grid. The original variable is named petroleum_s and is collected from PRIO-GRID, based on Klein Goldewijk et al. (2011)

2.7 Gross cell product (gcp_li_mer).

The gross cell product, measured in USD, based on the G-Econ dataset v4.0, last modified May 2011. The original G-Econ data represent the total economic activity at a 1x1 degree resolution, so when assigning this to PRIO-GRID we distribute the total value across the number of contained PRIO-GRID land cells. In border areas, the G-Econ 1x1 degree cells might overlap with PRIO-GRID cells allocated to a neighboring country.To minimize bias, PRIO-GRID only extracts G-Econ data for cells that have the same country code as the G-Econ cell represents. This variable is only available for five-year intervals since 1990 (Nordhaus, 2006). Missing data is filled in through linear interpolations.

2.8 Infant mortality rate(imr_mean).

Mean infant mortality rate. Collected from PRIO-GRID, based on raster data from the SEDAC Global Poverty Mapping project (Storeygard et al., 2008).

2.9 Proportion of mountainous terrain (mountains_mean).

Proportion of mountainous terrain within the cell based on elevation, slope and local elevation range. Collected from PRIO-GRID, taken from a high-resolution mountain raster developed for UNEPś Mountain Watch Report. (Blyth, 2002).

2.10 Urban area (urban_ih_li).

Percentage area of the cell covered by urban area. Collected from PRIO-GRID, based on ISAM-HYDE landuse data(Meiyappan and Jain, 2012). Missing data is filled in through linear interpolations.

2.11 Excluded groups (excluded_li).

Number of excluded ethnic groups (discriminated or powerless) in the grid cell for the given year. Collected from PRIO-GRID, derived from the GeoEPR/EPR 2014 update 2 dataset (Vogt et al., 2015). Missing data is filled in through linear interpolations. This is used for country-level using EPR's non-georeferenced data.

2.12 Agricultural area (agri_ih_li).

Percentage area of the cell covered by agricultural area. Collected from PRIO-GRID, based on ISAM-HYDE landuse data(Meiyappan and Jain, 2012). Missing data is filled in through linear interpolations.

2.13 Barren area (barren_ih_li).

Percentage area of the cell covered by barren area. Collected from PRIO-GRID, based on ISAM-HYDE landuse data (Meiyappan and Jain, 2012). Missing data is filled in through linear interpolations.

2.14 Forest area (forest_ih_li).

Percentage area of the cell covered by forest area. Collected from PRIO-GRID, based on ISAM-HYDE landuse data (Meiyappan and Jain, 2012). Missing data is filled in through linear interpolations.

2.15 Grasslands (savanna_ih_li).

Percentage area of the cell covered by grasslands. Collected from PRIO-GRID, based on ISAM-HYDE landuse data (Meiyappan and Jain, 2012). Missing data is filled in through linear interpolations.

2.16 Shrublands (shrub_ih_li).

Percentage area of the cell covered by shrublands. Collected from PRIO-GRID, based on ISAM-HYDE landuse data (Meiyappan and Jain, 2012). Missing data is filled in through linear interpolations.

2.17 Pasture land (pasture_ih_li).

Percentage area of the cell covered by pasture area, based on ISAM-HYDE land-use data. In PRIO-GRID, this indicator is available for the years 1950, 1960, 1970, 1980, 1990, 2000, and 2010 (Meiyappan and Jain, 2012). Missing data is filled in through linear interpolations.

2.18 State based conflict event (ged_dummy_sb).

Dummy variable indicating whether there was a conflict event (at least one battle related death) in a given grid cell in a given month. State based conflicts, that is intra-state, inter-state, internationalized and extrasystemic conflicts (Sundberg and Padskocimaite, 2010). In the country models this is aggregated up to the country level using cShapes v.0.4-2. (Weidmann, Kuse, and Gleditsch, 2010).

2.19 Non-state conflict event (ged_dummy_ns).

Dummy variable indicating whether a non-state conflict event (at least one battle related death) had occurred within the given grid cell. Non-state conflicts, where non-state armed groups are in conflict with one another (Sundberg and Padskocimaite, 2010).

2.20 One-sided conflict event (ged_dummy_os).

Dummy variable indicating whether there was a conflict event (at least one death) in a given grid cell in a given month. One-sided violence regards cases where government forces or non-state armed groups engage in violence against civilians (Sundberg and Padskocimaite, 2010).

2.21 Protest event (acled_dummy_pr).

Dummy variable indicating whether there was a protest event as defined by ACLED (Armed Conflict Location and Event Dataset) in a given grid cell in a given month (Raleigh and Hegre, 2005).

2.22 Lagged state-based conflict event (l1_12_ged_dummy_sb).

Dummy variable for whether a state-based conflict event had occurred within the given grid cell. First to twelve order temporal lag (previous grid-month). Calculated in dynasim based on ged_dummy_sb.

2.23 Lagged non-state conflict event (l1_12_ged_dummy_ns).

Dummy variable for whether a non-state conflict event had occurred within the given grid cell. First to twelve order temporal lag (previous grid-month). Calculated in dynasim based on ged_dummy_ns.

2.24 Lagged one-sided conflict event (l1_12_ged_dummy_os).

Dummy variable for whether a one-sided conflict event had occurred within the given grid cell. First to twelve order temporal lag (previous grid-month). Calculated in dynasim based on ged_dummy_os.

2.25 Lagged protest event (l1_12_acled_dummy_pr).

Dummy variable for whether a protest event as defined by ACLED (Armed Conflict Location and Event Dataset) had occurred within the given grid cell. First to twelve order temporal lag (previous grid-month). Calculated in dynasim based on acled_dummy_pr.

2.26 Months since last state-based conflict (decay_12_cw_ged_dummy_sb_0).

Exponential decay function for months since the last state based conflict had occurred within the given grid cell. 12 month half-life. Calculated in dynasim based on ged_dummy_sb.

2.27 Months since last non-state conflict (decay_12_cw_ged_dummy_ns_0).

Exponential decay function for months since the last non-state conflict had occurred within the given grid cell. 12 month half-life. Calculated in dynasim based on ged_dummy_ns.

2.28 Months since last one-sided conflict (decay_12_cw_ged_dummy_os_0).

Exponential decay function for months since the last one-sided conflict had occurred within the given grid cell. 12 month half-life. Calculated in dynasim based on ged_dummy_os.

2.29 Spatial lag of state-based conflict event (q_1_1_l1_l3_ged_dummy_sb).

Variable capturing whether a state-based conflict event occurred in the specified previous month in the grid cells surrounding a given grid cell (first to third order). Specified as 'queen', meaning the variable is in effect for all straight and diagonal directions. Calculated in dynasim based on ged_dummy_sb.

2.30 Spatial lag of non-state conflict event (q_1_1_l1_l3_ged_dummy_ns).

Variable capturing whether a non-state conflict event occurred in the specified previous month in the grid cells surrounding a given grid cell (first to third order). Specified as 'queen', meaning the variable is in effect for all straight and diagonal directions. Calculated in dynasim based on ged_dummy_ns.

2.31 Spatial lag of one-sided conflict event (q_1_1_l1_l3_ged_dummy_os).

Variable capturing whether a one-sided conflict event occurred in the specified previous month in the grid cells surrounding a given grid cell (first to third order). Specified as 'queen', meaning the variable is in effect for all straight and diagonal directions. Calculated in dynasim based on ged_dummy_os.

2.32 Spatial lag of protest event (q_1_1_l1_l3_acld_dummy_pr).

Variable capturing whether a protest event as defined by (Armed Conflict Location and Event Dataset) occurred in the specified previous month in the grid cells surrounding a given grid cell (first to third order). Specified as 'queen', meaning the variable is in effect for all straight and diagonal directions. Calculated in dynasim based on acled_dummy_pr.

2.33 Months since last state based conflict (proxevent_sb_lag1).

Exponential decay function for months since the last state based conflict had occurred within the given grid cell. First order temporal lag. Calculated in dynasim based on ged_dummy_sb.

2.34 Months since last state based conflict (proxevent_sb_lag2).

Exponential decay function for months since the last state based conflict had occurred within the given grid cell. Second order temporal lag. Calculated in dynasim based on ged_dummy_sb.

2.35 Months since last non-state conflict (proxevent_ns_lag1)

Exponential decay function for months since the last non-state conflict had occurred within the given grid cell. First order temporal lag. Calculated in dynasim based on ged_dummy_ns.

2.36 Months since last one-sided violence (proxevent_os_lag1)

Exponential decay function for months since the last one-sided violence event had occurred within the given grid cell. First order temporal lag. Calculated in dynasim based on ged_dummy_os.

2.37 Diamond deposits (diamsec_s).

Dummy variable for whether secondary (alluvial) diamond deposits have been found within the given grid cell. Collected from PRIO-GRID, based on the Diamond Resources dataset v1a (Lujala, Gleditsch, and Gilmore, 2005).

2.38 Gem deposits(gem_c).

Dummy variable for whether gem deposits have been found within the given grid cell. Collected from PRIO-GRID, based on the GEMDATA datbase (Lujala, 2009).

2.39 Placer gold deposits (goldplacer_s).

Dummy variable for whether placer gold deposits have been found within the given grid cell. Collected from PRIO-GRID, based on the GOLDATA_L subset of the GOLDDATA v1.2 (Balestri, 2012).

2.40 Surface gold deposits (goldsurface_s).

Dummy variable for whether surface gold deposits have been found within the given grid cell. Collected from PRIO-GRID, based on the GOLDATA_S subset of the GOLDDATA v1.2 (Balestri, 2012).

2.41 Onshore petroleum deposits (petroleum_s).

Dummy variable for whether onshore petroleum deposits have been found within the given grid cell. Collected from PRIO-GRID, based on the Petroleum Dataset v.1.2 (Lujala, Rød, and Thieme, 2007).

2.42 Distance to nearest gem resource (log1p(dist_gem_s_wgs).

Captures the distance from the grid cell to the nearest gem deposit (static data). The distance is measured in logged WGS86 units (decimal degrees), a unit closely resembling the design choices for the overall grid. Based on the variable named gem_s in PRIO-GRID, which in turn gathers it from Lujala (2009).

2.43 Distance to nearest gold placer (log1p(dist_goldplace_s_wgs).

Captures the distance from the grid cell to the nearest gold placer (static data). The distance is measured in logged WGS86 units (decimal degrees), a unit closely resembling the design choices for

the overall grid. Based on the variable named goldplacer_s in PRIO-GRID, which in turn gathers it from Balestri (2012).

2.44 Distance to nearest gold surface resource (log1p(dist_goldsurface_s_wgs).

Captures the distance from the grid cell to the nearest gold surface resource (static data). The distance is measured in logged WGS86 units (decimal degrees), a unit closely resembling the design choices for the overall grid. Based on the variable named goldsurface_s in PRIO-GRID, which in turn gathers it from Balestri (2012).

2.45 Infant mortality rate(log1p(imr_mean).

Log-transformed mean infant mortality rate. Collected from PRIO-GRID, based on raster data from the SEDAC Global Poverty Mapping project (Storeygard et al., 2008).

References

Balestri, Sara (2012). "Gold and Civil Conflict Intensity: evidence from a spatially disaggregated analysis". In:

Blyth, Simon (2002). *Mountain watch: environmental change & sustainable developmental in mountains*. 12. UNEP/Earthprint.

Coppedge, Michael et al. (2015). *Varieties of Democracy (V-Dem) Country Coding Units v6*. Göteborg: V-Dem Institute.

Gleditsch, Kristian S. and Michael D. Ward (1999). "A Revised List of Independent States since the Congress of Vienna". In: *International Interactions* 25.4, pp. 393–413.

Klein Goldewijk, Kees et al. (2011). "The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years". In: *Global Ecology and Biogeography* 20.1, pp. 73–86.

Lujala, Paivi (2009). "Deadly combat over natural resources gems, petroleum, drugs, and the severity of armed civil conflict". In: *Journal of Conflict Resolution* 53.1, pp. 50–71.

Lujala, Päivi, Nils Petter Gleditsch, and Elisabeth Gilmore (2005). "A Diamond Curse? Civil War and a Lootable Resource". In: *Journal of Conflict Resolution* 49.4, pp. 538–562.

Lujala, Päivi, Jan Ketil Rød, and Nadja Thieme (2007). "Fighting over Oil: Introducing a New Dataset". In: *Conflict Management and Peace Science* 24.3, pp. 239–256.

Maddison, Angus (2010). *Background Note on "Historical Statistics"*. URL: `http://www.ggdc.net/MADDISON/Historical\_Statistics/BackgroundHistoricalStatistics\_03-2010.pdf`.

Meiyappan, Prasanth and Atul K Jain (2012). "Three distinct global estimates of historical land-cover change and land-use conversions for over 200 years". In: *Frontiers of Earth Science* 6.2, pp. 122–139.

Nordhaus, William D. (2006). "Geography and macroeconomics: New data and new findings". In: *Proceedings of the National Academy of Sciences of the United States of America* 103.10, pp. 3510–3517. DOI: `10.1073/pnas.0509842103`.

Raleigh, Clionadh and Håvard Hegre (2005). "Introducing ACLED: An Armed Conflict Location and Event Dataset". In: *IGCC Conference Disaggregating the Study of Civil War and Transnational Violence*. San Diego, CA.

Samir KC; Bilal Barakat, Vegard Skirbekk and Wolfgang Lutz (2008). *Projection of Populations by Age, Sex and Level of Educational Attainment for 120 Countries for 2005–2050. IIASA IR-08-xx*. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Storeygard, Adam et al. (2008). "The global distribution of infant mortality: a subnational spatial view". In: *Population, Space and Place* 14.3, pp. 209–229. DOI: `10.1002/psp.484`.

Sundberg Ralph, Mathilda Lindgren and Ausra Padskocimaite (2010). *UCDP GED Codebook version 1.5-2011*. Department of Peace and Conflict Research, Uppsala University.

Uchida, Hirotsugu (2009). "Agglomeration index: towards a new measure of urban concentration". In:

Vogt, Manuel et al. (2015). "Integrating Data on Ethnicity, Geography, and Conflict The Ethnic Power Relations Data Set Family". In: *Journal of Conflict Resolution*, p. 0022002715591215.

Weidmann, Nils B, Doreen Kuse, and Kristian Skrede Gleditsch (2010). "The geography of the international system: The CShapes dataset". In: *International Interactions* 36.1, pp. 86–106.

World Bank (2017). *World Development Indicators*. Washington DC: The World Bank.