The Work-Injury Introduction and Expansion Dataset

Nate Breznau, breznau.nate@gmail.com

Abstract

This workflow brings together various country-time indicators to construct a dataframe for us in the analysis. Many sources are theoretically relevant in the introduction and expansion of work-injury policy. These data are a combination of information about economic development, work-injury policy, the legal end of slavery and forced labor and democratization.

Setup

GWIP

The Global Work-Injury Policy Data.

qwip_v1.csv (Breznau and Lanver 2020)

Independent Nation States

We aim to measure when geographic territories, peoples or empires became independent nation states. This is important for several reasons. The primary reason is that social policy - what we understand as welfare state development - was only implemented by nation states. By definition a welfare state requires a national state governance. The second reason is that prior to nation states and internationally recognized borders, geographies shifted immensely. China, Germany, Poland and France for example included dramatically different geographic spaces during their empire(s).

The modern nation state arguably starts in Europe, possibly with the French Revolution. However, the wartime shifting borders and constant overthrowing of various powers during the prior to the early 19th century means coding 'national independence' during this period is misleading. Therefore our national independence variable has most European countries coded as 1816 (e.g., Germany and France) or later. The US is the oldest in the data with 1783 and China is much 'younger' with 1860 as this marked the moment when international conflicts pushed China into a reactive (and smaller) national policy. These measures are not intended to reflect prevailing historical narratives, rather to trace the orgins of nation states and their geographies as we know them today.

GDP per capita

mpd2018.csv Maddison 2018 version. (Bolt, Inklaar, de Jong and Luiten van Zanden 2018)

The goal is to have a variable that compares economic development (a proxy for industrialization, measured as GDP per capita) in 1900. In hindsight not ideal, but I simply added missing values to the Excel sheet rather than automate this.

1. Setup Frame To account for missing data the following changes were made to the file mpd2018.xlsx Included missing cells to complete time series for all countries (adjusted xlsx file, just added year rows) This was then saved as mpd2018i.xlsx

2. Missing Countries: GDP interpolations

Imputation via similar cases Although not perfect, missing data can be imputed using similar cases.

Target Country	Interpolation Case
Bhutan	Sri Lanka
Belize	Jamaica
Lie cht enstein	Luxembourg
Monaco	Luxembourg
$South \ Sudan$	Sudan
$San\ Marino$	Austria
Andorra	Austria

Lower GDP Islands, mostly S.E. Asia/Pacific

Target Country	Interpolation Case	Notes
Suriname Solomon Islands	St. Lucia Cambodia	1 for 1900 (adj for crash after 2013)
Fiji, Guyana, Samoa,	Indonesia (to 1950)	then 500 for 1900
Marshall Islands, Palau, Micronesia,		
Vanuatu		
Kiribati	Myanmar	(start at 500)

Goggle Data Query

Low/Mid GDP Islands, mostly Latin America/Carribean

Target Country	Interpolation Case	Notes
Bahamas & Brunei	No coparative case	Unique case (Brunei 500 in 1900) - 1945 manufacturing takes off, but strong pre-industrial economy; 1850 - 1000; after it is roughly double Barbados (note, 1994-2001 removed, and 2013-2016 removed due to peculiarities in the Maddison data; 2016 replaced with 35000)
St. Kitts & Nevis & Antiqua & Barbuda	Trinidad & Tobago	· · · · · · · · · · · · · · · · · · ·
St. Vincent & Grenadines & Grenada	Dominican Republic	
Papua New Guinea	Honduras	

Google Data Query

3. Partially Incomplete Time-Series For countries with GDP data that does not go all the way back to 1900, I interpolate 1900 and the linearly fill from this interpolation up to the earliest possible observation.

Interpolate 1900 = 1 Burundi, Gambia, Guinea, Guinea, Bissau, Equit. Guinea, Niger, Angola, Central Afr. Rep., Cameroon, Comoros, Kenya, Mozambique, Rwanda Mauritania, Somalia (The diff. btw. above and below categories has to do with time-series-interpolation over historical accuracy)

Interpolate 1900 = 500 Albania, Benin, Burkina Faso, Botswana, D.R. Congo, Congo, Cabo Verde, Cambodia, Djibuti, Dominica Namibia, Senegal, Sudan, Togo, Chad, Ethiopia, Gabon, Laos, St. Lucia, Liberia, Thailand, Taiwan, Uganda, Madagascar, Mali, Malta, Mongolia, Mauritius, Malawi, Nigeria, Sierra Leone, Sao Tome, Seychelles, Tanzania, Ukraine, Yemen, Zambia, Zimbabwe

Other Interpolations

Target Country	Unique Interpolation
Somalia	2016 = 1700
Trinidad and Tobago, Afghanistan, Bangladesh,	$Cyprus, Lybia, \ Oman, Pakistan$
Bahrain,	
UAE	1950 = 500
UAE	1958 (discovered oil)= 3000
UAE	1960 (oil) = 20000
Qatar	1938 = 500
Qatar	1940 (discovered oil) = 2000
Qatar	1943 (oil) = 10000
Kuwait	1900 = 500
Kuwait	1938 (discovered oil) = 2000
Kuwait	1940 (oil) = 6000
Kuwait	1950 = 9000
Nicaragua, Paraguay, El Salvador	1850 = 1
Guatemala, Honduras, Haiti, Hungary, Korea,	$Puerto\ Rico,\ Romania$
Myanmar,	
Singapore	1850 = 900
Malaysia	1850 = 1000
New Zealand	1840 = 1000
Panama	1850 = 1000
Russia	1850 = 1500
Barbados	1890 = 500
Iceland	1890 = 1000
Ivory Coast	1900 = 1200
Algeria	1870 = 900
Algeria	1900 = 1000
Algeria	1920 = 3500
Algeria	1950 = 6000
Estonia & Kazakhstan	(1946 = 4700)
Kyrgystan, Latvia, Lithuania, Tajikistan,	1950 = 5000
Turkmenistan	
Uzbekistan	1950 = 3000
Moldova	1950 = 1000
Israel	1940 = 3500 (trend)
Luxembourg	1900 = 5000 & 1890 = 3000

4. Russian Territories & USSR Russian territories 1890-1949 (take USSR GDP) *Estonia, Latvia & Lithuania*

Russian territories 1890-1930 (take USSR GDP) Kyrgyzstan, Croatia, Slovenia,

Russian USSR 'Territories' as of 1922 (use USSR 1900+GDP for predictive purposes)

Belarus, Ukraine, and TransCaucasia (Armenia, Azerbaijan, Georgia) (USSR GDP until 1958 to give a transition until 1973, 1st year available in Maddison)

Final data saved as mpd2018.csv

5. Import and Clean

6. Merge GWIP & GDP

Legal Slavery & Forced LLabor

Ola Rosling, Gapminder.org, gapm.io/d_lslave_v1

https://www.gapminder.org/data/documentation/legal-slavery/ file downloaded 24-04-2020 Data countries by year sheet saved as a .csv file $data_legal_slavery_country.csv$

```
abolish <- read.csv(here::here("data", "data_legal_slavery_country.csv"), header=T, check.names=F, stri.
abolish_long <- abolish
#convert strings to years
illegal <- which(abolish=="Illegal", arr.ind=T)</pre>
legal <- which(abolish=="Legal", arr.ind=T)</pre>
abolish[illegal] <-colnames(abolish)[illegal[,2]]
#use the number 3000 to make sure it does not conflict in the row minimum calculation
abolish[legal] <-3000
abolish[4:221] <- as.numeric(unlist(abolish[4:221]))</pre>
#rowmin = year of abolition
abolish$abol <- apply(abolish[,4:221], 1, FUN=min)
abolish$cow_code <- countrycode(abolish$name, "country.name", "cown")
abolish$cow_code <- ifelse(abolish$name=="Serbia", 345, abolish$cow_code)
abolish <- completeFun(abolish, "cow_code")</pre>
abolish <- select(abolish, cow_code, abol)</pre>
wi_df <- left_join(wi_df, abolish, by = "cow_code")</pre>
# Saudi Arabia missing
wi_df <- wi_df %>%
  mutate(abol = ifelse(cow_code==670, 1962, abol))
# Almost all countries had slavery prior to capitalism. This is about chattel and profit-extraction/exc
#Long format
```

Colonial Rule

QoQ Basic Data qog bas cs jan20.csv (Dahlberg et al. 2020)

Variable Description: This is a tenfold classification of the former colonial ruler of the country. We have excluded the British settler colonies (the US, Canada, Australia, Israel and New Zealand), and exclusively focused on Western overseas' colonialism. This implies that only Western colonizers (e.g. excluding Japanese colonialism), and only countries located in the non-Western hemisphere 'overseas' (e.g. excluding Ireland & Malta), have been coded. Each country that has been colonized since 1700 is coded. In cases of several colonial powers, the last one is counted, if it lasted for 10 years or longer. The categories are the following:

- 0. Never colonized by a Western overseas colonial power
- 1. Dutch
- 2. Spanish
- 3. Italian
- 4. US
- 5. British
- 6. French
- 7. Portuguese
- 8. Belgian
- 9. British-French
- 10. Australian

I add Russia as a colonizer (prior to 1922)

11. Russia

```
colony <- read.csv(here::here("data", "qog_bas_cs_jan20.csv"), header=T, stringsAsFactors = F)
colony$cow_code <- colony$ccodecow</pre>
```

```
colony <- select(colony, cow_code, ht_colonial)

# Estonia, Latvia and Lithuania were Russian 'colonies' as territories (from 1710) prior to 1918; Kyrgy
colony$ht_colonial <- ifelse(colony$cow_code == 366 | colony$cow_code == 367 | colony$cow_code == 368 |

wi_df <- left_join(wi_df,colony,by="cow_code")

# code NA as not colonial for now, check later
wi_df$ht_colonial <- ifelse(is.na(wi_df$ht_colonial),0,wi_df$ht_colonial)

# make colonial dummy
wi_df$fcolony <- ifelse(wi_df$ht_colonial==0,0,1)</pre>
```

Merge new additions

Polity IV (Democratization)

PolityData.rds (Polity IV scores) Saved as this file to preserve replicability (as Polity scores get updated regularly)

```
# Last retreived 07-Apr-2020
# saved in case of version changes, see "A02 Data Compile Polity IV.R"
PolityData <- readRDS(here::here("data", "PolityData.rds"))

# fix Germany & Vietnam
PolityData$cow_code <- ifelse(PolityData$cow_code==260,255,PolityData$cow_code)
PolityData$cow_code <- ifelse(PolityData$cow_code==816,817,PolityData$cow_code)

# Introduce longitudinal data
wi_df_long <- select(wi_df, -c(gdpi, gdp_7yr_rate, gdp_7yr_a))
wi_df_long <- left_join(gdpm, wi_df_long, by = c("cow_code"))</pre>
```

```
# set up time-series data
wi_df_long <- wi_df_long %>%
 mutate(indep = 1,
         firstlaw = 1,
         firstrisk = 1,
         firstsocins = 1,
         bluefull = 1)
wi_df_long <- wi_df_long %>%
  group_by(cow_code) %>%
  mutate(indep = ifelse(year < independence, 0, indep),</pre>
         firstlaw = ifelse(year < labor_workinjury_firstlaw,0,firstlaw),</pre>
         firstrisk = ifelse(year < labor_workinjury_firstins,0,firstrisk),</pre>
         firstsocins = ifelse(year < labor_workinjury_first_socins, 0, firstsocins),</pre>
         bluefull = ifelse(year < labor_workinjury_firstlaw_bluecollar_fullcoverage,0,bluefull))</pre>
wi_df_long <- left_join(wi_df_long,PolityData, by = c("cow_code","year"))</pre>
#clean up
wi_df_long <- select(wi_df_long, -c(iso2c,standardized_country,ccode,scode))</pre>
# make colonial rule a form of authoritarian governance
# some cases not coded properly for fcolony
# interpolate polityIV
wi_df_long <- wi_df_long %>%
  group_by(cow_code) %>%
 mutate(fcolony = max(fcolony),
         polity2i = ifelse(year <= independence & fcolony == 1 & is.na(polity2),-10,polity2),</pre>
         polity2i = na.approx(polity2i, maxgap=15, rule = 2)) %>%
  ungroup()
wi_df_long$polity2i <- round(wi_df_long$polity2i, 0)</pre>
wi_df_long <- select(wi_df_long, cow_code, country.x, year, polity2, polity2i, gdpi, independence, ever
rm(gdpm, gdpi,abol_long,illegal, legal,colony)
wi_df_long <- select(wi_df_long, fcolony, everything())</pre>
```

Import/Clean

Missing

Freedom House after indep, 2019 interpolation

COW Code	Name	Interpolated Score	Freedom House Score
31	Bahamas	8	(9/10)
53	Barbados	9	(9.5/10)
54	Dominica	8	(9/10)
55	Grenada	8	(9/10)
56	St Lucia	8	(9/10)
57	St Vincent	8	(9/10)

COW Code	Name	Interpolated Score	Freedom House Score
58	Antigua	7	(8.5/10)
60	St Kitts	8	(9/10)
80	Belize	7	(8.5/10)
346	Bosnia & H.	0	(5/10)
403	Sao Tome	6	(8/10)
591	Seychelle	4	(7/10)
835	Brunei	-4	(3/10)
935	Vanuatu	6	(8/10)
946	Kiribati	8	(9/10)
983	Marshall Is	8	(9/10)
987	Micronesia	8	(9/10)
990	Samoa	6	· /

Other Missing

$\overline{\text{COW}}$		
Code	Name	Unique Interpolations
6	Puerto Rico	- 1900 change to -8 (Foraker Act), 1917 -6 (Jones Act), 1943 2
		(Statehood), 1952 6 Constitution
115	Suriname	1955-1974 linear interpolate
255	Germany	prior to 1868 -6 (simply extend series)
375	Finland	code as Russia prior to 1917
310	Hungary	make equivalent to Germany 1851-1866
205	Ireland	make equivalent to UK prior to 1921
223	Liechtenstein	use Germany
221	Monaco	use France
331	San Marino	use France
232	Andorra	France
290	Poland	1901-1917 (make -10 for theoretical reasons, war, generations)
395	Iceland	make equivalent to Sweden (as analogy) from 1918 onward, -6 prior
338	Malta	8 (9/10) 1946-2016, 0 prior to 1946
712	Mongolia	-7 prior to 1923 (extend series)
339	Albania	-10 prior to indep
731	North Korea	prior to 1947 -10 (remove cases prior to 1901)
732	Korea	prior to 1947 -10
670	Saudi Arabia	extend -10 back before 1926
713	Taiwan	extend -8 back 20 years
817	Viet nam	-10 for missing

British rule where Europeans took over $\,$ Score 2 prior to indep

20 Canada 660 Lebanon - somewhere in between - code 0 prior to 1943 (maybe remove fcolony=1?) $\,900$ Australia

Parts of Yugoslavia (recode 1918-2005)

341 Montenegro

Parts of the USSR use Russian scores 1922-1990 (all) 1918-1921 (only if missing)

316 Czech Rep, 343 Macedonia, 359 Moldova, 366 Estonia, 367 Latvia, 368 Lithuania, 369 Ukraine take Russia for all missing, 370 Belarus, 372 Georgia, 373 Azerbaijan, 701 Turkmenistan, 702 Tajikistan, 703 Kyrgyzstan (make sure to replace imputed values!), 704 Uzbekistan, 705 Kazakhstan

```
# Use case_when in the future
wi df long <- wi df long %>%
  mutate(polity2i = ifelse(cow_code==31 & is.na(polity2i), 8, polity2i),
  polity2i = ifelse(cow_code==53 & is.na(polity2i), 9, polity2i),
  polity2i = ifelse(cow_code==54 & is.na(polity2i), 8, polity2i),
  polity2i = ifelse(cow_code==55 & is.na(polity2i), 8, polity2i),
  polity2i = ifelse(cow_code==56 & is.na(polity2i), 8, polity2i),
  polity2i = ifelse(cow_code==57 & is.na(polity2i), 8, polity2i),
  polity2i = ifelse(cow_code==58 & is.na(polity2i), 7, polity2i),
  polity2i = ifelse(cow_code==60 & is.na(polity2i), 8, polity2i),
  polity2i = ifelse(cow_code==80 & is.na(polity2i), 7, polity2i),
  polity2i = ifelse(cow_code==346 & is.na(polity2i), 0, polity2i),
  polity2i = ifelse(cow_code==403 & is.na(polity2i), 6, polity2i),
  polity2i = ifelse(cow_code==591 & is.na(polity2i), 4, polity2i),
  polity2i = ifelse(cow_code==835 & is.na(polity2i), -4, polity2i),
  polity2i = ifelse(cow_code==935 & is.na(polity2i), 6, polity2i),
  polity2i = ifelse(cow_code==946 & is.na(polity2i), 8, polity2i),
  polity2i = ifelse(cow_code==983 & is.na(polity2i), 8, polity2i),
  polity2i = ifelse(cow_code==987 & is.na(polity2i), 8, polity2i),
  polity2i = ifelse(cow_code==990 & is.na(polity2i), 6, polity2i),
  polity2i = ifelse(cow_code==115, na.approx(polity2i,maxgap=Inf), polity2i),
  polity2i = ifelse(cow_code==255 & year<1869,-6,polity2i),</pre>
  polity2i = ifelse(cow_code==290 & year<1918 & year >1900,-10,polity2i),
  polity2i = ifelse(cow_code==338 & year > 1945,8,polity2i),
  polity2i = ifelse(cow_code==338 & year < 1946,0,polity2i),</pre>
  polity2i = ifelse(cow_code==712 & year < 1924,-7,polity2i),</pre>
  polity2i = ifelse(cow_code==339 & year < 1914,-10,polity2i),</pre>
  polity2i = ifelse(cow_code==731 & year > 1900 & year<1948,-10,polity2i),</pre>
  polity2i = ifelse(cow_code==732 & year < 1948,-10,polity2i),</pre>
  polity2i = ifelse(cow_code==670 & year < 1926,-10,polity2i),</pre>
  polity2i = ifelse(cow_code==713 & year < 1949 & year > 1929,-8,polity2i),
  polity2i = ifelse(cow code==817 & is.na(polity2i),-10,polity2i),
  polity2i = ifelse(cow_code==20 & year < 1867,2,polity2i),</pre>
  polity2i = ifelse(cow_code==660 & year < 1943,0,polity2i),</pre>
  polity2i = ifelse(cow_code==900 & year < 1901,2,polity2i),</pre>
  polity2i = ifelse(cow_code==355 & year < 1879,-4,polity2i),</pre>
  polity2i = ifelse(cow_code==365 & year > 1940 & year < 1948,-10,polity2i))</pre>
wi_df_long <- wi_df_long %>%
  mutate(polity2i = ifelse(cow_code == 6 & year > 1899 & year < 1917, -8, ifelse(cow_code == 6 & year >
         polity2i = ifelse(cow_code == 986 & year < 1947, -10, ifelse(cow_code == 986 & year > 1946 & y
         polity2i = ifelse(cow_code == 713 & year < 1949,-10, ifelse(cow_code == 713 & year > 1948, 8,
wi_df_longx <- wi_df_long</pre>
```

```
wi_df_longx$polity2ii <- wi_df_longx$polity2i</pre>
wi_df_longx <- select(wi_df_longx,cow_code, year, polity2ii)</pre>
# fix Montenegro
wi_m_341 <- subset(wi_df_longx, cow_code==345 & year<2007)</pre>
wi_m_341$cow_code <- 341
# fix Czech/Slovak - Austro-Hungary
wi_m_316a <- subset(wi_df_longx, cow_code==305 & year<1918)</pre>
wi_m_316a$cow_code <- 316
wi_m_317a <- subset(wi_df_longx, cow_code==305 & year<1918)</pre>
wi_m_317a$cow_code <- 317
# Russian/Soviet entities
wi_df_long_R <- subset(wi_df_longx, cow_code==365 & year<1991)</pre>
wi_m_375 <- subset(wi_df_long_R, year<1917)</pre>
wi_m_375$cow_code <- 375
wi_m_316 <- subset(wi_df_long_R, year>1917)
wi_m_316$cow_code <- 316
wi_m_317 <- subset(wi_df_long_R, year>1917)
wi_m_317$cow_code <- 317
wi_m_343 <- subset(wi_df_long_R, year>1917)
wi_m_343$cow_code <- 343
wi_m_359 <- subset(wi_df_long_R, year>1917)
wi_m_359$cow_code <- 359
wi_m_366 <- subset(wi_df_long_R, year>1940)
wi_m_366$cow_code <- 366
wi_m_367 <- subset(wi_df_long_R, year>1940)
wi_m_367$cow_code <- 367
wi_m_368 <- subset(wi_df_long_R, year>1940)
wi_m_368$cow_code <- 368
wi_m_369 <- wi_df_long_R
wi_m_369$cow_code <- 369
wi_m_370 <- wi_df_long_R
wi_m_370$cow_code <- 370
wi_m_372 <- subset(wi_df_long_R, year>1918)
wi_m_372$cow_code <- 372
wi_m_371 <- subset(wi_df_long_R, year>1918)
wi_m_371$cow_code <- 371
wi_m_373 <- subset(wi_df_long_R, year>1918)
```

```
wi_m_373$cow_code <- 373
wi_m_701 <- wi_df_long_R
wi_m_701$cow_code <- 701
wi_m_702 <- wi_df_long_R
wi_m_702$cow_code <- 702
# 703 Kyrgystan (make sure to replace imputed values!)
wi_m_703 <- wi_df_long_R</pre>
wi_m_703$cow_code <- 703
wi_m_704 <- wi_df_long_R
wi_m_704$cow_code <- 704
wi_m_705 <- wi_df_long_R
wi_m_705$cow_code <- 705
# Other Entities
wi_m_310 <- subset(wi_df_longx, cow_code==255 & year > 1851 & year < 1867)
wi_m_310$cow_code <- 310
wi_m_223 <- subset(wi_df_longx, cow_code==255)</pre>
wi_m_223$cow_code <- 223
wi_m_205 <- subset(wi_df_longx, cow_code==200 & year < 1921)</pre>
wi m 205$cow code <- 205
wi_m_221 <- subset(wi_df_longx, cow_code==220)</pre>
wi_m_221$cow_code <- 221
wi_m_331 <- subset(wi_df_longx, cow_code==220)</pre>
wi_m_331$cow_code <- 331
wi_m_232 <- subset(wi_df_longx, cow_code==220)</pre>
wi_m_232$cow_code <- 232
wi_m_395 <- subset(wi_df_longx, cow_code==380 & year > 1850)
wi_m_395$polity2ii <- ifelse(wi_m_395$year < 1918, -6, wi_m_395$polity2ii)
wi_m_395$cow_code <- 395
wi_merge <- do.call("rbind",list(wi_m_205,wi_m_221,wi_m_223,wi_m_331,wi_m_232, wi_m_310,wi_m_316,wi_m_3
rm(list = ls(pattern="^wi\\ m\\ "))
# join interpolated values with main data
wi_df_long <- left_join(wi_df_long,wi_merge,by=c("cow_code","year"))</pre>
wi_df_long <- wi_df_long %>%
  mutate(polity2i = ifelse(cow_code==703 & year<1991,polity2ii,polity2i),</pre>
         polity2i = ifelse(cow_code==316 & year<1993 & year>1989,0,polity2i),
         polity2i = ifelse(cow_code==317 & year<1993 & year>1989,0,polity2i))
rm(wi_df_longx,wi_df_long_R,PolityData,abolish_long,abolish)
```

```
wi_df_long$polity2i <- ifelse(is.na(wi_df_long$polity2i),wi_df_long$polity2ii,wi_df_long$polity2i)</pre>
```

Add Scores

Slavery Adjustments to Rosling data

1. Abolition adjustment

Min center by colonial powerj when otherwise unknown, then add the year the colonial power abolished slavery. It is a rough proxy.

2. Other adjustments

600 Morocco 1961 Slavery 'abolished' in the constitution Goodman, R. David. 2012. "Demystifying 'Islamic Slavery': Using Legal Practices to Reconstruct the End of Slavery in Fes, Morocco." History in Africa 39: 143–74. |- 411 Eq Guinea, effective date 1945, prior to independence

In 1 Dutch (1863)

In 5 British (1834 +4yrs transition period) 52 Trin and Tob 1838 53 Barbados 1838

Dominica it was abolished and then effectively restored 55 Grenada 1838 60 St Kitts 1838

all others except Sierra Leone (1928) do min centering

In 6 France do min centering (1905)

540 Angola recode to 1945

In 8 Belgium do mean centering (1890 in practice)

```
# 1 - 1863
# 5 - 1838
# 6 - 1905
# 8 - 1890
wi_df_long <- wi_df_long %>%
  group_by(ht_colonial) %>%
    mutate(sladj = mean(abol, na.rm = T),
           sladm = min(abol, na.rm=T),
           abolC = abol-sladj,
           abolM = abol-sladm)
wi_df_long <- wi_df_long %>%
           mutate(aboli = ifelse(ht_colonial==1,(abol-abolM)+1863,abol),
           aboli = ifelse(ht_colonial==5,abolC+1838,aboli),
           aboli = ifelse(ht_colonial==5 & aboli<1838,1838,aboli),</pre>
           aboli = ifelse(ht_colonial==6,abolC+1905,aboli),
           aboli = ifelse(ht_colonial==6 & aboli<1905,1905,aboli),</pre>
           aboli = ifelse(ht_colonial==8,abolC+1890,aboli),
           aboli = ifelse(ht_colonial==8 & aboli<1890,1890,aboli),</pre>
           aboli = ifelse(cow_code==52 | cow_code==53 | cow_code==55 | cow_code==60,1838,aboli),
           aboli = ifelse(cow_code==540 | cow_code==411,1945,aboli),
           aboli = ifelse(cow_code==600, 1961, aboli),
           aboli = ifelse(abol<1906,abol,aboli),</pre>
```

```
aboli = round(aboli))

#replace abol in country-level data
wi_df_long_abol <- aggregate(wi_df_long, by = list(wi_df_long$cow_code), FUN = mean, na.rm=T)
wi_df_long_abol <- select(wi_df_long_abol, cow_code, abol)
wi_df_long <- select(wi_df_long, -c(abol))

wi_df_long <- left_join(wi_df_long, wi_df_long_abol)

# Fix cross-sectional data

wi_df <- wi_df %>%
    mutate(abol = ifelse(cow_code==52 | cow_code==53 | cow_code==55 | cow_code==60,1838,abol),
        abol = ifelse(cow_code==540 | cow_code==411,1945,abol),
        abol = ifelse(cow_code==600, 1961, abol),
        abol = ifelse(abol == 3000, 2020, abol), # set three countries that have not made slavery ille
        abol_indep_gap = abol - independence)
```

Potentially useful for future analyses

Countries strongly impacted by the Revolutions of 1848. Countries sending delegates to the IWA's First International 1864-1867 (Austria, Great Britain, France, Germany, Belgium, Spain, US, Italy, the Netherlands, Switzerland and Russia (informally via Bakhuin))

Final Recodes

Former Communist Recodes The communist/socialist variable basically follows the list here:

Iliev, Peter, and Louis Putterman. 2007. "Social Capability, History and the Economies of Communist and Postcommunist States." Studies in Comparative International Development 42(1):36-66. doi: 10.1007/s12116-007-9002-8.

But excludes Zimbabwe and Nicaragua.

Zimbabwe not considered commuist/socialist because the government did not control the means of production and no socialist government was instilled in the 1980 revolution.

Makaye, Peter, and Constantine Munhande. 2013. "Zimbabwe's Socialist Development Experiment 1980-1989." Journal Of Humanities And Social Science 18:63–68. doi: 10.9790/0837-1826368.

Nicaragua was never really socialist, or never fully run by a socialist government as sections of it were changing hands due to Cold War forces.

fmrcmnst.csv constructed as a result

```
# former/current communist list
fmrc <- read.csv(here::here("data", "fmrcmnst.csv"), header = T)
fmrc$cow_code <- countrycode(fmrc$country, "country.name","cown")
fmrc$cow_code <- ifelse(fmrc$country == "Serbia", 345, fmrc$cow_code)
fmrc$commst <- 1
fmrc <- select(fmrc, cow_code, commst)

wi_df_long <- left_join(wi_df_long,fmrc, by = "cow_code")
wi_df_long$commst <- ifelse(is.na(wi_df_long$commst),0,wi_df_long$commst)

wi_df <- left_join(wi_df,fmrc, by = "cow_code")</pre>
```

```
wi_df$commst <- ifelse(is.na(wi_df$commst),0,wi_df$commst)</pre>
wi_df_long <- wi_df_long[order(wi_df_long$cow_code,wi_df_long$year),]</pre>
wi_df_long <- wi_df_long %>%
  group_by(cow_code) %>%
        mutate(slavery = as.numeric(slavery),
        noslave = (slavery - 1)*-1,
         # slavery didn't fill all cells (fix this in earlier code someday)
         slavery = mean(slavery, na.rm = T),
         noslave = mean(noslave, na.rm = T))
wi_df_long <- wi_df_long %>%
  mutate(gdpD7 = gdp_7yr_rate,
         #try period before Soviet
         period1 = ifelse(labor_workinjury_firstlaw<1923,1,0),</pre>
         independence = as.numeric(independence),
         # indep should be truncated for this analysis
         independence = ifelse(independence<1850,1850,independence),</pre>
         # truncate at 1981
         abol = ifelse(abol>1982,1982,abol),
         # year of abolition does not make sense for non-slave production countries
         # but take it as a counterfactual with mean replacement for NA
         abol = ifelse(is.na(abol), mean(abol, na.rm=T),abol),
         # Vietnam & Taiwan missing (Vietnam has history slave production)
         abol = ifelse(cow code==713,1948,abol),
         abol = ifelse(cow code==817,1956,abol),
         noslave = ifelse(cow_code == 713, 1, noslave),
         noslave = ifelse(cow_code == 817, 1, noslave))
wi_df_long <- wi_df_long %>%
  group_by(cow_code) %>%
  mutate(polityD10 = (polity2i+10)-(lag(polity2i, 10)+10),
         polityD5 = (polity2i+10)-(lag(polity2i, 5)+10),
         gdpD10 = gdpi-lag(gdpi, 10),
         blueins = ifelse(firstrisk == 1 & bluefull==1,1,0),
         # if indep is past scutx, recode blueins to indep year
         blueins = ifelse(blueins==1 & year<independence,0,blueins),</pre>
         scut = ifelse(blueins==1 & lag(blueins)==0, year, 0),
         scutf = mean(scut, na.rm = T),
         first2fullf = scut-labor_workinjury_firstlaw,
         first2fullf = ifelse(first2fullf<0,NA,first2fullf),</pre>
         first2fullf = mean(first2fullf, na.rm=T),
         first2fullf = ifelse(is.na(first2fullf), 45, first2fullf),
         f2flog = ifelse(!is.na(first2fullf), first2fullf+1, first2fullf),
         f2flog = log(f2flog),
         first2full = ifelse(first2fullf>40,40,first2fullf),
         scut = ifelse(scut == 0, NA, scut),
         first2full = ifelse(is.na(scut),NA, first2full)) %>%
  ungroup()
```

```
# put first2full back into country-level data
wi_df_long_merge <- aggregate(wi_df_long, by = list(wi_df_long$cow_code), FUN = mean, na.rm = T)</pre>
wi_df_long_merge <- select(wi_df_long_merge, cow_code, first2full, scut)</pre>
wi_df <- left_join(wi_df, wi_df_long_merge, by = "cow_code")</pre>
# some NAs, why?
wi_df_long <- completeFun(wi_df_long, "year")</pre>
wi_df_longa <- wi_df_long %>%
  group_by(cow_code) %>%
  mutate(polity2i_first = ifelse(year == (labor_workinjury_firstlaw - 2), polity2i, NA),
         scut2 = max(scut, na.rm=T),
         polity2i_riski = ifelse(year == (scut2 - 2), polity2i, NA),
         polity2i_risk = mean(polity2i_riski, na.rm = T)) %>%
  ungroup()
\# wi_df_longa \leftarrow select(wi_df_longa, scut, scut2, year, polity2i, polity2i_risk, everything())
wi_df <- select(wi_df, -c(abol))</pre>
wi_df_longa <- select(wi_df_longa, cow_code, polity2i_first, polity2i_risk, noslave, abol)</pre>
wi_df_longa <- completeFun(wi_df_longa, "polity2i_first")</pre>
wi_df <- left_join(wi_df, wi_df_longa, by = 'cow_code')</pre>
# make a variable to include countries that did not yet adopt a full-coverage blue-collar risk-pooling
wi_df <- wi_df %>%
 mutate(first2full = ifelse(first2full == "NaN", NA, first2full),
         scut = ifelse(scut == "NaN", NA, scut),
         scut_na = ifelse(is.na(scut), 2016, scut),
         first2full_na = ifelse(is.na(first2full), 65, first2full)
  )
# Fix period, pre-1923
wi_df$period1 <- ifelse(wi_df$labor_workinjury_firstlaw < 1923, 1, 0)
# for mathematical reasons it is fair to consider pre-1890 law introductions as outliers because democr
wi_df$labor_workinjury_firstlaw_i <- ifelse(wi_df$labor_workinjury_firstlaw < 1891, 1890, wi_df$labor_w
# Some correlations
wi_df$independence <- as.numeric(wi_df$independence)</pre>
# Make non-colonies, non-slavery-based labor forces, abol = NA
wi_df$abol2 <- ifelse(wi_df$fcolony==1,wi_df$abol,NA)</pre>
# years since indep
wi_df$yrs_indep <- wi_df$labor_workinjury_firstlaw - wi_df$independence
# make negative numbers equal to 0 (colonies)
wi_df$yrs_indep <- ifelse(wi_df$yrs_indep < 0, 0, wi_df$yrs_indep)</pre>
```

Communism and Enslave Labor Simultaneously For countries that have both communism and slavery we have to assume that communist movements came later, and that they cannot be effective until they began. Therefore we recode the first law year to be 2 years prior to the independence of the state in the 11 cases where communism and a slave production coexist in a country.

To check this logic we also run a model with Communist and slave societies treated as independent entites where slavery societies are kept in the original coding format.

```
wi_df2 <- wi_df %>%
  mutate(commst = ifelse(commst == 1 & noslave == 0, 0, commst))
wi df3 <- wi df %>%
  subset(commst == 1 & noslave == 0) %>%
  mutate(labor_workinjury_firstlaw2 = ifelse(independence > labor_workinjury_firstlaw, independence - 2
         ffull = ifelse(!is.na(labor_workinjury_firstins) & !is.na(labor_workinjury_firstlaw_bluecollar
         ffull = ifelse(ffull < independence, independence, ffull),</pre>
         first2full_2 = ffull - labor_workinjury_firstlaw2,
         first2full_2 = ifelse(first2full_2 > 40, 40, first2full_2),
         first2full_2 = ifelse(is.na(first2full), NA, first2full_2),
         first2full = first2full,
         labor_workinjury_firstlaw = labor_workinjury_firstlaw,
         noslave = 1,
         cow_code = cow_code+1000) %>%
  select(-c(labor_workinjury_firstlaw2, ffull, first2full_2))
wi_df_comb <- rbind(wi_df2, wi_df3)</pre>
rm(wi_df2, wi_df3)
wi_df <- wi_df %>%
  mutate(labor_workinjury_firstlaw2 = ifelse(commst == 1 & noslave == 0 & independence > labor_workinju
         ffull = ifelse(!is.na(labor_workinjury_firstins) & !is.na(labor_workinjury_firstlaw_bluecollar
         ffull = ifelse(ffull < independence, independence, ffull),</pre>
         first2full_2 = ffull - labor_workinjury_firstlaw2,
         first2full_2 = ifelse(first2full_2 > 40, 40, first2full_2),
         first2full_2 = ifelse(is.na(first2full), NA, first2full_2),
         first2full_na_2 = ifelse(is.na(first2full_2), 65, first2full_2)
```

```
# Adjust GDP to the log, interpolate NA
wi_df <- wi_df %>%
  mutate(gdpik = gdpi/1000,
         gdplog = log(gdpik),
         gdplogm = mean(gdplog, na.rm=T),
         gdplogc = gdplog-gdplogm,
         gdpm = mean(gdpik, na.rm=T),
         gdpc = gdpik-gdpm,
         first2full2 = first2full,
       first2full = ifelse(is.na(first2full) & !is.na(commst), 50, first2full),
         abol2 = ifelse(abol_indep_gap < -50, -50,</pre>
                 ifelse(abol_indep_gap > 50, 50, abol_indep_gap)),
         abol2 = ifelse(abol_indep_gap < - 50, 2,</pre>
                 ifelse(abol_indep_gap < -40, 3,
                ifelse(abol_indep_gap >= -40 & abol_indep_gap < -20, 4,
                ifelse(abol_indep_gap >= -20 & abol_indep_gap <= 0, 5,
                ifelse(abol_indep_gap > 0 & abol_indep_gap < 20, 6,
```

GDP Log and Abolition Time

Work-Injury Law Visualization

```
# Start sequence four years before the first ever law
vr = 1870
yr = seq(yr, 2017, 1)
fig1 <- data.frame(year=yr,</pre>
                 numc=0,
                 numany=0,
                 nuins=0,
                 nusocins=0.
                 numcS=0,
                 numanyS=0,
                 stringsAsFactors=FALSE)
for (year in 1870:2017){
  fig1$numc <- ifelse(fig1$year==year, length(which(wi_df$independence <= year)), fig1$numc)</pre>
  fig1$numany <- ifelse(fig1$year==year, length(which(wi_df$labor_workinjury_firstlaw <= year & wi_df$i
  fig1$nuins <- ifelse(fig1$year==year, length(which(wi_df$labor_workinjury_firstins <= year & wi_df$in
 fig1$nusocins <- ifelse(fig1$year==year, length(which(wi_df$labor_workinjury_first_socins <= year & w
fig1 <- fig1 %>%
  mutate(numcS = numc - (numany-nuins) - nuins,
         numanyS = numany-nuins, # how many have any law but not risk pooling
         numanySA = nuins - nusocins) # how many have risk pooling but not social insurance
```

Prep Data

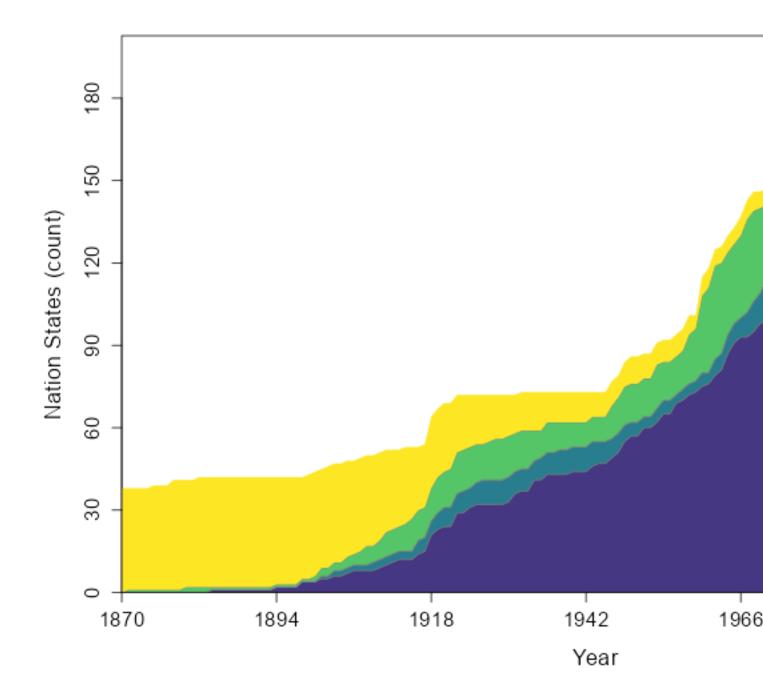
```
agg_png(file = here::here("results", "Fig_Data_Viz.png"), width = 800, height = 550, units = 'px')
ylim <- c(0, 1.09*max(fig1$nuins + fig1$numanyS + fig1$numcS))
xx <- c(fig1$year, rev(fig1$year))
yysrc2 <- c(rep(0, nrow(fig1)), rev(fig1$nuins))

plot(x=fig1$year, y=fig1$nusocins, ylim=ylim, type='l', yaxt ='n',xaxt='n',col = 'white',
ylab='Nation States (count)', xlab='Year',xaxs="i",yaxs="i",cex.lab=1.4)
polygon(xx, yysrc2, col='#453781FF', border = 'grey50')

yysrc0 <- c(fig1$nusocins, rev(fig1$nusocins) + rev(fig1$numanySA))</pre>
```

```
polygon(xx, yysrc0, col='#287D8EFF', border = 'grey50')
rev
Render
```

```
## function (x)
## UseMethod("rev")
## <bytecode: 0x000001af1c7f2408>
## <environment: namespace:base>
yysrc1 <- c(fig1$nuins, rev(fig1$nuins) + rev(fig1$numanyS))</pre>
polygon(xx, yysrc1, col='#55C667FF', border = 'grey50')
rev
## function (x)
## UseMethod("rev")
## <bytecode: 0x000001af1c7f2408>
## <environment: namespace:base>
yysrc3 <- c(fig1$nusocins + fig1$numanyS + fig1$numanySA, rev(fig1$nusocins) + rev(fig1$numanyS) + rev(
polygon(xx, yysrc3, col='#FDE725FF', border = NA)
axis(2, at=c(0,30,60,90,120,150,180),cex.axis=1.3,cex.lab=1.5)
axis(1, at=c(1870,1894,1918,1942,1966,1990,2014), cex.axis=1.3,cex.lab=1.5)
legend(x=1975, y=60, title="Work-Injury Law Type", title.adj = 0.2, c('No Law', 'Employer Liability', '
dev.off()
## pdf
##
```



Save Data for Analysis

```
# Select variables of interest
wi_df <- wi_df %>%
select(cow_code, country_name, independence, labor_workinjury_firstlaw, labor_workinjury_firstlaw_na,
```

```
wi_df_comb <- wi_df_comb %>%
  mutate(gdpik = gdpi/1000,
         gdplog = log(gdpik)) %>%
  select(cow_code, country_name, independence, labor_workinjury_firstlaw, scut, first2full, gdplog, pol
# Create codebooks
wi_df_codebook <- as.data.frame(t(wi_df[1,]))</pre>
wi_df_codebook <- wi_df_codebook %>%
  mutate(variable = row.names(.))
wi_df_codebook <- wi_df_codebook %>%
  mutate(label = ifelse(variable == "cow_code", "Correlates of War numeric 3-Digit ID code",
                 ifelse(variable == "country_name", "Name of country, English",
                 ifelse(variable == "independence", "Nation statehood independence year",
                 ifelse(variable == "labor_workinjury_firstlaw", "First known work-injury law of any ty
                 ifelse(variable == "labor_workinjury_firstlaw_na", "labor_workinjury_firstlaw recoded "
                 ifelse(variable == "scut", "First instance of a risk-pooling law with de jure coverage
                 ifelse(variable == "scut_na", "scut with NAs recoded to 2016",
                 ifelse(variable == "first2full", "Years between first law (labor_workinjury_firstlaw)
                 ifelse(variable == "first2full_2", "first2full recoded so that it is the time since in
                 ifelse(variable == "first2full_na_2", "first2full_2 recoded non-adopters into 65 year
                 ifelse(variable == "gdplog", "Logged GDP per capita in 1900 (Maddison project)",
                 ifelse(variable == "polity2i_first", "Polity IV score at the time of independence (mis
                 ifelse(variable == "abol", "Year of abolition of enslaved/forced labor",
                 ifelse(variable == "abol_indep_gap", "Year of abolition (abol) - year of nationhood in
                 ifelse(variable == "noslave", "No local enslaved production or trading",
                 ifelse(variable == "period1", "Before 1923 = 1, after 1922 = 0",
                 ifelse(variable == "commst", "Successful communist revolution (at least 4 years stable
         ) %>%
  select(variable, label)
write.csv(wi df, file = here::here("data", "wi df.csv"), row.names = F)
write.csv(wi_df_comb, file = here::here("data", "wi_df_comb.csv"), row.names = F)
write.csv(wi_df_codebook, file = here::here("data", "wi_df_codebook.csv"), row.names = F)
\# write.csv(wi_df_long, file = here::here("data", "wi_df_long.csv"), row.names = F)
```

References

Bolt, Jutta, Robert Inklaar, Herman de Jong and Jan Luiten van Zanden (2018), "Rebasing 'Maddison': new income comparisons and the shape of long-run economic development", Maddison Project Working paper 10

Breznau, Nate, and Felix Lanver. 2020. Global Work-Injury Policy Database (GWIP): Project Overview and Codebook. 4. Bremen, Germany: Collaborative Research Center SFB 1342 "The Global Dynamics of Social Policy."

Dahlberg, Stefan, Sören Holmberg, Bo Rothstein, Natalia Alvarado Pachon, and Sofia Axelsson. 2020. The Quality of Government Basic Dataset, Version Jan20. Sweden: University of Gothenburg: The Quality of Government Institute. http://www.qog.pol.gu.se doi:10.18157/qogbasjan20

Rosling, Ola. 2020. "Year of Legal Slavery Abolition". Gapminder.org, gapm.io/d_lslave_v1 https://www.gapminder.org/data/documentation/legal-slavery/ file downloaded 24-04-2020

Colophon

sessionInfo()

```
## R version 4.2.2 (2022-10-31 ucrt)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 19045)
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=English_United States.utf8
## [2] LC_CTYPE=English_United States.utf8
## [3] LC_MONETARY=English_United States.utf8
## [4] LC NUMERIC=C
## [5] LC_TIME=English_United States.utf8
## attached base packages:
## [1] stats
                 graphics grDevices utils
                                               datasets methods
                                                                    base
##
## other attached packages:
  [1] ragg_1.2.5
                          kableExtra_1.3.4 foreign_0.8-83
                                                               ggplot2_3.4.0
   [5] knitr_1.41
                          zoo_1.8-11
                                            countrycode_1.4.0 xlsx_0.6.5
##
  [9] tidyr_1.2.1
                          dplyr_1.0.10
                                            pacman_0.5.1
##
## loaded via a namespace (and not attached):
## [1] tidyselect_1.2.0 xfun_0.36
                                            purrr_1.0.1
                                                               rJava_1.0-6
   [5] lattice_0.20-45
                          colorspace_2.0-3 vctrs_0.5.1
                                                               generics_0.1.3
## [9] htmltools_0.5.4
                          viridisLite_0.4.1 yaml_2.3.6
                                                               utf8_1.2.2
## [13] rlang_1.0.6
                          pillar_1.8.1
                                            glue_1.6.2
                                                               withr_2.5.0
## [17] DBI_1.1.3
                                            stringr_1.5.0
                                                               munsell_0.5.0
                          lifecycle_1.0.3
## [21] gtable_0.3.1
                          rvest_1.0.3
                                            evaluate_0.19
                                                               fastmap_1.1.0
## [25] fansi_1.0.3
                          xlsxjars_0.6.1
                                            scales_1.2.1
                                                               webshot_0.5.4
## [29] systemfonts_1.0.4 textshaping_0.3.6 digest_0.6.31
                                                               stringi_1.7.12
## [33] rprojroot_2.0.3
                          grid_4.2.2
                                            here_1.0.1
                                                               cli_3.6.0
## [37] tools_4.2.2
                          magrittr_2.0.3
                                            tibble_3.1.8
                                                               pkgconfig_2.0.3
## [41] xml2 1.3.3
                          assertthat 0.2.1 rmarkdown 2.19
                                                               svglite 2.1.1
## [45] httr 1.4.4
                          rstudioapi_0.14
                                            R6_2.5.1
                                                               compiler_4.2.2
```