**CODEBOOK for the dataset used in Besley and Kudamatsu (2006) “Health and Democracy”, *American Economic Review Papers and Proceedings*, 96(2), pp.313-318.**

For any inquiry, email Masayuki Kudamatsu at <m.kudamatsu@lse.ac.uk>.

**Description of variables (see Section V “Data Appendix” of the long version of the paper for more details)**

CODE: Three letter country code (as used in Penn World Table)

COUNTRY: Country name (as used in World Development Indicators)

CCODE: Numerical country code (as used in POLITY IV)

YEAR: Years from 1900 through 2002 (from 1990 through 2002 for Germany and Yemen)

CODEYEAR: CODE+YEAR

In the following list of variables in the alphabetical order, the data source is indicated in parentheses for variables obtained from the existing datasets. Those variables without the data source are derived ones (see *cleaning\_pap.do for how these are constructed)*. Variables with \* are not actually used in the econometric analysis, but included for the ease of tracing how variables are constructed.

BALANCED: Dummy for countries with 9 observations on life expectancy over time that were autocratic in 1956 and then democratized only once by 2002 without reversal to autocracy

CDUMx: Country dummies

\*CGDP: Nominal GDP per capita in international dollars (Penn World Table 6.1)

\*COLONY: Dummy for colonial years (ie. POLITY2 is missing)

COLONYF0055: COLONY for 1900-1955 (ie, fraction of colonial years during 1900 to 1955)

DEMOCF0055: DEMOCRACY for 1900-1955 (ie. fraction of democratic years during 1900 to 1955)

\*DEMOCRACY: Dummy for democracy (ie. POLITY2>0)

DEMOCS56: DEMOCRACY since 1956 (ie. fraction of democratic years from 1956 to the present year)

DEMOCSTM4: DEMOCRACY since t-4 (ie. fraction of democratic years from 4 years ago to the present year)

\*EAP: Dummy for East Asia and Pacific

ECA: Dummy for Eastern Europe and Central Asia

\*GDPDEF: GDP Deflator (ie. CGDP/RGDPCH)

\*HEXPGOVPPP: Government health expenditures per capita in current international dollars (World Health Reports 2002, 2004, and 2005)

INFMORT: Infant mortality (World Development Indicators 2005)

\*IMM\_DPT: DPT immunization (World Development Indicators 2005)

\*IMM\_MSL: Measles immunization (World Development Indicators 2005)

INCOME: INCOME (ie. RGDPCH1000 averaged for years t-4 to t)

INCOME2: INCOME squared

LAC: Dummy for Latin America and the Caribbean

LDEMOCS56: DEMOCS56 lagged five years.

\*LEGOR\_FR: Dummy for French legal origin (La Porta et al. 1999)

LEGOR\_GE: Dummy for German legal origin (La Porta et al. 1999)

LEGOR\_SC: Dummy for Scandinavian legal origin (La Porta et al. 1999)

LEGOR\_SO: Dummy for Socialist legal origin (La Porta et al. 1999)  
LEGOR\_UK: Dummy for British legal origin (La Porta et al. 1999)

LIFEEXP: Life expectancy at birth (World Development Indicators 2005)

MEANGOVHEXP: Average real government health expenditures per capita during 1996-2000

MEANIMM\_DPT: IMM\_DPT averaged over 5-year periods (1981-5, 1986-90, 1990-5, 1996-2000)

MEANIMM\_MSL: IMM\_MSL averaged over 5-year periods (1981-5, 1986-90, 1990-5, 1996-2000)

MENA: Dummy for Middle East and North Africa

\*POLITY2: POLITY score (POLITY IV version 2003)

\*REALGOVHEXP: Real government health expenditures per capita in constant 1996 international dollars

\*RGDPCH: Real GDP per capita in constant 1996 international dollars (Penn World Table 6.1)

\*RGDPCH1000: Real GDP per capita in 1,000 constant 1996 international dollars

SA: Dummy for South Asia

SANITATION: Access to improved sanitation facilities (World Development Indicators 2005)

SCHOOLING60: Average years of schooling in the population aged over 15 in 1960 (Barro and Lee 2000)

SSA: Dummy for Sub-Saharan Africa

WATER: Access to clean water sources (World Development Indicators 2005)

WESTERN: Dummy for Western Europe and North America

YDUMx: Dummy for year 1900+x where x ranges from 1 to 102