

# Codebook for “Why do leaders nationalize the oil industry?” (*Energy Policy* 75 (Dec 2014): 228–243)

Paasha MAHDAVI

[Refer to manuscript for full bibliographical information of data sources]

**code** World Bank three-letter country identifier.

**id** Correlates of War numeric country identifier.

**year** Numeric year. The **time** variable is a transformation of the **year** variable, where  $time = year - 1945$ .

**NOCyear** Binary indicator for year of nationalization as defined by the timing of NOC establishment. Source: author.

**NOCsum1** Sum of nationalizations in the prior year for all countries in the sample in that year.

**NOCtotal** Cumulative sum of nationalizations in all prior years for all countries in the sample.

**guriev** Binary indicator for an act of expropriation in a given year. Source: Guriev, Kolotilin, and Sonin (2011).

**oilprod** Annual oil production in metric tonnes per year. Source: Ross (2012) via the United States Geological Survey.

**logoilprod** Logged oil production (natural log). Calculated from **oilprod**.

**dogprod** Year-to-year growth in oil production. Calculated from **oilprod**.

**yrsfirstprod** Years since first commercial production of oil. Calculated using **oilprod**.

**residp** De-trended oil price. The formula for creating this residual is  $\ln p_t = \alpha * \ln p_{t-1} + \beta_1 + \beta_2 * t + \beta_3 * t^2 + \epsilon_t$  where  $p_t$  is the price of oil at time  $t$  and  $p_{t-1}$  is the lagged price of oil, for each year  $t \in [1945, 2005]$ . The deviation from this price trend is the corresponding price shock, so we can estimate the shock by computing yearly residuals,  $\epsilon_t$ . Source for nominal price data: British Petroleum *Statistical Review of Energy*.

**opec** Binary indicator for whether a given country is a member of the Organization of the Petroleum Exporting Countries in a given year. Source: [opec.org](http://www.opec.org).

**regime** Indicator of regime type, measured using the Polity IV coding. Re-scaled to range from 0 to 10 (original scale ranges from -10 to 10, re-scaled by adding 10 and dividing by 2). Source: Marshall and Jaggers (2006).

**agereg** Indicator of regime age, measured as the duration in years of the current governing regime. Source: Cheibub, Gandhi, and Vreeland (2010).

**lngdp** Logged gross domestic product per capita (natural log). Source: Maddison (2007).

**growth** Annual growth (percent) of GDP per capita. Calculated from **lngdp** after exponentiating.