

# Take What You Can: Property Rights, Contestability and Conflict

Thiemo Fetzer and Samuel Marden

## Replication Read-me and Supplementary Data Appendix

### 1. Replication Read-Me

The data and do files required to replicate the contents of the tables and figures in the paper accompany this document. Table 1, below outlines which do file creates what, and the accompanying data file. Each replication and do file are a self-contained pair.

| Table 1.        |                    |  |
|-----------------|--------------------|--|
| Do file         | Accompanying Dta   | Tables and Figures Created                                   |
| replication1.do | municipalpanel.dta | Tables 1, 2A, 3A, 4A, 5A, A1, A2, A4A A5A A6A A7A, Figure 3. |
| replication2.do | event.dta          | Tables 2B, 4B, 5B, A3, A4B, A5B, A6B, A7B, Figure 5          |
| replication3.do | eventspill.dta     | Tables 3B, 3C.   |
| replication4.do | pixel.dta          | Table 6, A10, A11, A12, A14.                                 |

### 2. Supplementary Data Appendix

The supplementary data appendix describes the source of the data for each of the variables in the dta's listed above as well as any processing of the raw data that has taken place. Our main conflict outcome data was digitised from pdf sources and many of our variables of interest were generated from geographical data using ArcGIS and Google Earth Engine. We describe the steps required to process the data and the original sources here and in Online Appendix A. These two documents should be read together.

#### 2a. Variables in municipalpanel.dta

*GEOCODIGO*: The 7 digit municipality code which uniquely identifies municipalities in the data and can be used to link to other data obtained from IBGE (Brazilian Institute of Geography and Statistics) and other official sources.

*id*: an alternative municipal identifier

*state*: 2 digit numeric code indicating the state.

*statecode*: 2 letter code indicating the state.

*year*: The year.

*stateyear*: State-by-Year Identifier

*all\_area*: Area (sqkm) protected by either ecological or indigenous areas in a municipality. This, and the following two variables were created using maps of Brazilian protected areas in ArcGIS as described in Online Appendix A. All areas were calculated based on the *SAD69/Brazil Polyconic* projection.

*indi\_area*: Area (sqkm) protected by either indigenous areas. These are areas classified as the following ('Acquired Indigenous Area' or 'Dominial Indigenous Area' or 'Indigenous Area' or 'Indigenous Reserve').

*eco\_area*: Area (sqkm) protected by either ecological areas. These are areas not classified as indigenous areas as described above.

*AREA\_M*: Area (sqkm) of the municipality. These areas were calculated in ArcGIS using the census 2000 municipal boundaries shapefile available from IBGE (see Appendix A). Consequently, these areas may differ slightly from official estimates of municipal area.

*all\_areaSHARE*: =  $\text{all\_area}/\text{AREA\_M}$

*indi\_areaSHARE*: =  $\text{indi\_area}/\text{AREA\_M}$

*eco\_areaSHARE*: =  $\text{eco\_area}/\text{AREA\_M}$

*homicides*: Homicides (count) obtained from the Mortality Information System and downloaded from <http://tabnet.datasus.gov.br/cgi/deftohtm.exe?sim/cnv/obt10br.def> on February 24, 2016

*indigenoumurder*: Homicides of indigenous peoples (count) obtained as above.

*conflicts* = the CPT “disputes” variable described in Appendix A.

*murders*: the CPT “murders” variable described in Appendix A.

*attemptedmurders*: the CPT “attempted murders” variable described in Appendix A.

*deaththreats*: the CPT “Attempted Murders” variable described in Appendix A.

*conflictevents*: the “Escalations” variable, =  $\text{murders} + \text{attemptedmurders} + \text{deaththreats}$ .

*murdersattempted*: the “Violence” variable, =  $\text{murders} + \text{attemptedmurders}$ .

*anyviolence*: dummy variable indicating  $\text{murdersattempted} > 0$ .

*logPIB*: Log Municipal GDP, underlying data from IBGE ([http://downloads.ibge.gov.br/downloads\\_estatisticas.htm](http://downloads.ibge.gov.br/downloads_estatisticas.htm)).

*logAGRI*: Log Municipal Agricultural GDP, underlying data from IBGE as above.

*logVA\_PUBLIC*: Log Municipal Public Sector Value Added, underlying data from IBGE as above.

*finest*: count of total environmental fines issued in municipality in each year. Data extracted from the database at <https://servicos.ibama.gov.br/ctf/publico/areasembargadas/ConsultaPublicaAreasEmbargadas.php> on March 27, 2016.

*florafines*: count of total environmental fines issued for “flora” offences in municipality in a given year. Obtained as above.

*everblacklisted*: Dummy indicating whether Municipality was ever ecologically blacklisted. List of ever blacklisted municipalities obtained from Cisneros et al. (2015).

*nonprotincaareashare*: Share of municipal land in year  $t$  which is an INCRA designated settlement but not ecologically or indigenously protected. Calculated on an annual basis from a shapefile indicating the location and

boundaries of INCRA settlements downloaded from <http://painel.incra.gov.br> on February 19, 2015. For each year, we intersected the extent of INCRA settlements with the extent of protected areas and the municipal boundaries shapefile and then calculated the share of each municipality that was both protected (i.e. part of *all\_areaSHARE* and unprotected. The variable is the latter share if municipal area. The variable below is the former.

*protincaareashare*: Share of municipal land in year t which is an INCRA designated settlement and ecologically or indigenously protected. See previous variable for details.

*logtotalexp*: Log Public Sector Expenditure (2000-2010) underlying data from Brollo et al. (2013).

*prd010*: the share of the municipality with 10km of a protected area. This variable was created using ArcGIS using the census 2000 municipal boundaries and protected area shapefiles described above.

*present\_event*: dummy indicating whether the municipality appears in the 'event study' data.

*shareicmsecologico*: share of parent state ICMS revenues transferred on the basis of environmental services. Data obtained from <http://www.icmsecologico.org.br/site/> on November 20, 2015.

*all\_area\_share\_state*: Municipal Share of State Protected Area (MSP) calculated from protected area and boundaries shapefiles described above.

*transferproxy*: = *shareicmsecologico* x *all\_area\_share\_state*

## 2b. Variables in event.dta

Variables which have the same name and description as above are omitted.

*protareaid*: Protected Area Identifier

*mid*: Municipality Identifier

*protgeocodigo*: Municipality X Protected Area Identifier.

*protidyr*: Protected Area X Year Identifier.

*intent*: the share of the municipality protected by the associated protected area. This variable calculated using ArcGIS and the boundary and protected area data described above.

*treatintent*: intent X dummy indicating the protected area has been established.

*indigtreatintent*: intent X dummy indicating the protected area has been established X dummy indicating the protected area was a tribal protected area.

*ecotreatintent*: intent X dummy indicating the protected area has been established X (1 - dummy indicating the protected area was a tribal protected area).

*treatintent2*: as *treatintent*, except intent is the share of land in the municipality that was unprotected at t-1.

*tttXX\_intent*: intent X time to establishment dummy (where 'time to establishment' = XX-14).

*timetotreat3d*: dummy indicating protected area was established between 2000-2007.

*Incraaffected*: dummy indicating  $nonprotincaareashare + protincaareashare > 0$ .

*ecotrandum*: dummy indicating at least one municipality intersecting with the protected area lies in a state that at some point has provided ICMS transfers to municipalities based on ecological services. Data on state level ICMS transfers as above.

### 2c. Variables in eventspill.dta

Variables which have the same name and description as above are omitted.

This dataset stacks the three sets data required for Table 2b (columns 1-4), 3a and 3b, to facilitate statistical tests comparing the coefficients.

*degrees*: identifier indicating whether this is data required for 0, 1<sup>st</sup> or 2<sup>nd</sup> degree adjacency regressions (see p. 22 of the main paper for a description of what is meant by adjacency) .

*geo\_ob\_grp*: *protgeocodigo* X *degrees* identifier.

*protidyr\_grp*: *protidyr* X *degrees* identifier.

*GEOCODIGO*:*\_grp* municipality X *degrees* identifier.

*objectid\_grp*: protected area X *degrees* identifier.

*treatintent1*: this is the standard *treatintent* variable described above if *degrees*=0. Takes a value of 0 otherwise.

*treatintent2*: this is the standard *treatintent* variable, but in a dataset where first degree adjacent municipalities are included in the set of control municipalities as described on p. 22. of the main paper. As first degree municipalities are not affected by the protected area in question, adjacent areas take a value of 0. Takes a value of 0 if *degrees*!=1.

*treatintent3* this is the standard *treatintent* variable, but in a dataset where first and second degree adjacent municipalities are included in the set of control municipalities as described on p. 22. of the main paper. As first and second degree municipalities are not affected by the protected area in question, these adjacent areas take a value of 0. Takes a value of 0 if *degrees*!=2.

### 2c. Variables in pixel.dta

Sample of 793, 928 randomly drawn points (pixels) in the Amazon states. The unit of observation is a single point. Data processed in ArcGIS and Google EarthEngine. Unless otherwise stated, distances calculated as straightline distance based from maps projected into SAD1969.

*x*: Longitude (x) of the pixel.

*y*: Latitude (y) of the pixel.

*dprotected0110*: dummy indicating whether the pixel was newly protected between 2001 and 2010. Location of protected areas obtained from shapefiles described above.

*dprotected0310*: Newly Protected Between 2003 and 2010.

*dprotected0308*: Newly Protected Between 2003 and 2008.

*dprotected0107*: Newly Protected Between 2001 and 2007.

*dprotected0109*: Newly Protected Between 2001 and 2009.

*dprotected0108*: Newly Protected Between 2001 and 2008.

*elev*: Mean elevation of the gridcell the pixel falls in. Data from Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010). Provides a grid for elevation and variance in elevation at a 7.5 arc second resolution. Downloaded Wednesday 2 September, 2015 [http://topotools.cr.usgs.gov/GMTED\\_viewer/](http://topotools.cr.usgs.gov/GMTED_viewer/).

*elev\_var*: Variance of elevation for the gridcell the pixel falls in. Source as *elev*.

*RIVER\_D*: distance of ln(KM to River + 1). Larger rivers only (>10m). GIS map of river location from Natural Earth <http://www.naturalearthdata.com/downloads/10m-physical-vectors/10m-rivers-lake-centerlines/>. Version 3.0. Downloaded on September 3, 2015.

*geocodigo*: identifier indicating the municipality the pixel falls in, based on 2000 census boundaries shapefile as described above.

*syrr*: state identifier, based on 2000 census boundaries shapefile as described above.

*ss\_1*: dummy indicating pixel in AC.

*ss\_2*: dummy indicating pixel in AM.

*ss\_3*: dummy indicating pixel in AP.

*ss\_4*: dummy indicating pixel in MA.

*ss\_5*: dummy indicating pixel in MT.

*ss\_6*: dummy indicating pixel in PA.

*ss\_7*: dummy indicating pixel in RO.

*ss\_8*: dummy indicating pixel in RR.

*ss\_9*: dummy indicating pixel in TO.

*forested*: dummy indicating pixel was forested in 2001. Based on MODIS landcover rasters at native resolution (Channan et al., 2014). A cell is classified as forested if it falls in MODIS categories 1-5. Data available from <http://glcf.umd.edu/data/lc/>.

*cropland*: dummy indicating cell was cropland in 2001 (MODIS classifications 12 and 14). Source as before.

*shrubland*: dummy indicating cell was shrubland or grassland in 2001 (MODIS classification 6-10). Source as before.

*dforested0110*: change in MODIS forest cover status between 2001 and 2010 (as defined for *forested*). 1 = transition to forested status. 0 = no change. -1 = deforestation.

*longtermmeanrainfall*: Annual Rainfall Average computed from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS), available from <http://chg.geog.ucsb.edu/data/chirps/>

*longtermmedtemp*: Annual Temperature Average computed from the MODIS Daily Global Surface Temperature (MOD11a1) product, available from 2001 onwards, available from [https://lpdaac.usgs.gov/dataset\\_discovery/modis/modis\\_products\\_table/mod11a1](https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod11a1)

*sdlongtermrainfall*: Annual Rainfall Standard Deviation, standard deviation computed from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS), available from <http://chg.geog.ucsb.edu/data/chirps/>

*Indist*: The natural log of distance (in KM + 1) to a national highway in 2000. Based on 2013 network shapefile available from DNIT <http://www.dnit.gov.br/mapas-multimodais/shapefiles>, accessed March 17, 2015. Road network rolled back to 2000 using annual descriptions of past road networks contained in excel files available from DNIT and downloaded from <http://www.dnit.gov.br/sistema-nacional-de-viacao> March 17, 2015.

*Inq\_casv*:  $\ln(\text{GAEZ Yield Cassava} + 1)$ . All yields obtained from raster available from the FAO's GAEZ database (download from <http://gaez.fao.org/Main.html#>). We use yields based on "Intermediate Level Inputs" and "Rain Fed Agriculture" over the baseline period.

*Inq\_maiz*:  $\ln(\text{GAEZ Yield Maize} + 1)$

*Inq\_ricw*:  $\ln(\text{GAEZ Yield Wetland Rice} + 1)$

*Inq\_ricd*:  $\ln(\text{GAEZ Yield Dryland Rice} + 1)$

*Inq\_soyb*:  $\ln(\text{GAEZ Yield Soybean} + 1)$

*Inq\_oilp*:  $\ln(\text{GAEZ Yield Oil Palm} + 1)$

*Inpq\_max*:  $\ln(\max\{\text{GAEZ Yield} \times \text{Price}\} + 1)$  for the set of crops above. Yields obtained from GAEZ as before. Prices from FAOstat in Brazil for 2000 except for Oilpalm prices which are the average of Mexico, Peru and Ecuador. GAEZ yields converted from dry to wet weights using conversion ratios from the GAEZ model documentation (Fisher et al., 2012, p. 98).

*distnlightpxls*: Distance to Nightlights. Raster derived from discretized night light image capturing lit pixels available from <https://www.ngdc.noaa.gov/eog/dmsp/downloadV4composites.html>.

*distnlightpxls2*: Distance to Nightlights Squared [see above].

*pdiff*: difference in propensity scores for matched pixels (no replacement). Treated pixels are *dprotected0110*=1, control *dprotected0110*=0. Propensity scores from regression in Table A10.

*dforestuse0108*: Change in Forested Status 2001-05 vs 2008-12. Based on changes in classifications of sequences of MODIS land cover data (data source as before) between 2001-05 and 2008-12. 1=change from an un-forested sequence to a forested one. 0=no change in classification. -1=a change from forested to un-forested. See Appendix A3 for further details on classification of sequences.

*dforestuse0107*: Change in Forested Status 2001-06 vs 2007-12.

*dforestuse0109*: Change in Forested Status 2001-04 vs 2009-12.

*drovingnarrows0108*: Change in Temp. Def. Status 2001-05 vs 2008-12 (narrow definition of Temp. Def). 1=temporary deforestation in 2<sup>nd</sup> period, not in 1<sup>st</sup>. 0=temporary deforestation in both periods or in neither period. =1=temporary deforestation in 1<sup>st</sup> period but not 2<sup>nd</sup>. See Appendix A3 for further details on classification of sequences.

*dovingnarrow0107*: Change in Temp. Def. Status 2001-06 vs 2007-12 (narrow definition of Temp. Def).

*dovingnarrow0109*: Change in Temp. Def. Status 2001-04 vs 2009-12 (narrow definition of Temp. Def).

*dovingbroad0108*: Change in Temp. Def. Status 2001-05 vs 2008-12 (broad definition of Temp. Def).

*dstationarybroad0108*: Change in Perm. Def. Status 2001-05 vs 2008-12 (broad definition of Perm. Def)  
1="permanent" deforestation in 2<sup>nd</sup> period, not in 1<sup>st</sup>. 0="permanent" deforestation in both periods or in neither period. =1="permanent" deforestation in 1<sup>st</sup> period but not 2<sup>nd</sup>. See Appendix A3 for further details on classification of sequences.

*dstationarybroad0107*: Change in Perm. Def. Status 2001-06 vs 2007-12 (broad definition of Perm. Def)

*dstationarybroad0109*: Change in Perm. Def. Status 2001-04 vs 2009-12 (broad definition of Perm. Def)

*dstationarynarrow0108*: Change in Perm. Def. Status 2001-05 vs 2008-12 (narrow definition of Perm. Def)

## References

Brollo, F., Nannicini, T., Perotti, R. and Tabellini, G. (2013). 'The political resource curse', *American Economic Review*, vol. 103(5), pp. 1759–1796.

Channan, S. K., Collins, K., & Emanuel, W. R. (2014). Global mosaics of the standard MODIS land cover type data. *University of Maryland and the Pacific Northwest National Laboratory, College Park, Maryland, USA, 30*.

Cisneros, E., Zhou, S. L., & Börner, J. (2015). Naming and shaming for conservation: Evidence from the Brazilian Amazon. *PloS one*, 10(9), e0136402.

Fischer, G., Nachtergaele, F. O., Prieler, S., Teixeira, E., Tóth, G., Van Velthuisen, H., ... & Wiberg, D. (2012). Global Agro-ecological Zones (GAEZ v3. 0)-Model Documentation.