Supplemental Information (SI)

SI-1. Variable Definition Table

Variable Name	Definition	Range/Data Type	Source
propids	Unique identifier for each property (unit of analysis)	1-181 (nominal data)	SUDAM, "Perfil da ocupação do solo e subsolo Paraenses" (Profile of the occupation of the soil and subsurface claims of the State of Pará). Brasília, Brazil 1990.
year	Chronological year	1984-2010 (ordinal)	N/A
propsize	Size of the property, measured in hectares. UTM Zone 22S, WGS1984 Datum.	564-56035 (ratio data)	Geographic Information Systems calculation.
d2cities	Distance to urban centers, measured in kilometers	3.89-79.3 (ratio data)	Cities locations provided by the Brazilian Institute of Geography and Statistics (IBGE), Euclidian distance calculated with Geographic Information Systems.
lhcontrol	Binary variable indicating periods of largeholder control.	0-1 (nominal data)	Newspaper analysis (see SI-2)
numdalr	Number of conflict events (conflict of any kind).	0-56 (ratio data)	Newspaper analysis (see SI-2)
numdeaths	Number of deaths recorded due to land conflict.	0-24 (ratio data)	Newspaper analysis (see SI-2)
soilbin	Binary variable indicating if soils are suitable for smallholder agriculture (i.e., produce)	0-1 (nominal data)	Soil map provided by Brazilian Agricultural Research Corporation (EMBRAPA), recoded using EMBRAPA soil descriptions.
def1whectares	Deforested area measured in hectares, no regrowth of forests allowed.	0-55013 (ratio data)	Classified Landsat 5 scenes, land cover extracted to SUDAM map (see first row of this table).
settlementformed	Binary variable indicating if an agrarian reform settlement was formed.	0-1 (nominal data)	National Agrarian Reform Institute (INCRA) agrarian reform settlement boundaries overlain on SUDAM map (see first row of this table)
dalrever	Binary variable indicating if property ever experienced conflict.	0-1 (nominal data)	Recoding of numdalr variable.
balsamp	Binary variable indicating if observation should be excluded (exclude = 1) according to panel balancing (see methods discussion).	0-1 (nominal data)	See methods section of main paper.
precipannual	Amount of precipitation per year.	1347-2826 (ratio data)	NOAA/OAR/ESRL PSD, Boulder, Colorado, USA
period	Conflict epoch.	1-3 (ordinal data)	Generated by authors based on history of conflict, Newspaper analysis (see SI-2).
run1	Number of years since last land conflict event.	0-27 (ratio data)	Newspaper analysis (see SI-2)
ddefha	First-difference deforested area (def1whectares).	0-5481 (ratio data)	Derived from def1whectares variable.
dndalr	First-difference number of land conflicts (numdalr).	0-56 (ratio data)	Derived from numdalr variable.
dndeaths	First-difference number of deaths (numdeaths).	0-24 (ratio data)	Derived from numdeaths variable.
dlhcontrol	First-difference largeholder control (lhcontrol).	0-1 (ratio data)	Derived from lhcontrol variable.
dprecip	First-difference precipitation (precipannual).	0-1011 (ratio data)	Derived from precipannual variable.
dsettlementformed	First-difference settlement formed (settlement formed).	0-1 (ratio data)	Derived from settlementformed variable.

SI-2. Deforestation trajectory on property that was excluded from our analyses after the balancing procedure.

Average Total Deforested Area (Ha)
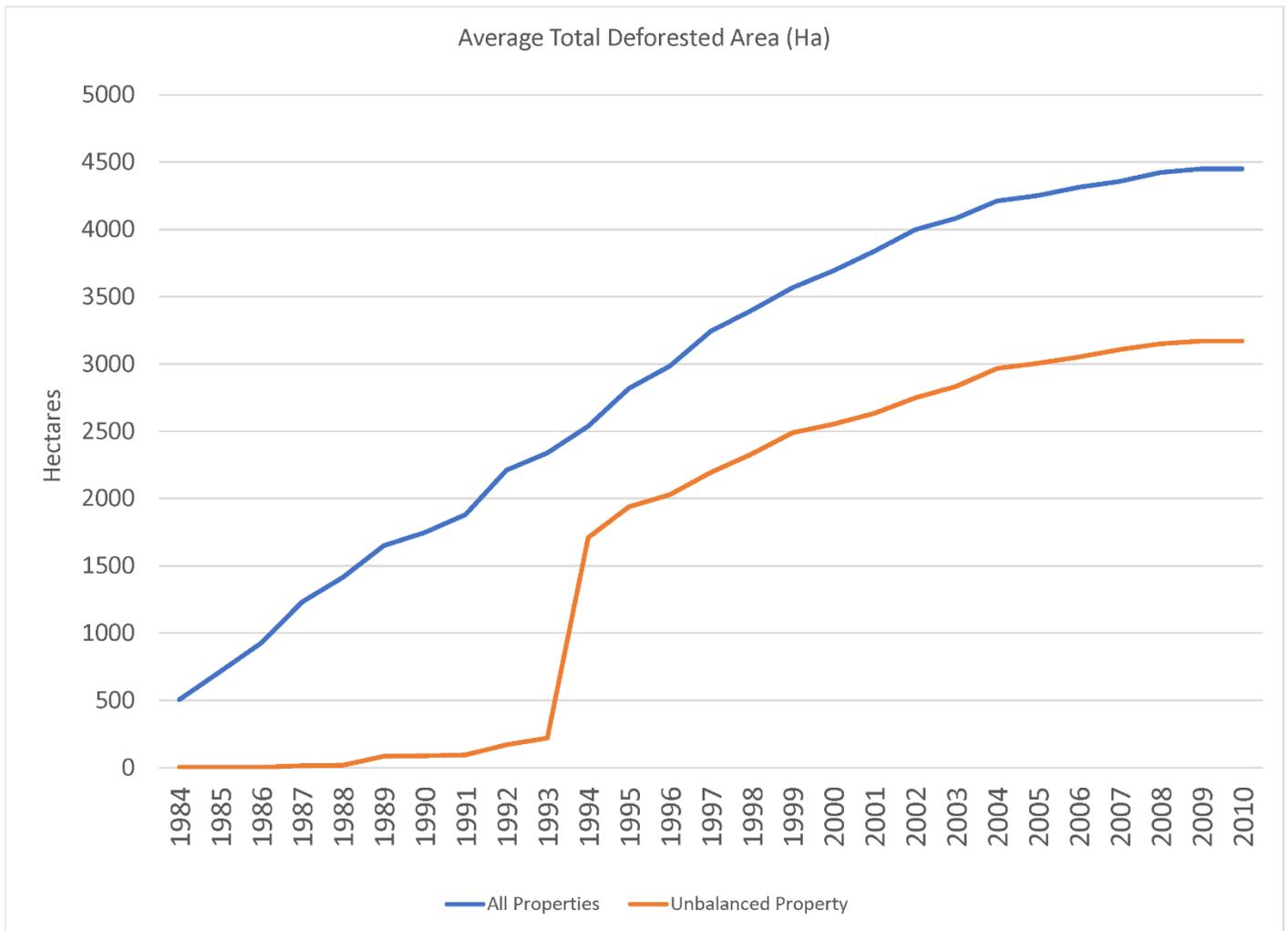
SI-3. Procedures for Newspaper Data

The data on conflict events were accumulated from newspaper accounts collected from the two newspapers that were consistently published in the region during the study period (1984-2010). Both newspapers, *O Correio do Tocantins* (now known as "*O Correio do Carajás*") and *Opinão!* were published in Marabá, with a minimum of two issues per week (in the 1980s) up to daily publication (starting in 2001). These newspapers report exclusively in Portuguese, and one (*O Correio do Tocantins*) was subjectively more critical of land reform on its editorial pages than the other (*O Opinião!)*, which was more populist in its coverage of land conflict and land reform and its editorial pages were more likely to support landless social movements (although was still critical of their actions, overall).

The newspaper pages were photographed by author Aldrich at the *Casa da Cultura* (Cultural House), an archive and museum maintained by the Municipal government of Marabá in 2006, in 2010, 2011, and 2014. Photographs were taken using natural light (i.e., no flash, by an open window) at six megapixels using a Pentax K100D and a 50mm manual-focus lens, with supervision by the archivist on staff. Only pages with discussion of land tenure, land conflict, agrarian violence, or land reform-related events were photographed.

After the entire set of newspaper pages were photographed (a total of 8,572 photographs), each article was coded by author Aldrich (over 6,000 pages coded by Aldrich) or one of two Brazilian graduate assistants (approximately 2,572 coded by graduate assistants) under the supervision of author Castro. The coding process commenced with a complete reading of the article, start to finish, at which point the article would be assessed as being "codable" -- articles which did not includes specific information, such as op-eds, reports on policy discussions, or cases where specific events were mentioned to support an argument (rather than documenting those specific events) were considered codable. Codable articles were then manually read for keywords which frequently had quantitative data attached to them (the "salient details" we mention in the article). The Portuguese words for "police," "death," "attack," "expulsion," "eviction," "ranch/farm," "municipality," "gunmen," "injury," and "complaint," are all such keywords, though it should be stated that we did not develop a formal list of these keywords.

For each codable article, a new row in a standardized spreadsheet was created, and data on the event described was recorded. The headers of that spreadsheet are (in order): Date of Publication, Date of Event, Type of Event (organically developed list of descriptors, not used in analysis), Gunmen Present, Police Present, Number of Deaths, Ranch Name, Municipality, Largeholder Name, Description (a one-to-two sentence paraphrasing of the event), Newspaper Photo Filename.

Once the entire dataset of conflict details was complete, we cross referenced conflict events between both newspapers (based on date, location, and description of the events), and only those events described in both newspapers were included in the analysis presented in this paper.

Linking conflict events to location involved a manual hierarchical matching process. First, ranch names were compared (e.g., the newspaper might mention a ranch named "Cabaceiras") and we would review the ranches in the Brazil Nut Polygon that may have that name (as indicated by the cadastral map, described in the manuscript). If there was a match based on ranch name, we would then review whether that ranch was in the same Municipality as indicated in the newspaper article (e.g., newspaper indicated the ranch "Cabaceiras" is in the municipality of Marabá). If these two data points matched we would consider the conflict event to have taken place in the location indicated on the cadastral map. In the Brazil Nut Polygon there are twelve cases where different properties have the same name, but only two cases where those same names are located in the same municipality. In these cases we then compared largeholder names between newspaper events and the cadastral maps to refine the match. In the limited cases where this information was not available, we discarded the event from our analysis as it could not be geolocated appropriately.

Dep. Variable:	Pooled OLS, First Difference Deforestation (Hectares), Detrended for Time	Fixed-Effects, First Difference Deforestation (Hectares)
Regression Characteristics	n = 4619, F [10, 4608]	n = 4619, F [7, 178]
	Prob > F = 0.0000	Prob > F = 0.0000
	R^2 = 0.3276	R^2 = 0.0661
Variable Name	Coefficient (t-value)	Coefficient (t-value)
Number of Conflicts (Lagged)	-0.639 (-0.26)	-41.800 (-1.76)*
Number of Deaths (Lagged)	3.054 (0.74)	2.975 (0.83)
Settlement Formed (Lagged)	47.356 (6.79)***	69.785 (2.16)**
Years Since Last Conflict	-0.831 (-1.95)*	-2.109 (-1.81)*
Annual Precipitation	-0.056 (-4.82)***	-0.051 (-5.20)***
Property Size (Hectares)	0.029 (1.02)	
Distance to Cities (Km)	0.217 (1.02)	
Soil Binary	-12.849 (-2.35)**	
Time Period 1992-2001	-44.167 (-4.39)***	-41.257 (-3.05)**
Time Period 2001-2010	-157.604 (-14.63)***	-152.987 (-8.58)***
Constant	170.799 (5.13)***	335.740 (14.83)***

Notes: Statistical significance indicated as follows: * = 0.10, ** = 0.05, *** = 0.000.

SI-4. Lagged models for Hypothesis 1. Note that lagging does not change the sign of the coefficient for conflict and agrarian reform settlement variables.

Group:	N	Mean	St.Err
Never Expropriated	43	3839.639	301.312
Expropriated	56	5669.159	986.013
Combined	99	4874.519	577.869
Difference		-1829.521	1157.011
t	-1.581		

H: Difference < 0	H: Difference <> 0	H: Difference > 0
Pr(T<t) = 0.0585	Pr(|T|>|t|) = 0.1171	Pr(T>t) = 0.9415

SI-5. t-test on deforestation totals (1984-2010, measured in hectares) on contentious properties between those which were expropriated for settlement formation and those that were not. Among properties with conflict those which are expropriated have significantly greater deforestation than those that were not. Note that 99 of the 180 properties had land conflict over the study period.

Dep. Variable:	Number of Conflict Events	Number of Conflict Events	Number of Conflict Events
Regression Characteristics	n = 4860	n = 4860	n = 4860
	Prob > F = 0.0405	Prob > F = 0.0406	Prob > F = 0.0333
	R^2 = 0.0007	R^2 = 0.0007	R^2 = 0.0.0010
Variable Name	Coefficient (t-value)	Coefficient (t-value)	Coefficient (t-value)
Year	0.001 (2.05)**		
Period (1 = 1984-1992, 2=1992-2001, 3=2001-2010)		0.809 (2.05)**	
Period 2 (1992-2001)			0.191 (2.42)**
Period 3 (2001-2010)			0.162 (2.05)**
Constant	-16.644 (-2.01)**	0.127 (1.49)	0.017 (3.06)**

Notes: Statistical significance indicated as follows: * = 0.10, ** = 0.05, *** = 0.000.

SI-6. Conflict and the passage of time are correlated (meaning that conflict is episodic, but concentrated in certain time periods). Related to this, deforestation and time are correlated, too.

Group:	N	Mean	St.Err
No Contention	81	3503.346	227.902
Contention	99	4305.141	462.301
Combined	180	3944.333	275.119
Difference		-801.795	551.294
t	-1.454		

H: Difference < 0	H: Difference <> 0	H: Difference > 0
	Pr(|T|>|t|) = 0.1476	
Pr(T<t) = 0.0738		Pr(T>t) = 0.9262

SI-7. t-test on deforestation totals (1984-2010, measured in hectares) between properties with conflict and those without. Properties without contention have significantly less deforestation than those that do, an outcome that supports H1.