

**Appendix to**  
**“Introducing the Lynching in Latin America (LYLA) Dataset”**

Enzo Nussio and Govinda Clayton

*Center for Security Studies, ETH Zürich*

**Contents**

1. Other data on lynching in Latin America
2. Lynching in Latin American law
3. Data collection procedures
4. Validation procedures
5. More detailed description of Application analysis
6. References

## 1. Other data on lynching in Latin America

In Latin America, several data collection efforts use different sources and varied means of operationalizing lynching. Here, we list some of the most important data collection efforts:

- In Brazil, two data collection efforts based on newspaper reports captured 1179 lynching events from 1980 to 2006 (NEV/USP 2007), and 2028 cases with 2579 victims spanning the period from 1945 to 1998 respectively (Martins 2015).
- In Guatemala, the United Nations Mission to the Guatemalan Peace Process collected evidence of 577 lynchings with 250 persons killed between 1996 and 2003 (MINUGUA 2000; see Mendoza 2008).
- In Mexico, researchers captured 1206 attempted and completed lynchings between 1988 and 2018, with 279 cases in 2018 alone (Rodríguez Guillén and Veloz Ávila 2019). State records in Mexico suggest that this might be a low estimate, as in the state of Puebla alone, a government agency recorded 305 lynching incidents between 2017 and 2018, with 42 people killed and 418 “rescued” (Puebla Hoy 2018). The report of the National Commission on Human Rights in Mexico further contains figures about lynching prevalence in four regions, drawing on a survey. Roughly 16% of the surveyed citizens were aware of a lynching in their locality in the previous 12 months (CNDH 2019, 71).
- The 2017 Venezuelan Violence Observatory annual report registered 2.4 people killed in lynchings per week, thus more than 120 lethal victims of lynching (Observatorio Venezolano de Violencia 2018).
- In Bolivia, Luna Acevedo (2016) identified 199 lynching events with 373 victims from 2005 to 2011 based on local newspaper reports.
- In Colombia, researchers found evidence of 102 lynchings only in the capital Bogotá in the month of August 2014 (Ariza 2019). This figure is based on the availability of detailed police reports about citizen arrests. The extremely large number of lynching incidents for a short period of time and exclusively focusing on Bogotá suggests that Ariza and his team pick up a lot of small incidents that are not covered in newspapers.<sup>1</sup>

---

<sup>1</sup> Ariza and his team study police reports about legal citizen arrests from the mayor’s office in Bogotá for the month of August 2014. This was the only month that contained sufficient qualitative description of events to classify them as potential lynchings. According to their coding criteria, 102 cases (of a total of 1236 citizen arrest cases) constituted lynchings in the month of August 2014. Unfortunately, this kind of information is not available on a wider scale and does not allow for systematic comparison across time and space.

- Other studies report numbers on Ecuador (Guerrero 2000; Santillán 2008), Peru (Castillo Claudett 2000), Argentina (Gamallo 2020; González, Ladeuix, and Ferreyra 2011) and Bolivia (Vilas 2008).

Some of these data sources are used in our “Country reports” which are described below and used for validation of the LYLA dataset.

## 2. Lynching in Latin American law

In the course of this research, we reviewed the existing legal codes in Latin America and did not find a typified crime corresponding specifically to lynching (see also CNDH 2019, 188). However, legal codes typify related practices (article numbers refer to the respective penal codes of each country).

The most conspicuous example is the Mexican constitution, which states: “No person can make justice on their own or use violence to claim a right” (art. 17). The related article 131 of the penal code also specifies the crime of a “motín” (akin to mutiny). Also, as a federal state, Mexico has additional legislation on the state level. In the state of Hidalgo, for example, a police protocol was officially adopted in 2019 to attend lynchings (Periódico Oficial del Estado de Hidalgo 2019). Furthermore, congressman José Porfirio Alarcón Hernández proposed to change article 321 of the Mexican penal code in line with what he described as lynching (Diario de los Debates de la Cámara de Diputados 2004). This initiative, which did not pass, was in response to the famous 2004 Tláhuac lynching of three policemen.

The Peruvian Penal code recognizes a crime of arbitrary justice administration, which specifically prohibits making self-justice (art. 417).

The Venezuelan penal code prohibits self-justice, but the punishment depends on the types of violence inflicted, for example homicide or injury (art. 271).

Guatemala (art. 39) and Uruguay (art. 65) specify a crime of “muchedumbre”, which involves the participation in a tumultuous assembly including the commission of crimes, which could amount to a lynching. In both cases, all the material participants of such an assembly are legally liable while the others are exempt from punishment.

A similar crime is specified in the Paraguayan penal code, which calls it disturbance of public peace (art. 234).

Several states typify in their penal codes injuries and homicides resulting from fights (“riñas”), for example Argentina (art. 35), Bolivia (art. 259), Costa Rica (art. 139), Ecuador (art. 470), Honduras (art. 119 and 137), and Nicaragua (art. 158).

The Penal Code of the Dominican Republic furthermore specifies a crime of “barbarism” which involves torture and may be related to lynching (art. 303).

In Brazil, there are mitigating circumstances for injuries and homicides perpetrated due to a relevant “social or moral value” (art. 65), which may be related to lynching.

The Colombian code has a wide-ranging specification of legitimate self-defense whereby the defense has to be proportional to the aggression (art. 32.6), in contrast to the more common legal prescription that the defense has to be proportional to the means necessary to defend

oneself, regardless of the type of aggression committed. This may be relevant for lynching events.

In Bolivia, there was a proposal to specifically include the crime of lynching in the penal code in 2013, which was rejected (Opinión Bolivia 2013). The proposing lawyer argued that the common practice of prosecuting lynchings with the crime of homicide was insufficient.

### 3. Data collection procedures

#### 3.1. Search string

Within the Factiva news archive, we restricted our search to articles that could potentially describe lynching events. Hence, we used relevant search strings. These search strings were largely identical across countries but included additional country-specific criteria to reduce noise. This way, we could produce a manageable amount of newspaper articles that our research assistants would go through and code according to a pre-established codebook. This is the basic search string we used on the Factiva website:

```
( (Lynch*) or (Linch*) or (mano near5 propia) or (popular* near5 justic*) or (quema* near5 vivo*) or (atad* near5 poste*) or (Turba) or (tumulto) or (Mob) or (Stoning) or (Immolating) or (Apedrea*) or (Lapida*) or (vigilant* same justic* or kill* or attack*)) or (Hanging near100 (dead or death or kill* or body)) ) not (Merrill Lynch) not (Merrill Lynch) not (Meril Lynch) not (Merril Lynch) not (Merill Linch) not (Merrill Linch) not (Meril Linch) not (Merril Linch) not (Meryll Lynch) not (Meryll Lynch) not (Meryl Lynch) not (Merryl Lynch) not (Meryll Linch) not (Merryl Linch) not (Meryl Linch) not (Merryl Linch) not (Larry Lynch) not (David Lynch) not (James Lynch) not (Michael Lynch) not (Peter Lynch) not (Gabriel Lynch) not (Titulares de los diarios latinoamericanos) not (in=I814) not (fds=PEMEKS) not (fds=BCMEKS)
```

#### 3.2. News articles covered for each country and year

The coding of lynching events depends on the source one uses. To a certain extent, the identification of lynching events is a function of available news sources. This [document](#) presents the total amount of articles for each country contained in Factiva and the amount of articles that have a chance to contain information about lynching events (results generated with our search string). Varying levels of newspaper coverage depend mainly on the size and international importance of a given country. Mexico, Brazil, and Argentina are well covered while Central American countries are least well covered. Researchers can find information on a set of variables for each country and year from 2010 to 2019 (for Mexico, the years 2000-2009 and 2020-2021 are also covered).

The first variable (n\_doc\_ly) corresponds to the articles that research assistants read and coded (roughly 80,000 for the period between 2010-2019). These variables can be introduced for example into country-year analysis to adjust for coverage variation across time and space.

#### 3.3. Coding procedures

Events were coded with two versions of the codebook. First, Mexico, Guatemala, Brazil, and Colombia were coded with an extended version, covering a large amount of variables. Second, all other countries were coded with a shorter codebook covering less questions. Coverage for each variable is clearly stated in the separate Codebook. The most important variables covered are the date and coordinate of a given lynching event. The selection of news reports and review of articles by coders was the same for both types of countries.

As general guideline, we do not blindly follow language used by journalists but code events as lynchings depending on our definitional criteria, including the presence of (1) a group of civilians, (2) using violence, (3) against an alleged wrongdoer, (4) in a public display. The boundary condition for the violence used is a clear threat of lynching violence. Therefore, events that some may consider “attempted” lynching are also included in our dataset. Researchers who use the LYLA data are free to set a more demanding criterion for inclusion (for example only focusing on cases with a resulting injury or death).

#### **4. Validation procedures**

In this Appendix, we examine the validity of the Lynching in Latin America (LYLA) dataset. The phenomenon we intend to capture with our dataset are lynching events in Latin America, defined as “publicly displayed physical violence executed by a group of civilians against alleged wrongdoers”. Given that there is no readily available systematic information about this phenomenon, we decided to collect data ourselves using the Factiva news sources repository.

##### **4.1. Note on potential biases for future users of the LYLA data**

The real amount of lynching events in Latin America is unknown. Our data can thus not be considered a full representation of actual lynchings. They reveal a low estimate of actual lynching events, given that many cases are not reported. This is why we consistently speak of “reported lynchings”. Also, our data may represent a low estimate as we deliberately set a bar for classifying an event as lynching, perhaps higher than other analysts and journalists.

For most analysis focusing on relationships between lynching and other variables, underreporting is not the main problem, but systematic bias across units. Our data on *reported lynchings* covers the tip of the iceberg of the underlying phenomenon of *actual lynchings*. To draw valid inferences from this data, the units that we compare (for example years and geographical units) should have the same relationship between reported and actual lynchings (i.e. between the tip of the iceberg and the hidden part of the iceberg).

This assumption is not always satisfied. For example, there can be systematic differences in the relationship between actual and reported lynchings if the sources of information change from one year to the next (see description of Factiva source material above). Also, systematic differences can arise from differential news coverage across space (we have estimates of sources across countries). Cities, for example, have more journalists than rural areas, which may lead to an urban bias in our data. These sources of bias need to be addressed using appropriate statistical tools.

##### **4.2. Validation using external data sources, Factiva coverage, and qualitative studies**

The validation of the LYLA data proceeds in three steps. First, we present external data on lynching in Latin America created by other researchers and institutions to assess how the LYLA data compares to these other sources. From these comparisons, we can make an informed guess about the validity of our data. Given that most additional data sources are based on national datasets, we undertake this data validation process for each country. For some countries, we can compare our data to several additional data sources. If appropriate data is



available, we present timelines and tables that compare our data to other data sources. While comparisons across different datasets sound relatively straightforward, we need to be aware of different lynching definitions and operationalizations. Generally speaking, we compare yearly numbers of lynching events (or lynching victims, depending on availability) and numbers of lynching events across provinces. With this procedure, we can assess the two crucial dimensions of variation for our analysis: space and time.

Second, we assess the overall amount of newspaper evidence on which our data is based, using the total amount of Factiva articles covering each country. This is important as encountering a lynching event is, in part, a function of the number of articles reviewed. With this procedure, we thus gain additional measures of lynching prevalence on the national and yearly scale and can compare these measures across countries.

Third, we qualitatively review relevant literature to identify whether we capture similar underlying phenomena. This is especially important for countries for which there is no additional data source.

As a result of this process, we compiled a country report for all countries with relevant additional data sources (we collapsed some of the countries with no additional information). These reports can be accessed [here](#).

### **4.3. Survey-based validation in Mexico City**

One alternative to capture the prevalence of lynching is to tap into the local knowledge of residents. We therefore fielded a representative survey with 2183 adult Mexico City residents in February 2022, in partnership with a Mexican opinion survey company. We employed multi-stage sampling, first selecting 340 colonias (Mexico City has a total of 1800 colonias or neighborhoods) with probability proportional to size sampling (Skinner 2016), and then randomly selecting six or more households within each colonia.

The questionnaire contained a series of questions including one about local knowledge of a lynching-style incident. After presenting respondents with a vignette about the modal type of lynching (a male thief being punished by a group of bystanders), we asked them: “*Do you know or have you heard of such a type of event in your colonia, meaning neighbors punishing a criminal?*” Overall, 30.7% of the respondents in the Mexico City sample responded affirmatively.

We can use the average affirmative response to this question in each colonia to further validate our lynching event dataset. We correlate the average colonia-level response with different specifications of colonia-level lynching events drawn from the LYLA data: (1) a

binary indicator of whether there was a reported lynching or not in the whole period (2000-2022 February: 24% of colonias had a lynching), (2) the number of reported lynchings per colonia over the whole period (mean: 0.4), (3) whether there was a reported lynching from 2017 to February 2022 (16% of colonias), (4) whether there was a reported lynching from 2019 to February 2022 (12% of colonias), (5) lynchings per million inhabitants (mean: 30), and (6) the natural log of lynchings per million inhabitants.

Table A1 shows the respective correlation coefficients for the 340 covered colonias. While the coefficients are not large, there is a systematic correlation between the survey reported lynching measure and our lynching event data based on newspaper reports.

*Table A1. Correlation of LYLA event counts and survey responses (N=340 colonias)*

Variables	Surveyees who know
Lynchings yes/no	0.214***
Number of lynchings	0.174***
Lynchings after 2016	0.114**
Lynchings after 2018	0.112**
Lynchings per mio	0.172***
Log Lynchings per mio	0.210***

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

When interpreting these figures, readers should be aware of two important challenges for a validation process based on local survey measures: First, to tap into local knowledge, the survey question needs to refer to a small unit – in our case the colonia or neighborhoods. Using larger units – such as municipalities or states – is a less promising strategy as the sampling process may happen to produce a sample from a lynching prone-neighborhood in one unit and a less lynching-prone neighborhood in another unit. Requirements for sample sizes within units would therefore be very demanding for larger units with larger within-unit variation. Also, knowledge of what happens at the municipality or state level may be less accurate.

Second, without incurring significant costs, it is impossible to capture both a large number of units and a large number of individuals within each unit. We struck a balance at 340 colonias within Mexico City, which is an unusually large coverage for an opinion survey, and at least 6 respondents in each colonia.

Third, at the outset, it is unclear how accurate local knowledge about lynching events effectively is. If local knowledge is highly accurate across individuals, a small sample would suffice to capture the actual occurrence of lynching events in a given unit. However, within-unit variation is large in our case. Hence, the estimates we recover for each colonia are affected by random variation.

Despite these limitations, Table A1 shows that our lynching event data is systematically correlated with residents' knowledge of lynching on the Mexico City colonia level.

## 5. More detailed description of Application analysis

### 5.1. Variables and their sources

The following variables are used in the province level analysis:

- *State legitimacy indicators (trust in government, trust in police, trust in justice, Courts are fair)*: Measured on 1 to 7 Likert scale. Each individual province mean is calculated based on all available individuals living in that province. For trust in police, all provinces are based on at least 22 individuals, 95% of the provinces are based on 50 observations, and 91% on 100 observations. Source: <https://www.vanderbilt.edu/lapop/about-americasbarometer.php>
- *Area in km2*: Area of the province in square kilometers, calculated from polygon information.
- *Population size*: Mean population number from 2000-2019. Source: Instituto Igarapé. <https://homicide.igarape.org.br/>
- *Distance to capital*: Calculated as the distance from the province capital to the country capital, as a great circle distance.
- *Homicide rate*: Rate of homicides per 100,000 inhabitants. Source: Instituto Igarapé. <https://homicide.igarape.org.br/>
- *Road density*: Calculated as kilometers of roads divided by area in km2. Source for road kilometers: Center for International Earth Science Information Network - CIESIN - Columbia University, and Information Technology Outreach Services - ITOS - University of Georgia. <https://sedac.ciesin.columbia.edu/data/set/groads-global-roads-open-access-v2>
- *Owning a car*: Average response to the LAPOP question about household car owning. Source: Latin American Public Opinion Project: <https://www.vanderbilt.edu/lapop/about-americasbarometer.php>
- *Urbanity*: Calculated from coverage of urban areas in each PRIO grid cell. Original source: Globcover 2009. Accessed via: <https://grid.prio.org/#/download>

### 5.2. Additional analysis

Table A2. Full table output for Figure 7 (fixed effects model)

	(1)	(2)	(3)	(4)	(5)	(6)
Legitimacy index	-0.32*** (0.07)					
First component of legitimacy items		-0.31*** (0.07)				
Trust in government			-0.18* (0.08)			
Trust in police				-0.41*** (0.08)		

Trust in justice					-0.19**	
					(0.07)	
Courts are fair						-0.26***
						(0.06)
Population size	0.00	0.00	0.00	0.00	0.00*	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	0.89***	0.91***	0.98***	0.88***	0.97***	0.94***
	(0.18)	(0.18)	(0.19)	(0.18)	(0.18)	(0.18)
<i>N</i>	338	338	338	349	348	349
adj. <i>R</i> <sup>2</sup>	0.47	0.47	0.45	0.44	0.42	0.43

Standard errors in parentheses

OLS models with country fixed effects and adjusting for population size. *N* varies due to non-response.

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

*Table A3. Full table output for Figure 7 (control variables model)*

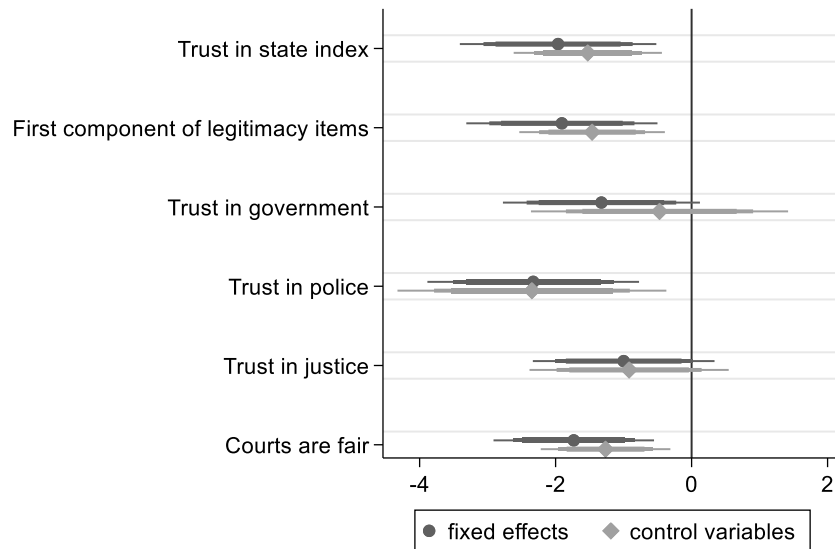
	(1)	(2)	(3)	(4)	(5)	(6)
Legitimacy index	-0.29**					
	(0.09)					
First component of legitimacy items		-0.29**				
		(0.09)				
Trust in government			-0.13			
			(0.13)			
Trust in police				-0.38**		
				(0.11)		
Trust in justice					-0.21	
					(0.10)	
Courts are fair						-0.24**
						(0.08)
Population size	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Area km <sup>2</sup>	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Distance to Capital	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Homicide rate	-0.00	-0.00	-0.00	-0.01	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Road density	2.67	2.79	2.37	2.50	2.85	3.02
	(1.74)	(1.74)	(1.87)	(1.56)	(1.78)	(1.76)
Owning a car	0.67	0.67	0.78	0.24	0.63	0.47
	(0.65)	(0.65)	(0.70)	(0.65)	(0.62)	(0.65)
Urbanity	0.18	0.18	0.22	0.15	0.20	0.18
	(0.13)	(0.13)	(0.12)	(0.11)	(0.13)	(0.14)
Constant	0.51	0.49	0.52	0.71*	0.50	0.49
	(0.27)	(0.27)	(0.31)	(0.32)	(0.30)	(0.28)
<i>N</i>	338	338	338	349	348	349
adj. <i>R</i> <sup>2</sup>	0.18	0.17	0.12	0.22	0.14	0.15

Standard errors in parentheses

OLS models without clustered standard errors on country level. *N* varies due to non-response.

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

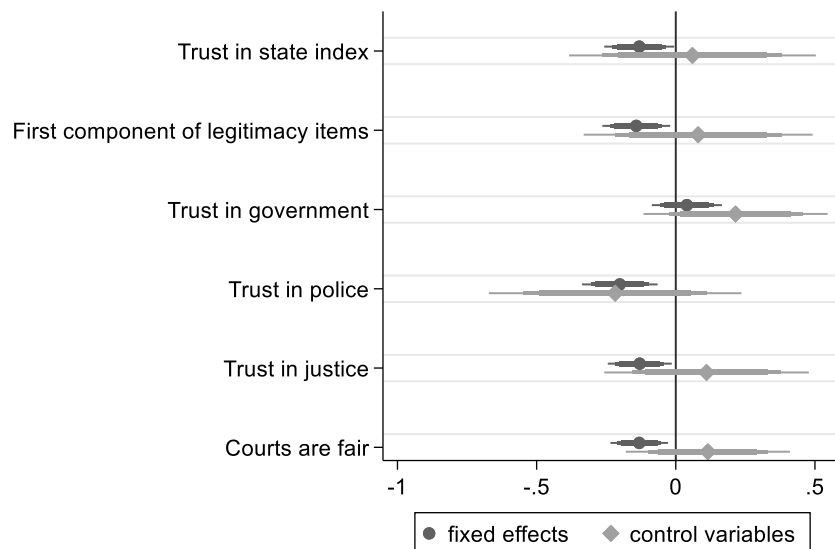
Figure A1. State legitimacy and lynching per million inhabitants



Note: OLS regression coefficients with 99, 95 and 90% confidence intervals. Fixed effects model adjusts for population size and country fixed effects. Control variables model includes series of control variables and clustered standard errors. N = 338 to 349 for all models.

Figure A1 displays the same model as Figure 7 in the main paper without logging the dependent variable.

Figure A2. State legitimacy and homicide rate per 100,000 inhabitants (log)



Note: OLS regression coefficients with 99, 95 and 90% confidence intervals. Fixed effects model adjusts for population size and country fixed effects. Control variables model includes series of control variables and clustered standard errors. N = 338 to 349 for all models.

Figure A2 shows that indicators of state legitimacy are not robustly related to homicide rates. Results are similar for non-logged homicide rates.

## 6. References

- Ariza, Rosemberth. 2019. "Linchamientos En Bogotá: ¿Violencia Urbana Legítima o Consolidación de Prácticas de Odio Social?" *Análisis Político* 32 (96): 83–102.
- Castillo Claudett, Eduardo. 2000. "La justicia en tiempos de la ira: linchamientos populares urbanos en América Latina." *Ecuador Debate* 51: 207–26.
- CNDH. 2019. "Informe Especial Sobre Los Linchamientos En México." Mexico: Comisión Nacional de Derechos Humanos.
- Diario de los Debates de la Cámara de Diputados. 2004. "Iniciativa Que Reforma El Artículo 321 Del Código Penal Federal." Mexico City: Cámara de Diputados.
- Gamallo, Leandro. 2020. "Collective Violence and Politics in Argentina." *New Global Studies* 14 (2): 157–64.
- González, Leandro, Juan Iván Ladeuix, and Gabriela Ferreyra. 2011. "Acciones colectivas de violencia punitiva en la Argentina reciente." *Bajo el Volcán* 10 (16): 165–93.
- Guerrero, Andrés. 2000. "Los linchamientos en las comunidades indígenas (Ecuador): ¿la política perversa de una modernidad marginal?" *Bulletin de l'Institut français d'études andines* 29 (3): 463–89.
- Luna Acevedo, Héctor. 2016. "Los Actos de Linchamiento y La Inseguridad Ciudadana En Bolivia." *Temas Sociales*, no. 38: 155–80.
- Martins, José de Souza. 2015. *Linchamentos: a justiça popular no Brasil*. São Paulo: Editora Contexto.
- Mendoza, Carlos. 2008. "Linchamientos en México y Guatemala: reflexiones para su análisis comparado." *El Cotidiano*, no. 152: 43–51.
- MINUGUA. 2000. *Los linchamientos: un flagelo contra la dignidad humana*. Guatemala City: MINUGUA.
- NEV/USP. 2007. "Banco de Datos Da Imprensa - Número de Casos de Linchamento." Universidade de São Paulo. <https://nev.prp.usp.br/dados/banco-de-dados-da-imprensa-sobre-as-graves-violacoes-de-direitos-humanos-dados-por-violacao-linchamento/>.
- Observatorio Venezolano de Violencia. 2018. "2017: Informe Anual de Violencia." Caracas: OVV. <https://observatoriodeviolencia.org.ve/news/2017-informe-ovv-de-violencia/>.
- Opinión Bolivia. 2013. "Proponen que linchamiento sea delito en Código Penal." October 12, 2013. <https://www.opinion.com.bo/articulo/informe-especial/proponen-linchamiento-sea-delito-codigo-penal/20131012234200664373.html>.
- Periódico Oficial del Estado de Hidalgo. 2019. "Protocolo de Actuación Policial Para El Control de Multitudes Ante El Riesgo de Violencia Colectiva." Pachuca: Gobernación de Hidalgo.
- Puebla Hoy. 2018. "Reporta SGG Más de 300 Intentos de Linchamiento," December 4, 2018. <https://pueblahoy.net/reporta-sgg-mas-de-300-intentos-de-linchamiento/>.
- Rodríguez Guillén, Raúl, and Norma Ilse Veloz Ávila. 2019. "Linchamientos En México: Una Puesta al Día." *El Cotidiano* 34 (214): 87–94.
- Santillán, Alfredo. 2008. "Linchamientos urbanos. 'Ajusticiamiento popular' en tiempos de la seguridad ciudadana." *Íconos* 12 (2): 57–69.
- Skinner, Chris J. 2016. "Probability Proportional to Size (PPS) Sampling." In *Wiley StatsRef*, 1–5. New York: John Wiley & Sons, Ltd.
- Vilas, Carlos M. 2008. "Lynchings and Political Conflict in the Andes." *Latin American Perspectives* 35 (5): 103–18.