Supplementary Material

"Personalization of Power and Mass Uprisings in Dictatorships"

Contents

Appendix A: Data

List of nonviolent protest campaigns and autocratic regime cases

Country	*Year*	*Never*	*Navco*	*MEC*	*Country*	*Year*	*Never*	*Navco*	*MEC*	*Country*	*Year*	*Never*	*Navco*	*MEC*	*Country*	*Year*	*Never*	*Navco*	*MEC*
Albania	1989	0	1	0	Kyrgyzstan	2010	1		1	Ecuador	1966	1	0	0	Senegal	1982	0	0	1
Albania	1990	1	0	1	Madagascar	1972	1	0	1	Egypt	2000	1	1	1	Senegal	1983	1	0	0
Algeria	1992	0	1	0	Madagascar	1991	1	1	1	Egypt	2007	1		1	Senegal	2000	1	1	1
Argentina	1977	1	1	1	Malawi	1992	1	1	1	El Salvador	1967	1	0	0	Serbia	1996	1	1	1
Armenia	1996	1	0	0	Malaysia	1998	1	0	1	El Salvador	1977	1	1	1	Serbia	2000	1	0	0
Armenia	2003	1	0	0	Malaysia	2007	1		0	Ethiopia	2005	1	0	0	S. Africa	1952	1	1	
Armenia	2007	1		1	Mali	1990	1	1	1	Gabon	1990	1	0	0	S. Africa	1984	1	0	1
Azerbaijan	1992	1	0	0	Mexico	1968	1	0	0	Gabon	1993	1	0	0	S. Vietnam	1964	1	0	1
Azerbaijan	2000	1	0	0	Mexico	1987	1	1	1	Georgia	2003	1	1	1	S. Vietnam	1966	1	0	1
Azerbaijan	2002	1	0	0	Mongolia	1989	1	1	1	Germany East	1956	0	1	1	Soviet Union	1987	1	0	1
Azerbaijan	2005	1	0	0	Morocco	1999	1	0	1	Germany East	1989	1	1	1	Soviet Union	1988	1	0	1
Bangladesh	1987	1	1	1	Morocco	2010	1		1	Ghana	2000	1	1	1	Soviet Union	1989	1	0	1
Belarus	1996	1	0	0	Myanmar	1988	1	1	1	Greece	1973	1	1	1	Soviet Union	1990	1	1	1
Belarus	1999	1	0	0	Myanmar	2007	1		1	Guatemala	1993	1	0	0	Soviet Union	1991	1	0	0
Belarus	2004	1	0	0	Nepal	1990	1	1	1	Guinea	2007	1		1	Spain	1975	1	0	0
Belarus	2006	1	1	1	Nepal	2006	0	1	0	Haiti	1946	1	0		Spain	1976	1	0	0
Belarus	2010	1		0	Nicaragua	1978	0	0	1	Haiti	1956	1	0	0	Sudan	1964	1	0	0
Benin	1963	1	0	0	Niger	1990	1	0	1	Haiti	1985	1	1	1	Sudan	1985	1	1	1
Benin	1989	1	1	1	Niger	1991	0	1	0	Haiti	1986	1	0	0	Taiwan	1979	1	1	1
Bolivia	1977	0	1	1	Nigeria	1990	1	1	1	Haiti	1987	0	0	1	Taiwan	1985	1	0	0
Bolivia	1978	1	0	0	Nigeria	1993	1	1	1	Haiti	1990	1	0	0	Taiwan	1990	1	0	0
Brazil	1968	1	0	0	Pakistan	1968	1	0	1	Haiti	2003	1	0	1	Tanzania	1992	1	1	1
Brazil	1977	1	0	0	Pakistan	1969	1	0	0	Hungary	1956	1	1	1	Thailand	1973	1	1	1
Brazil	1983	1	0	0	Pakistan	1977	1	0	1	Hungary	1989	1	1	1	Thailand	1992	1	1	1
Brazil	1984	0	1	1	Pakistan	1983	1	1	1	Indonesia	1956	0	1	0	Thailand	2007	1		1
Bulgaria	1989	1	1	1	Pakistan	1986	1	0	1	Indonesia	1974	0	1	0	Togo	1991	1	0	1
Burkina Faso	1966	1	0	0	Pakistan	2007	1		1	Indonesia	1979	0	0	1	Togo	2005	1	0	1
Cambodia	1998	1	0	1	Panama	1951	1	0		Indonesia	1989	1	0	1	Tunisia	2010	1		1
Cameroon	1991	1	0	0	Panama	1987	1	1	1	Indonesia	1997	1	1	1	Uruguay	1983	1	0	0
Cameroon	2008	1		0	Paraguay	1986	1	0	0	Indonesia	1999	1	0	1	Uruguay	1984	0	1	1
Chile	1983	1	1	1	Peru	2000	1	1	1	Iran	1977	1	1	1	Venezuela	1957	1	0	0
China	1956	0	1	1	Philippines	1983	1	1	1	Iran	1999	1	0	1	Venezuela	1958	0	1	1
China	1957	1	0	0	Poland	1956	1	1	1	Iran	2009	1		1	Yemen	2007	1		1
China	1976	1	1	1	Poland	1968	1	1	1	Kenya	1990	1	1	1	Yugoslavia	1968	1	1	1
China	1987	1	1	1	Poland	1970	1	1	1	Kenya	1997	1	0	0	Yugoslavia	1970	0	1	1
China	1989	1	1	1	Poland	1976	0	1	1	Korea S.	1960	1	1	1	Yugoslavia	1971	1	0	0
China	2008	1		0	Poland	1980	1	1	1	Korea S.	1979	1	1	1	Yugoslavia	1981	1	1	1
Colombia	1957	1	0	0	Poland	1988	1	0	0	Korea S.	1986	0	0	1	Yugoslavia	1988	1	0	1
Congo-Brz	1990	1	0	0	Portugal	1973	0	1	1	Korea S.	1987	1	1	0	Yugoslavia	1989	0	1	1
Czechoslovakia	1968	1	1	1	Portugal	1974	1	0	0	Kuwait	1989	1	0	0	Yugoslavia	1990	1	0	1
Czechoslovakia	1988	1	0	0	Romania	1987	1	1	1	Kuwait	2006	1	0	0	Zambia	1990	1	1	1
Czechoslovakia	1989	0	1	1	Russia	2007	1		1	Kyrgyzstan	2000	1	0	0	Zambia	2001	1	1	1
Dominican Rep	1961	1	0	0	Russia	2010	1		1	Kyrgyzstan	2005	1	1	1	Zimbabwe	2003	1	0	0

RE: NEVER, Chin (2017); Navco: Navco 2.0 Chenoweth and Lewis (2013); MEC: Mass Episodes of Contention, Chenoweth and Ulfelder (2017).

Figure A-1: Nonviolent protest campaign starts in dictatorships, 1946–2010

Autocratic regimes in test sample

Regime-case	*Begin*	*End*	*Regime-case*	*Begin*	*End*	*Regime-case*	*Begin*	*End*	*Regime-case*	*Begin*	*End*
Afghanistan 29-73	1946	1973	Congo-Brz 60-63	1961	1963	Ivory Coast 99-00	2000	2000	Poland 44-89	1946	1989
Afghanistan 73-78	1974	1978	Congo-Brz 63-68	1964	1968	Ivory Coast 00-NA	2001	2010	Portugal 26-74	1946	1974
Afghanistan 78-92	1979	1992	Congo-Brz 68-91	1969	1991	Jordan 46-NA	1947	2010	Romania 45-89	1946	1989
Afghanistan 96-01	1997	2001	Congo-Brz 97-NA	1998	2010	Kazakhstan 91-NA	1992	2010	Russia 93-NA	1994	2010
Afghanistan 09-NA	2010	2010	Congo/Zaire 60-97	1961	1997	Kenya 63-02	1964	2002	Rwanda 62-73	1963	1973
Albania 44-91	1946	1991	Congo/Zaire 97-NA	1998	2010	Korea North 48-NA	1949	2010	Rwanda 73-94	1974	1994
Algeria 62-92	1963	1992	Costa Rica 1948-49	1949	1949	Korea South 48-60	1949	1960	Rwanda 94-NA	1995	2010
Algeria 92-NA	1993	2010	Cuba 52-59	1953	1959	Korea South 61-87	1962	1987	Saudi Arabia 27-NA	1946	2010
Angola 75-NA	1976	2010	Cuba 59-NA	1960	2010	Kuwait 61-NA	1962	2010	Senegal 60-00	1961	2000
Argentina 43-46	1946	1946	Czechoslovakia 48-89	1949	1989	Kyrgyzstan 91-05	1992	2005	Serbia/Yugoslavia 91-00	1992	2000
Argentina 51-55	1952	1955	Domincan Rep 30-62	1946	1962	Kyrgyzstan 05-10	2006	2010	Sierra Leone 67-68	1968	1968
Argentina 55-58	1956	1958	Dominican Rep 63-65	1964	1965	Laos 59-60	1960	1960	Sierra Leone 68-92	1969	1992
Argentina 58-66	1959	1966	Dominican Rep 66-78	1967	1978	Laos 60-62	1961	1962	Sierra Leone 92-96	1993	1996
Argentina 66-73	1967	1973	Ecuador 44-47	1946	1947	Laos 75-NA	1976	2010	Sierra Leone 97-98	1998	1998
Argentina 76-83	1977	1983	Ecuador 63-66	1964	1966	Lesotho 70-86	1971	1986	Somalia 69-91	1970	1991
Armenia 94-98	1995	1998	Ecuador 70-72	1971	1972	Lesotho 86-93	1987	1993	South Africa 10-94	1946	1994
Armenia 98-NA	1999	2010	Ecuador 72-79	1973	1979	Liberia 44-80	1946	1980	South Vietnam 54-63	1955	1963
Azerbaijan 91-92	1992	1992	Egypt 22-52	1946	1952	Liberia 80-90	1981	1990	South Vietnam 63-75	1964	1975
Azerbaijan 93-NA	1994	2010	Egypt 52-NA	1953	2010	Liberia 97-03	1998	2003	South Yemen 67-90	1968	1990
Bangladesh 71-75	1972	1975	El Salvador 31-48	1946	1948	Libya 51-69	1952	1969	Soviet Union 17-91	1946	1991
Bangladesh 75-82	1976	1982	El Salvador 48-82	1949	1982	Libya 69-NA	1970	2010	Spain 39-76	1946	1976
Bangladesh 82-90	1983	1990	El Salvador 82-94	1983	1994	Madagascar 60-72	1961	1972	Sri Lanka 78-94	1979	1994
Bangladesh 07-08	2008	2008	Eritrea 93-NA	1994	2010	Madagascar 72-75	1973	1975	Sudan 58-64	1959	1964
Belarus 91-94	1992	1994	Ethiopia 1889-1974	1946	1974	Madagascar 75-93	1976	1993	Sudan 69-85	1970	1985
Belarus 94-NA	1995	2010	Ethiopia 74-91	1975	1991	Madagascar 09-	2010	2010	Sudan 85-86	1986	1986
Benin 60-63	1961	1963	Ethiopia 91-NA	1992	2010	Malawi 64-94	1965	1994	Sudan 89-NA	1990	2010
Benin 63-65	1964	1965	Gabon 60-NA	1961	2010	Malaysia 57-NA	1958	2010	Swaziland 68-NA	1969	2010
Benin 65-67	1966	1967	Gambia 65-94	1966	1994	Mali 60-68	1961	1968	Syria 46-47	1947	1947
Benin 67-69	1968	1969	Gambia 94-NA	1995	2010	Mali 68-91	1969	1991	Syria 49-51	1950	1951
Benin 69-70	1970	1970	Georgia 91-92	1992	1992	Mauritania 60-78	1961	1978	Syria 51-54	1952	1954
Benin 72-90	1973	1990	Georgia 92-03	1993	2003	Mauritania 78-05	1979	2005	Syria 57-58	1958	1958
Bolivia 43-46	1946	1946	Germany, East 49-90	1950	1990	Mauritania 05-07	2006	2007	Syria 62-63	1963	1963
Bolivia 46-51	1947	1951	Ghana 60-66	1961	1966	Mauritania 08-NA	2009	2010	Syria 63-NA	1964	2010
Bolivia 51-52	1952	1952	Ghana 66-69	1967	1969	Mexico 15-00	1946	2000	Taiwan 49-00	1950	2000
Bolivia 52-64	1953	1964	Ghana 72-79	1973	1979	Mongolia 21-93	1946	1993	Tajikistan 91-NA	1992	2010
Bolivia 64-69	1965	1969	Ghana 81-00	1982	2000	Morocco 56-NA	1957	2010	Tanzania 64-NA	1965	2010
Bolivia 69-71	1970	1971	Greece 67-74	1968	1974	Mozambique 75-NA	1976	2010	Thailand 44-47	1946	1947
Bolivia 71-79	1972	1979	Guatemala 54-58	1955	1958	Myanmar 58-60	1959	1960	Thailand 47-57	1948	1957
Bolivia 80-82	1981	1982	Guatemala 58-63	1959	1963	Myanmar 62-88	1963	1988	Thailand 57-73	1958	1973
Botswana 66-NA	1967	2010	Guatemala 63-66	1964	1966	Myanmar 88-NA	1989	2010	Thailand 76-88	1977	1988
Brazil 64-85	1965	1985	Guatemala 66-70	1967	1970	Namibia 1990-NA	1991	2010	Thailand 91-92	1992	1992
Bulgaria 44-90	1946	1990	Guatemala 70-85	1971	1985	Nepal 1846-1951	1946	1951	Thailand 06-07	2007	2007
Burkina Faso 60-66	1961	1966	Guatemala 85-95	1986	1995	Nepal 51-91	1952	1991	Togo 60-63	1961	1963
Burkina Faso 66-80	1967	1980	Guinea 58-84	1959	1984	Nepal 02-06	2003	2006	Togo 63-NA	1964	2010
Burkina Faso 80-82	1981	1982	Guinea 84-08	1985	2008	Nicaragua 36-79	1946	1979	Tunisia 56-NA	1957	2010
Burkina Faso 82-87	1983	1987	Guinea 08-10	2009	2010	Nicaragua 79-90	1980	1990	Turkey 23-50	1946	1950
Burkina Faso 87-NA	1988	2010	Guinea Bissau 74-80	1975	1980	Niger 60-74	1961	1974	Turkey 57-60	1958	1960
Burundi 1962-1966	1963	1966	Guinea Bissau 80-99	1981	1999	Niger 74-91	1975	1991	Turkey 60-61	1961	1961
Burundi 66-87	1967	1987	Guinea Bissau 02-03	2003	2003	Niger 96-99	1997	1999	Turkey 80-83	1981	1983
Burundi 87-93	1988	1993	Haiti 41-46	1946	1946	Nigeria 66-79	1967	1979	Turkmenistan 91-NA	1992	2010
Burundi 96-03	1997	2003	Haiti 50-56	1951	1956	Nigeria 83-93	1984	1993	UAE 71-NA	1972	2010
Cambodia 53-70	1954	1970	Haiti 57-86	1958	1986	Nigeria 93-99	1994	1999	Uganda 66-71	1967	1971
Cambodia 70-75	1971	1975	Haiti 86-88	1987	1988	Oman 1741-NA	1946	2010	Uganda 71-79	1972	1979
Cambodia 75-79	1976	1979	Haiti 88-90	1989	1990	Pakistan 47-58	1948	1958	Uganda 80-85	1981	1985
Cambodia 79-NA	1980	2010	Haiti 91-94	1992	1994	Pakistan 58-71	1959	1971	Uganda 86-NA	1987	2010
Cameroon 60-83	1961	1983	Haiti 99-04	2000	2004	Pakistan 75-77	1976	1977	Uruguay 73-84	1974	1984
Cameroon 83-NA	1984	2010	Honduras 33-56	1946	1956	Pakistan 77-88	1978	1988	Uzbekistan 91-NA	1992	2010
Cen African Rep 60-65	1961	1965	Honduras 63-71	1964	1971	Pakistan 99-08	2000	2008	Venezuela 48-58	1949	1958
Cen African Rep 66-79	1966	1979	Honduras 72-81	1973	1981	Panama 49-51	1950	1951	Venezuela 05-NA	2006	2010
Cen African Rep 79-81	1980	1981	Hungary 47-90	1948	1990	Panama 53-55	1954	1955	Vietnam 54-NA	1955	2010
Cen African Rep 81-93	1982	1993	Indonesia 49-66	1950	1966	Panama 68-82	1969	1982	Yemen 18-62	1946	1962
Cen African Rep 03-NA	2004	2010	Indonesia 66-99	1967	1999	Panama 82-89	1983	1989	Yemen 62-67	1963	1967
Chad 60-75	1961	1975	Iran 25-79	1946	1979	Paraguay 40-48	1946	1948	Yemen 67-74	1968	1974
Chad 75-79	1976	1979	Iran 79-NA	1980	2010	Paraguay 48-54	1949	1954	Yemen 74-78	1975	1978
Chad 82-90	1983	1990	Iraq 32-58	1946	1958	Paraguay 54-93	1955	1993	Yemen 78-NA	1979	2010
Chad 90-NA	1991	2010	Iraq 58-63	1959	1963	Peru 48-56	1949	1956	Yugoslavia 44-90	1946	1990
Chile 73-89	1974	1989	Iraq 63-68	1964	1968	Peru 62-63	1963	1963	Zambia 67-91	1968	1991
China 49-NA	1950	2010	Iraq 68-79	1969	1979	Peru 68-80	1969	1980	Zambia 96-NA	1997	2010
Colombia 49-53	1950	1953	Iraq 79-03	1980	2003	Peru 92-00	1993	2000	Zimbabwe 80-NA	1981	2010
Colombia 53-58	1954	1958	Ivory Coast 60-99	1961	1999	Philippines 72-86	1973	1986			

Figure A-2: Sample autocratic regimes, 1946–2010

TABLE A-1: Summary statistics, untransformed

Variable	Mean	Std. Dev.	Min.	Max.	N
Non-viol. protest campaign	0.038	0.192	0	1	4510
Year	1979.67	16.536	1946	2010	4510
Personalization	0.425	0.276	0	1	4510
Party personalization	0.235	0.269	0	1	4510
Security personalization	0.466	0.293	0	1	4510
Time since last protest	11.958	12.04	0	64	4510
Leader tenure (log)	1.867	1.021	0	4.043	4510
Population (log)	9.109	1.371	5.605	14.099	4510
Region NVC onsets (log)	0.295	0.448	0	2.079	4510

Summary statistics

TABLE A-2: Summary statistics, standardized

Variable	Mean	Std. Dev.	Min.	Max.	N
Non-viol. protest campaign	0.038	0.192	0	1	4510
Year	1979.67	16.536	1946	2010	4510
Personalization	0	1	-1.543	2.086	4510
Party personalization	0	1	-0.874	2.84	4510
Security personalization	0	1	-1.591	1.821	4510
Time since last protest	0	1	-0.993	4.322	4510
Leader tenure (log)	0	1	-1.829	2.132	4510
Population (log)	0	1	-2.556	3.639	4510
Region NVC onsets (log)	0	1	-0.657	3.979	4510

Security personalization

Previous measurement. One way of measuring security personalization employs a categorical regime typology, such as that developed by Geddes (2003). However, using a binary indicator of a personalist regime does not account for variation in each dictator's personalist power over time; nor does such an indicator allow us to compare personalist powers relative to the powers of ruling party or military because a binary indicator of regime type contains no information on the specific security personalization policies dictators implement during their tenure.

An alternative method is to construct a dynamic measure that captures security forces structure. One of the earliest attempts measures the number of effective organizations within the armed forces and their respective strength. Dictators engage in organizational proliferation of their security forces by creating rival parallel organizations, which increases the number of organizations and thus potential units necessary to coordinate a successful coup attempt. Intra-agency competition for greater resources, lack of information about the relative strength of various units, and simply a larger number of veto players all impede security forces coordination to oust their leader (Quinlivan 1999; Powell 2012; De Bruin 2018).

The earliest study that tests this logic systematically is Belkin and Schofer (2003). They measure the degree of counterbalancing by counting the number of military and paramilitary organizations and comparing the relative sizes of paramilitary forces to the total number of entire military. Similarly, Pilster and Böhmelt (2011, 2012) count the effective number of ground combat organizations and each organization's personnel. On the other hand, De Bruin (2020) constructs the "counterweight" forces data by identifying security forces that are operationally independent from the defense ministry and deployed within 60 miles of the capital.

While data on counter-balancing, the count of military and paramilitary organizations, and counterweight units probably best capture the concept of organizational proliferation, they often do not consider the wide range of security forces that may, in some contexts, help protect the dictator from coups or other internal threats.The effective number of organizations in the armed forces, for example, does not typically capture the proliferation of internal security organizations that remain outside the military, such as presidential guard units, secret police, or internal intelligence agencies. As importantly, these measures do not capture concepts related to *loyalty* to the dictator.

We acknowledge that loyalty appointments in the security sector can also be conceptualized as one of many (perhaps overlapping) coup-proofing strategies such as paying the military or creating multiple (perhaps counter-balancing) organizations (Pilster and Böhmelt 2011, 2012; De Bruin 2020). Conceptually, however, we focus not on the mere existence or number of security organizations but on the power in the relationship between the leader and the security organization. The creation of a paramilitary unit to fight a counter-insurgency with the unit commanded by a seasoned security officer may be something the military desires and enables. Such an unit might, in practice, counter-balance other units and increase the number of effective military organizations. However, such a unit is, to our mind, qualitatively different than the creation of a special military unit, paramilitary, or presidential guard (de facto) outside the normal chain of command, formed (and funded) against the wishes of the regular military, and led by a loyal family member of the regime leader. The latter, we posit gives the leader more power over the organizations in the security sector (i.e. the creation/appointment is an observed manifestation of that power); and as a result, the new unit is more loyal to the leader than other units.

We show that personalized security forces lower the likelihood of protest onset; and we argue that loyalty is a feature of SFP that links personalized forces to onset. We cannot measure de facto loyalty – a psychological state of mind – directly but we can measure observed moves to reshape the security apparatus that reflect more de facto relative power for the leader and should, we argue, breed more loyal forces.[46]

Security personalization. Our measure of *Security personalization* utilizes observable indicators from newly-collected and time-varying data on the dimensions of autocratic rule (Geddes, Wright, and Frantz 2018; Wright 2021). The data ranges from 1946 to 2010 and the following questions are asked for every regime-year observation as of January 1st:[47]

1. *Personal paramilitary*: Does the regime leader create paramilitary forces, a president's guard, or new security force loyal to himself? **(0/1)**

2. *Personal control*: Does the regime leader personally control the security apparatus or political police? **(0/1)**

3. *Personal appointment*: Does the regime leader have discretion over appointments to high office or appoint relatives in these positions? **(0/1)** × Did the regime leader come to power with military backing (mostly through coups)? **(0/1)**

4. *Personal purge*: Does the regime leader imprison or kill military officers from other groups without a reasonably fair trial? **(0/1)**

5. *Personal promotion*: Does the regime leader promote military officers loyal to himself or from his ethnic, tribal, regional groups rather than merit and seniority? **(0/1/2)**

The first four personalization policies are dichotomous (1 if answered yes to the question, 0 if no). We construct the *Personal appointment* item using information from (1) whether the leader has the discretion to appoint personnel to high office based on personal loyalty or to promote family members; and (2) whether the leader came to power with military backing as opposed to party or revolutionary group backing.[48] This is done in order to capture the cases where a (former) military officer defied the military as an institution in making personal appointments to high offices than a non-military backed leader.

On the other hand, *Personal promotion* item is coded on an ordinal scale.[49]

Because we have both binary and ordinal levels of measurement, we employ the hybrid Item Response Theory model to construct the latent scale of *Security personalization*. The IRT models are determine the relationship between the latent ability (in this case, the degree of security

[46]There may be other consequences of personalism as it relates to security forces, such as *fewer* budget resources allocated to the security forces overall (this would imply substitution between loyalty and material resources) but these other (possible) consequences of personalized forces are not the focus of our argument.

[47]Because the data is collected for January 1 of each calendar year, the measure picks up changes in these indicators in the prior calendar year, effectively lagging the relevant information by one year.

[48]The variable for (1), *officepers* picks up possible loyalty appointments in any high office positions, including the military and security apparatus as well as the supporting political party or even cabinet positions. Since appointments to the latter offices are not part of the concept we attempt to code, we attempt to isolate the loyalty appointments in the military and security apparatus. Our intuition is that the *officepers* variable is more likely to pick up security loyalty appointments than non-security related loyalty appointments when the leader comes from the military as opposed to cases where the leader comes from the party or a rebel movement because personalization typically entails first sidelining and purging members of the group that initially supports the leader (the military) and then proceeds to cast a wider net in organizations outside the initial seizure/launching group.

[49]This variable is coded 0 if the regime leader does not use loyalty in promotion AND no widespread forced retirement OR no military; coded 1 if promotions of top officers loyal to the regime leader or from his group; and coded 2 if the regime leader promotes officers loyal to him or from his ethnic, tribal, regional, or religious groups OR widespread forced retirement is used.

forces personalization) and the items (the observable personalization policies) (Reise and Waller 2009). The hybrid model allows us to fit different types of IRT models on subsets of items depending on how they are coded. The items that are coded as binary indicators are fitted using the two-parameter logistic model which allows each item to have varying levels of discrimination, while the ordinal item (i.e., *Personal promotion*) is fitted with a graded response model, which is an extension of the two-parameter model to ordered logistic model.

TABLE A-1: IRT Hybrid Model

	Item	Discrimination	Difficulty (θ)
Two-parameter model			
	Personal paramilitary	1.218	.630
	Personal control	1.468	-.405
	Personal appointment	1.280	1.486
	Personal purge	2.771	.397
Graded response model			
	Personal promotion	2.260	-.908 (*Personal promotion*≥1) 237 (*Personal promotion*=2)

Table A-1 presents the discrimination and difficulty parameters for each item. *Personal purge* has the highest level of discrimination, which is the ability to distinguish between lower and higher levels of security personalization. In other words, the dictator's ability to purge military officers based on personal loyalty provides the greatest amount of information on whether the dictator has highly personalized security forces. On the other hand, the difficulty parameter (θ) indicates the probability of positive observation for each personalization policy. For instance, having discretion over appointments to high office is the most difficult (i.e., least likely to find a positive observation) ability to achieve for the regime leaders than any other types of personalization policies.

Item response curves for binary outcome items in Figure A-3 (the first four panels) are the visual representation of the discrimination and difficulty parameters for each observable indicator. Steeper curves illustrate higher levels of discrimination. The curve for *Personal purge* is the steepest, revealing the greatest amount of information. Second, the difficulty parameter of each policy is located on the point at which the item response curve crosses the 0.5 probability of positive observation. Since a zero mean for θ (the latent ability) is assumed, relatively "easy" items are located on the left-side with negative difficulty parameter values while relatively "hard" items are located on the right-side with positive difficulty parameter values.

The last panel, for the *Personal promotion* item, has three outcomes and has two difficulty parameters and boundary characteristic curves as it is fitted using the graded response model. The graded response model is defined in terms of cumulative probabilities; therefore, each difficulty parameter indicates a point at which an observation with certain latent ability ($\theta = b_{ik}$) has a 50% chance of corresponding on category k or higher. Looking at the panel, we can see that an observation with latent ability of -0.908 has a 50% chance of responding to category 1 or greater in *Personal promotion* coding; an observation $\theta = 0.237$ has a 50% chance of response to category 2 in *Personal promotion* coding.

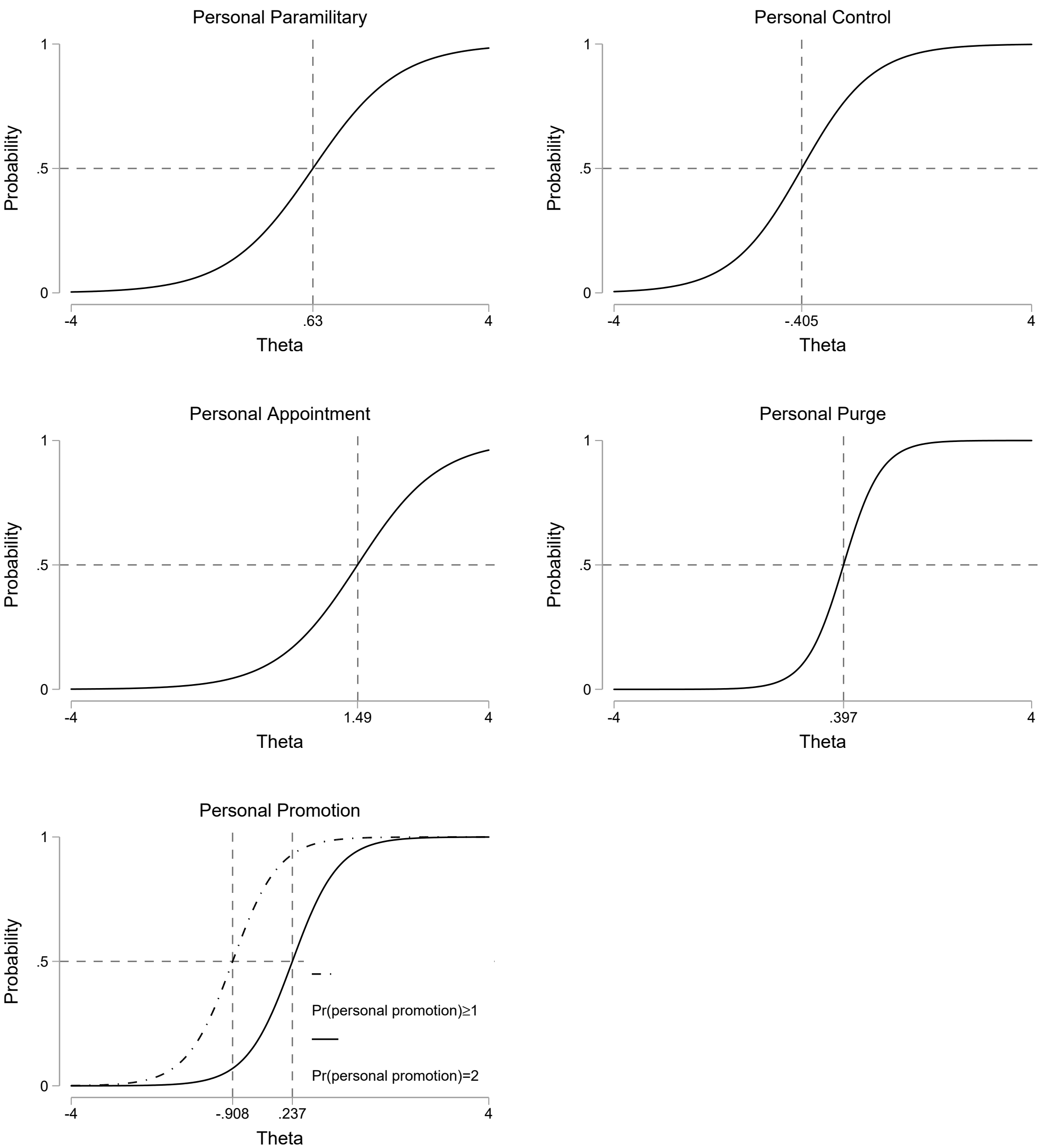

Figure A-3: Hybrid IRT Item Response Curves
Notes: *Personal promotion* panel shows boundary characteristic curves as the item has three outcomes and is fitted with the graded response model. Panels for rest of the items with two outcomes are shown item characteristic curves.

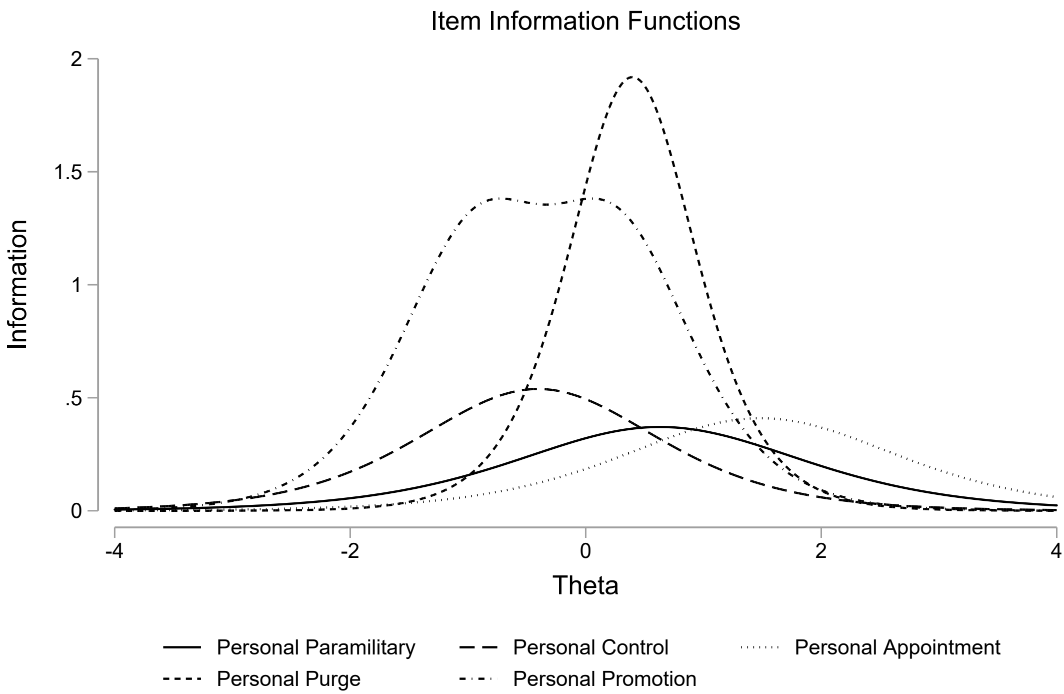

Figure A-4: Hybrid IRT Item Information Functions

The item information functions in Figure A-4 represent the amount of information for each item in the IRT model. Items that reveal more information are able to measure the latent ability around the estimated difficulty parameter with greater precision (i.e., taller and narrower curves). The amount of information is proportion to the discrimination parameter. The plot illustrates that each item information function's shape corresponds with the discrimination parameters in Table A-1 and the steepness of curves in Figure A-3.

A generalized structural equation model The IRT framework assumes that each item is independent of the other items; and thus estimates (and variances) for each item in the logit link function are assumed to be uncorrelated. Using a generalized structural equation model (SEM), we can probe this assumption. First, the generalized SEM allows for the model to estimate different link functions for each item. Thus it is possible to, for example, estimate some items using an ordered logit and some an ordinary logit, as in a hybrid-2PL-GRM described above. However, the generalized SEM also allows for gaussian models with identity link functions (which is an OLS regression). The benefit of this latter feature is that it allows the model to estimate correlated variances for items, thus relaxing a key (untested) assumption of the IRT framework.

If real-world information contained in some of items is the same the items would *not* independent. For example, creating a new personal paramilitary loyal to the leader may also indicate personal control over the security apparatus. If this is the case, then the two items would not be independent. Similarly, demoting or executing high-ranking officers deemed insufficiently loyal may indicate both control over the security apparatus and a purge. We thus examine a generalized SEM to test whether item variances are correlated; and we find they are, but for only one pair of items: *personal control* and *personal paramilitary*. We thus estimate a generalized SEM that is identical to the hybrid IRT model discussed above but instead use a Gaussian distribution with an identity link (instead of binomial distribution with a logit link) for these two items, allowing their errors to be correlated. In the reproduction files we show that allowing correlated errors produces a better fit than assuming them away (conditional on a Gaussian model with id link).

TABLE A-2: Generalized SEM

Item	family	link
Personal paramilitary	gaussian	identity
Personal control	gaussian	identity
Personal appointment	binomial	logit
Personal purge	ordinal	logit
Personal promotion	binomial	logit

The latent estimates of security personalism produced by the hybrid IRT and this generalized approach are nearly identical, as shown in Figure A-5. Further, when we create a binary treatment variable for analysis from the distribution of latent security personalism, we find that it identifies the exact same set of treated observations in the sample. Because the generalized SEM latent estimate accounts for this correlation in variances between two items, we use this generalized SEM modification of the hybrid IRT approach for the main treatment variable used in the analysis. However, we show in the reproduction files that none of the main results differ when using latent estimate from the hybrid IRT approach.

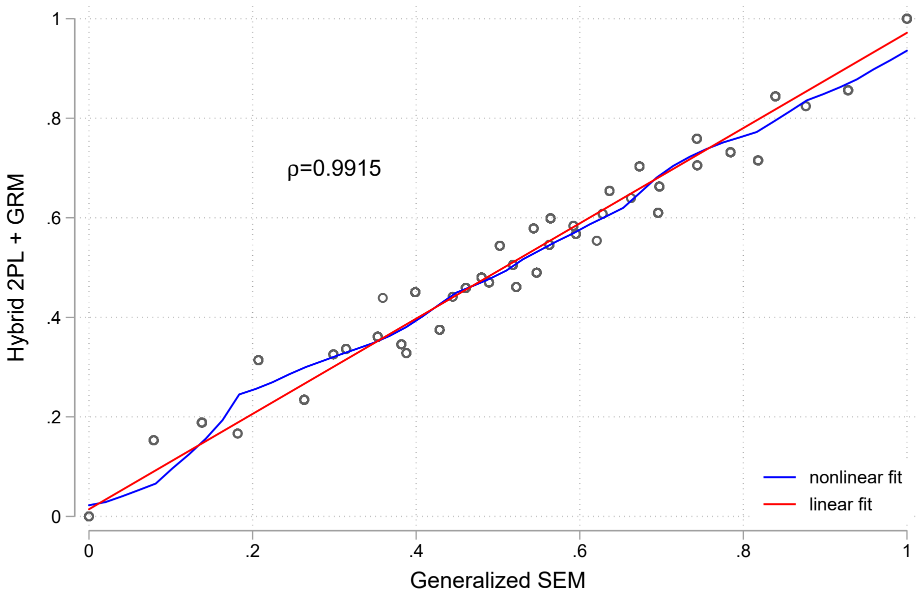

Figure A-5: Additional latent measures of SFP

Appendix B: Additional results for protest onset

Additional potential confounders

Our baseline specification includes a minimum of covariates that we posit are both pre-treatment and possibly confounders. In the baseline specification we adjust for three covariates: time since last onset, leader duration, and region protest. The first is necessary to mimic a survival model using a binary DV model; this is well established in the literature. The second, leader tenure, is correlated with both outcome and treatment and is arguably pre-treatment because dictators who have survived create different types of expectations – both for protesters and for the security apparatus. The literature, for example, posits that coup opportunities and timing (indicative of the expectations of security agents) are a function of leader tenure (Sudduth 2021). And leader's may become more unpopular over time, prompting protest (Chenoweth and Ulfelder 2017). Finally, region protest is exogenous; and the literature suggests regional contagion as pathway to explain protest mobilization. Further, dictators may strategically respond to external protests by altering the composition of their security forces. Thus region protest may influence both protest onset and security personalism.

Figure B-1 shows that the main result fore security personalization is robust to the inclusion any one of 42 additional covariates. The vertical axis is the estimated coefficient for $\beta_{SecurityPersonalization}$ for each of 42 tests, while the horizontal axis lists the names of the 42 additional covariates that are added to the base line specification, one at a time. The red horizontal line at -0.12 is the estimate for the base line model reported in the main text. This allows visual comparison of how changes to the model specification (adding the covariate listed on the horizontal axis) changes the estimate of interest relative to the estimate in the base line model. The plot shows that the main finding is robust for all the additional covariates; including a measure of *counter-weight* organizations and *heavily-armed counter-weights* in the military, from De Bruin (2018).

We make no claim that these covariates are pre-treatment; in fact, one could argue that some, like repression or the size and spending on the military, are post-treatment (i.e., endogenous to security personalization). We include plausible post-treatment variables to demonstrate robustness, in part because these variables are proxies for possible alternative theories of civil-military relations and protest mobilization.

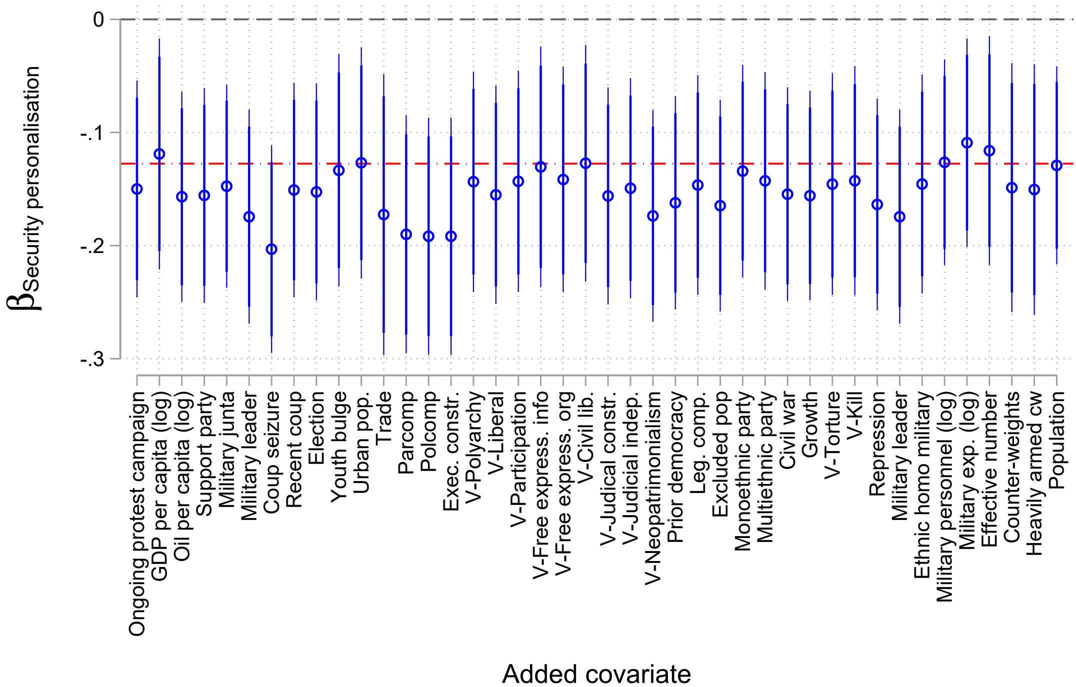

Figure B-1: Additional potential confounders

Semiparametric analysis

This section examines *Security personalization* and nonviolent campaign onset with a semiparametric model. We use Baltagi and Li's (2002) fixed-effects semiparametric estimator (hereafter BL-FE), which mixes a parametric component of a model with a non-parametric component. The advantage of this approach is to allow for (possible) non-linear relationships between a primary variable of interest (*Security personalization* without imposing a specific functional form on the relationship. The estimator stems from the following equation, where $x_{i,t}\theta$ is the parametric component of the model and $f(z_{i,t})$ is the non-parametric component; α_i are the unit fixed effects; $z_{i,t}$ is the main explanatory variable of interest, for which we do not want to impose a specific functional form; and $y_{i,t}$ is the outcome variable. The the estimator is thus "an additive partially linear model" (Li 2000, 1073).

$$y_{i,t} = x_{i,t}\theta + f(z_{i,t}) + \alpha_i + \varepsilon_{i,t} \tag{4}$$

The BL-FE estimator deals with α_i via differencing the equation and approximates $f(\cdot)$, a (possibly non-linear) link function, with splines that allow for many possible nonlinear functions relating the conditional variation in each series to the conditional variation in the other.[50] This approach yields values $\hat{u_{i,t}}$ from the following equation:

$$\hat{u_{i,t}} = y_{i,t} - x_{i,t}\hat{\theta} - \hat{\alpha_i} = f(z_{i,t}) + \varepsilon_{i,t} \tag{5}$$

The function $f(\cdot)$ can then be fit by regressing $\hat{u_{i,t}}$ on $z_{i,t}$ using a non-linear smoother.

Figure B-2 shows the non-linear fit for a specification that adjusts for minimal potential confounders (region NVCs, leader duration, and time since last onset) as well as regime-case and year fixed effects. The vertical axis measures the probability of campaign onset (conditional on the covariates, include the fixed effects) and the horizontal axis measures the level of security personalism. The dashed horizontal line at 0.039 is the average probability of onset in the sample, shown for reference. The polynomial fit, depicted as a dashed blue line and associated 95 percent confidence interval, declines from about 5.5 percent at low levels of security personalism to 3.5 percent onset risk at the median level of security personalism (roughly zero). Once the level of security personalism reaches about 0.5, the conditional probability of onset remains between 3 and 3.5 percent, never declining much further.

This plot suggests that security personalism reduces onset risk by over 2 percent (from 5.5 percent to between 3 and 3.5 percent), which is consistent with the marginal effects results we report in the main text. However, the marginal effect is concentrated in the bottom half of the distribution of security personalism, which suggests that moving from low to medium levels of security personalism benefits dictators by reducing protest campaign onset risk but there are few benefits to dictators in reducing this risk beyond this point. This result also indicates that a binary treatment variable that using the median value of *Security personalism* to distinguish low and high security personalism captures most of the continuous marginal effects.

[50] Libois and Verardi (2013) implement this estimator using a linear combination of a set of (kth degree) B-splines. We estimate B-splines with a power that stabilizes the non-linear relationship; see Appendix D for a discussion of degree selection.

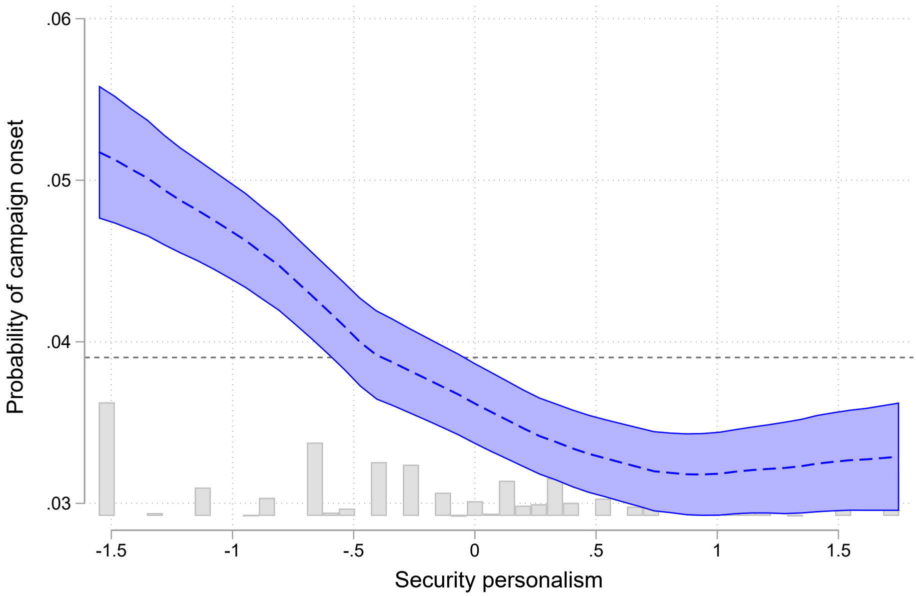

Figure B-2: Semiparametric relationship between *Security personalism* and NVC onset.

Linear probability models

Time-varying confounding Next, we test a series of fixed effects linear probability models (FE-LPM) with a binary treatment variable that allow us to examine possible bias from unmeasured time-varying confounders. An LPM is an appropriate modeling choice for binary dependent variables, especially when there are relatively few positive outcomes (Timoneda 2021). The treatment indicator is simply whether *security personalization* is greater than or equal to the sample median: half of sample observations are 'treated.'[51] In a bivariate test, NVC onset is 1.2 percent less likely in treated observations than untreated ones (3.4 versus 4.6 percent), a statistically significant difference. Across all years, treated dictatorships are 2.9 percent less likely to experience a protest campaign (including ongoing campaigns) than untreated ones (5.8 versus 8.7 percent), again a statistically significant difference.

First, we test a standard 2-way FE model, which accounts for regime-specific confounders and common time shocks. The estimate for *security personalism* in column (1) in Table B-1 is negative and significant: treatment reduces campaign onsets by about 2.5 percent. The 2-way FE model therefore produces similar estimates to the non-linear models tested thus far.[52]

Our first step to address any bias from *unobserved* time-varying confounders is to test an FE model with a nonlinear regime-specific time trend. This test, reported in column (2) of Table B-1, yields a stronger result, suggesting that treatment reduces onset by 3.3 percent. Next we test an interactive FE model, which allows time shocks to vary by dictatorship, ruling out confounding from, for example, world oil price shocks or the end of the Cold War and advent of Western democracy-promotion efforts influenced dictatorships differently. Column (3) reports this test and suggests that treatment reduces protest onsets by 2.1 percent.

The next set of 2-way FE models adopt a dynamic framework while still accounting for panel heterogeneity and common time shocks. Column (4) includes two lags of the past treatment to rule out the possibility that past treatments affect the outcome. The lags are both individually and jointly insignificant (p-value= 0.25 in the joint test). The estimate for treatment is substantially stronger (-0.044). Column (5) adjusts for lagged outcomes; again the lagged outcomes are insignificant and the treatment result remains (-0.026).

Although our outcome measures major nonviolent protest campaigns, anti-regime mobilization also occurs at much lowers levels in a much wider range of dictatorships. Smaller past protests might both spur the regime to restructure its security forces to make them more loyal and may be a harbinger of larger mass protest campaigns. Column (6) adjusts for the lagged time trend in a latent measure of *Mobilization for democracy* from the Varieties of Democracy project.[53] This variable, which is constructed from country-expert coders, captures both *small* and large mobilization events, such as demonstrations, strikes and sit-ins. The estimate for lagged pro-democracy mobilization is positive and significant, suggesting smaller scale protests are a precursor to the large mass mobilization campaigns that we model as the outcome.[54] But, again, the treatment estimate remains robust (-0.031).

[51]This operationalization of a binary treatment variable means that 75 of 280 regimes change treatment status; 97 regimes remain always treated; and 108 regimes are never treated. A semiparametric FE test indicates that selecting the median value as a cut-point for coding a binary treatment variable is appropriate because the (negative) marginal effect of security personalism tails off after crossing the median value; see Appendix B.2.

[52]A random effects probit model with minimal confounders (region protest, leader duration, and time since last onset) and adjusting for 5-year calendar time period effects yields a baseline marginal effect estimate of -0.023.

[53]See Pemstein et al. (2019); the variable name is `v2cademmob`.

[54]Including 3- and 4-year lags produces nearly identical results (Hamilton 2018).

The penultimate test adjusts for both lagged treatment (D_{t-1}) and outcome (Y_{t-1}), assuming each are endogenous and using 2- and 3-year lags of these variables as excluded instruments.[55] Assuming that the deeper lags are exogenous, this approach simultaneously blocks confounding via both past treatment and past outcome. Column (7) reports this models' results; and the estimate for our treatment variable is even stronger (-0.080).

The final test using the FE-LPM framework in column (8) relaxes linearity assumptions by testing a within-unit matching model (weighted fixed effects, WFE).[56] While the estimate size for *security personalism* is very strong, the standard error is also quite large and thus the estimate is not statistically different from zero. Throughout, though, the tests in Table B-1 rule out bias from unobserved time-varying confounding. Thus with appropriate caution given the large error estimate, the result in column (8) provides no evidence that results are biased away from zero.

[55]A test that includes these deeper lags, which assesses whether these longer lags can be treated as excluded instruments, indicates that they are individually and jointly *insignificant*.

[56]The within-unit proportion of treated observations is *not* constant across units, so this is an appropriate check on the results (Imai and Kim 2019, 475).

TABLE B-1: Security Personalization and NVC Onset, Linear Probability Models

	FE-LPM (1)	FE-LPM (2)	FE-LPM (3)	FE-LPM (4)	FE-LPM (5)	FE-LPM (6)	FE-LPM (7)	WFE (8)
Treatment	-0.025*	-0.033*	-0.021*	-0.044*	-0.026*	-0.031*	-0.080*	-0.088
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.03)	(0.33)
Region NVC onsets	0.013*	0.019*	0.016*	0.014*	0.015*	0.011*	0.015*	0.049
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.23)
Leader tenure (log)	0.006	0.004	0.006	0.008	0.009	0.014*	0.009	0.120
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.60)
Treatment$_{t-1}$				0.000			0.059	
				(0.02)			(0.03)	
Treatment$_{t-2}$				0.026				
				(0.01)				
NVC onset$_{t-1}$					-0.049		-0.007	
					(0.03)		(0.06)	
NVC onset$_{t-2}$					-0.014			
					(0.03)			
Mobilization$_{t-1}$						0.052*		
						(0.01)		
Mobilization$_{t-2}$						-0.022*		
						(0.01)		
N × T	4535	4559	4535	4007	4007	3155	3777	4535
Regime-cases	256	280	256	218	218	168	206	256
Regime-case FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓		✓	✓	✓	✓	✓	✓
Case-specific time trend		✓						
Interactive FE			✓					

Dependent variable is protest campaign onset. All specifications adjust for time since last NVC onset (log). Case-specific time trend in (2) is a non-linear calendar time trend ($time + time^2$) interacted with each regime-case fixed effect. Columns (1) and (2) drop singleton panels due to partialling out panel effects. F-statistic for excluded instruments in Column (7) is 98.6. Cluster-robust standard errors in parentheses. * $p < .05$.

Alternative unit effects approaches

This subsection reports results from tests that model (time invariant) unit effects in various ways. In the main text we report results from random effects (RE) and correlated random effects (CRE) estimators, with autocratic regime-cases as the cross-section unit.[57] The latter approach entails including the (unit) mean level of all covariates on right-hand side of the equation to isolate the time-varying (or 'within') effects of the explanatory variables.

Of the 280 regimes in the analysis, 172 regimes (1,957 observations) have no variation in the dependent variable during the sample period and 107 (2,553 observations) have variation.[58] A conditional logit estimator reduces the sample to the latter 107 regimes, excluding the other 172 regimes. This produces biased estimates of the marginal effects (Cook, Hays, and Franzese 2018): the baseline rate of onset in full sample is 4 percent; in the reduced sample it is 7 percent.

There are a number of ways to approach this issue. We opted for the correlated random effects (CRE) approach in the main text and showed that this produces stronger substantive effects than a random effects estimator. The CRE approach preserves the full sample, allowing estimates to incorporate information using the partial correlation between the over-time patterns and the unit-mean patterns for the full sample, not just the sample with regimes that experience onset at some point during the sample period. A series of linear probability model with regime-case fixed effects, reported in Appendix B.3, support the main finding.

Fixed effects logit An alternative to the CRE and FE-LPM is a logit maximum likelihood estimators, similar to the probit models in the main text. We do this so that we can compare results from these estimators with the estimate from a penalized maximum likelihood logit with unit effects. Cook, Hays, and Franzese (2018) suggest estimating a limited dependent variable model with a separate intercept for each unit that has within variation in the dependent variable and a common intercept for all cases with no variation in the outcome. They then estimate this model with a penalized logit that allows for total separation, which arises from the fact that there are units with no variation in the outcome. We adopt this approach but adapt it by incorporating unit means for each of the groups that have within variation and a separate (common) group mean for the cases with no variation. In doing so, we also include the mean levels for dependent variable on the right-hand side of the equation, which is the fixed intercept information from a fixed effects model:

$$Pr(Protest_{i,t} = 1) = \beta X_{i,t} + \delta \bar{X}_j + \phi \bar{Y}_j + \varepsilon_{i,t}; \varepsilon \sim N(0, 1) \tag{6}$$

In this equation, the units are regime cases indexed by i, but the group means are indexed by j, which denotes the separate groups for each i with outcome variation and a common group for all i's with no variation in the outcome. Further, $\bar{Y}_j$ mimics the intercepts for j in a model with fixed intercepts. We then estimate this specification with a Firth logit. Shown in Figure B-3, this approach yields a similarly-sized estimate to the RE estimator but with a higher variance.

Finally, we can implement this modified unit effects approach suggested by Cook, Hays, and Franzese (2018) directly into the probit model we examine in the main text:

[57]For some countries, such as China, the regime-case is the same unit as the country because there is only one regime-case during the sample period. In other countries, such as Iran, there are multiple regime-cases during the sample period. In Iran, for example, there are two regime-cases: the monarchic regime led by Mohammad Reza Pahlavi (the Shah) that ended in the 1979 revolution and the subsequent theocratic regime.

[58]Of the 172 regimes, five last only one year and always have a protest onset and 167 have no onsets.

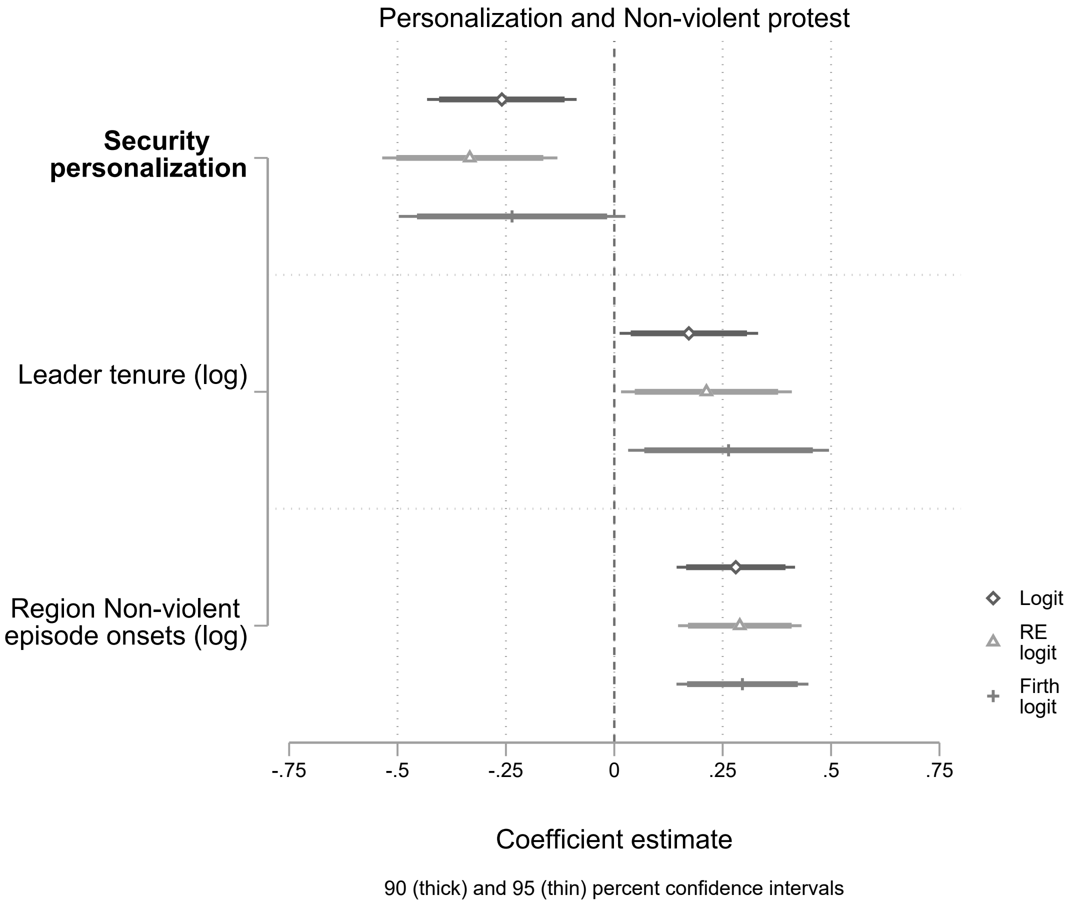

Figure B-3: Logit models

$$Pr(Protest_{i,t} = 1) = \alpha_j + \beta X_{i,t} + \varepsilon_{i,t}; \varepsilon \sim N(0,1) \tag{7}$$

Here, α_j means there is a separate intercept for each case with outcome variation (107 regimes) and separate common intercept for all other cases with no variation in the outcome (172 regimes). This approach yields an estimate of -0.122; however it is only statistically significant at the 0.181 level. For comparison, the main text reports that the probit estimate of interest is -0.108 while the RE probit is -0.128 – both statistically significant at the 0.05 level. In short, modeling the unit intercepts in the way suggested by Cook, Hays, and Franzese (2018) yields a similar estimate to the RE model, but with a substantially higher variance.

The main take-away from testing all these unit effects estimators is that the result for the estimated marginal effect reported in the main text is remarkably stable – even perhaps underestimated (absolutely). Approaches that use up substantial degrees of freedom, unsurprisingly yield higher variance estimates so some of the results reported here are not statistically significant at conventional levels. But these test should partially allay concerns about bias.

Time trends

Figure B-4 shows results from the main specification that address the possibility that the effects vary over calendar time. Importantly, we want to ensure that the main finding is not simply a product of Cold War super-power politics. The first two tests split the sample into the Cold War and the post-Cold War periods. While the latter estimate has wide confidence intervals, the point estimate is actually larger (in absolute size) than the estimate for the Cold War period. Next we test the full sample but model a common time trend in various ways: five-year period effects; a non-linear time trend (third order polynomial); and decade effects.

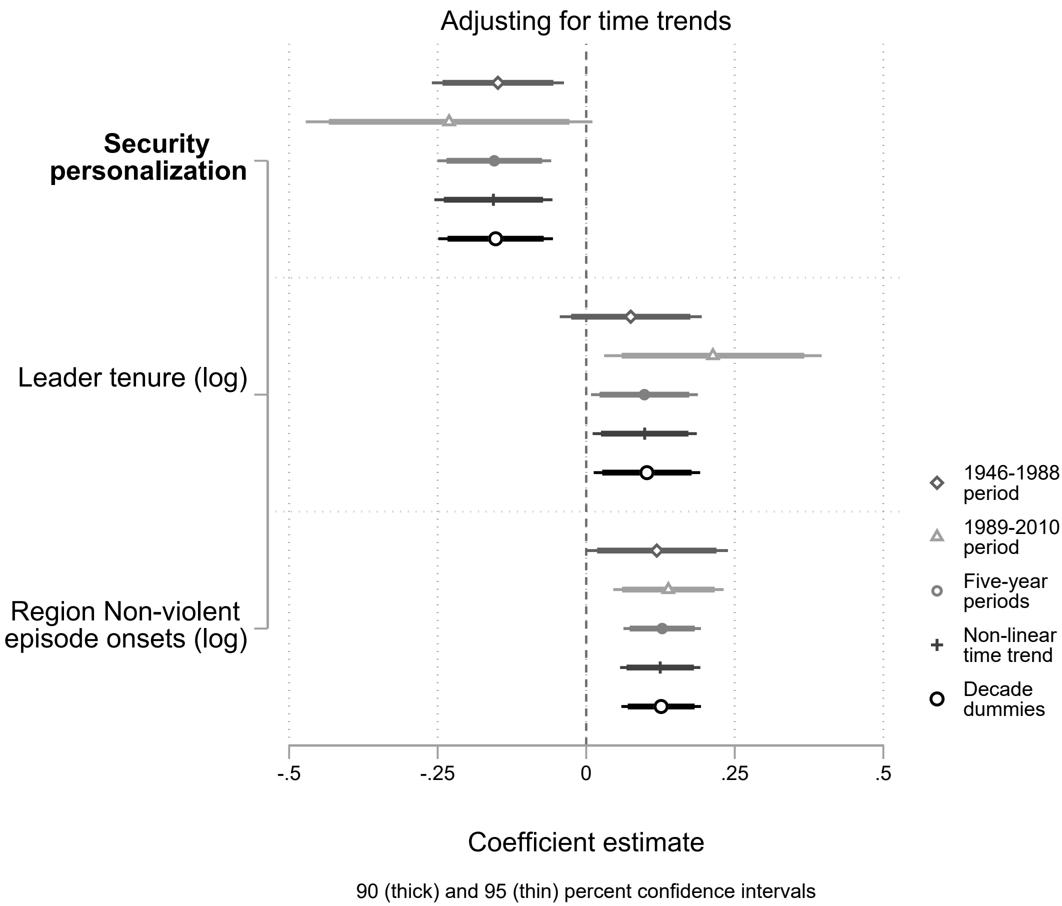

Figure B-4: Protest onset, modeling common time trends

Drop regions/decades

Figures B-5 and B-6 show that the main result is robust to dropping countries in any one of nine geographic regions and to dropping any one of six decade periods.

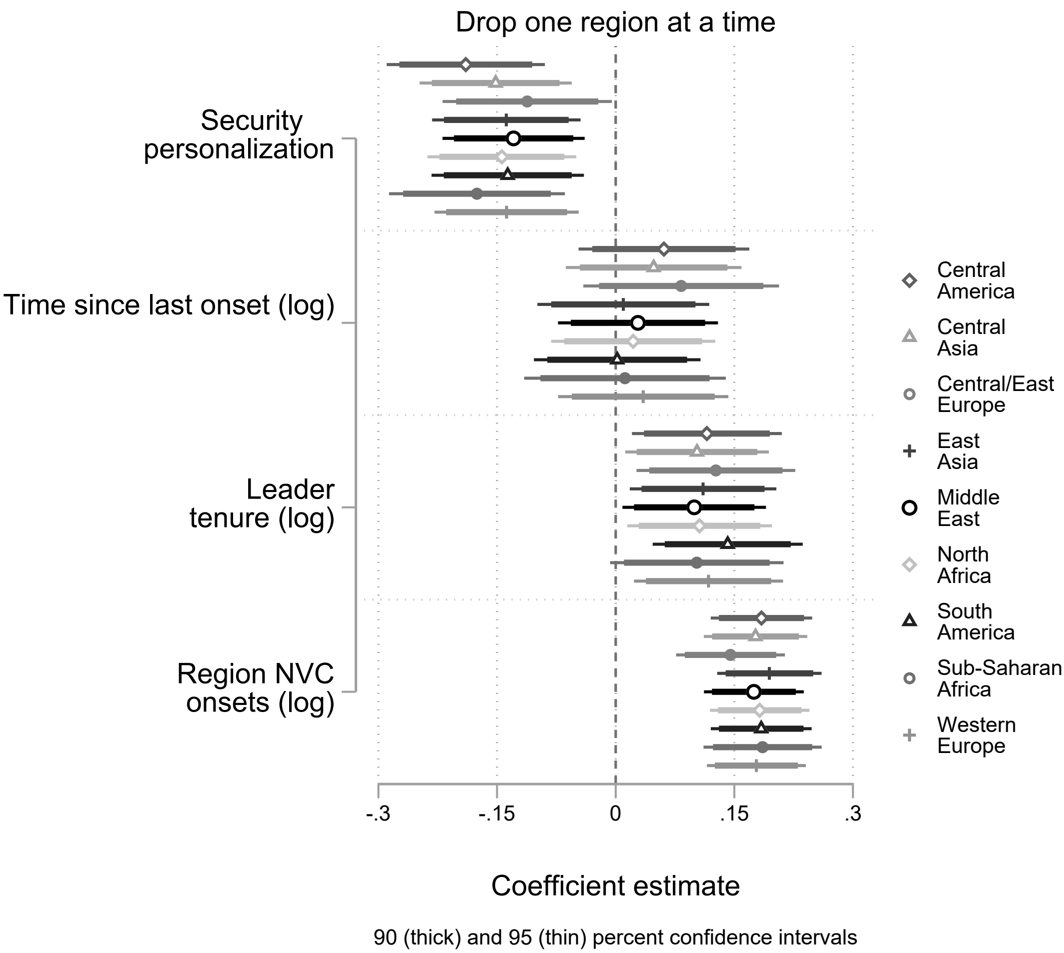

Figure B-5: Drop regions, one at a time

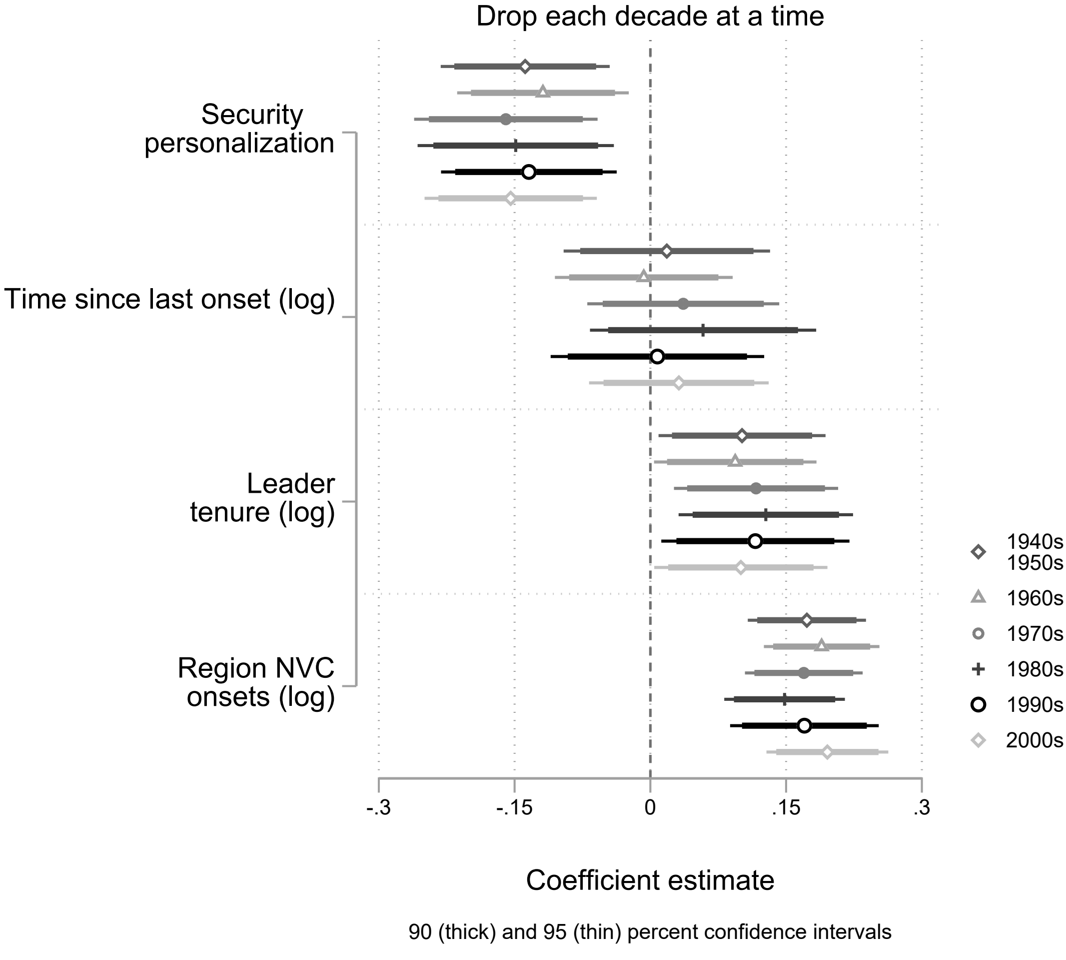

Figure B-6: Drop decades, one at a time

Alternate dependent variable coding

The main text examines models that include all nonviolent protest campaigns contained in any of three data sets: NAVCO, MEC, and NEVER. Figure B-7 shows that the main result for security personalization remains when we analyze data on protest campaigns from each of the data sets separately. That is, the main finding does not depend on adding campaigns from NEVER that are not included in NAVCO or MEC.

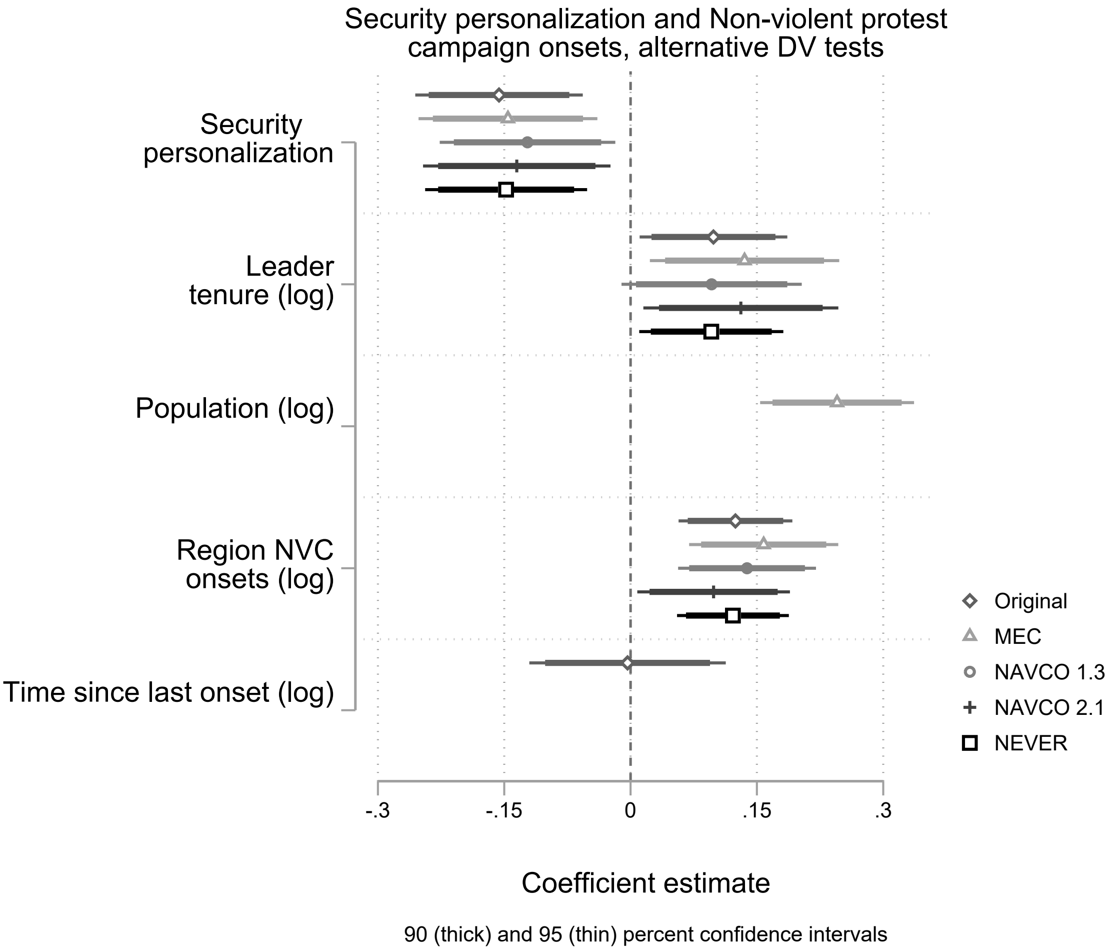

Figure B-7: Alternate measures of NVC onsets

Dropping covariates

Results reported in Table B-2 show that estimate for security personalization is negative when dropping any combination of variables from the base line specification used in the main text. The only estimate that is not statistically significant at the 0.05 level – reported in column (4) – is statistically significant at the 0.063 level. Further, specifying the time since last onset as a third-degree polynomial, instead of a natural log, yields an estimate that is significant at the 0.051 level. Thus the results reported in the main text are robust to specification changes.

TABLE B-2: Security personalization and NVC onset, dropping covariates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Security personalization	-0.155*	-0.095*	-0.137*	-0.106*	-0.112*	-0.116*	-0.155*	-0.145*	-0.112*
	(0.05)	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.04)
Time since last onset (log)	0.052	0.046	0.025	0.049	0.063	0.066	0.052	0.032	0.063
	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)
Cold war	-0.303*				-0.437*	-0.313*	-0.303*		-0.437*
	(0.11)				(0.10)	(0.11)	(0.11)		(0.10)
Leader tenure (log)	0.106*		0.121*				0.106*	0.112*	
	(0.05)		(0.05)				(0.05)	(0.05)	
Region NVC onsets (log)	0.139*			0.180*		0.139*	0.139*	0.178*	
	(0.03)			(0.03)		(0.03)	(0.03)	(0.03)	
(Intercept)	-1.674*	-1.810*	-1.819*	-1.844*	-1.559*	-1.661*	-1.674*	-1.853*	-1.559*
	(0.08)	(0.05)	(0.05)	(0.05)	(0.07)	(0.08)	(0.08)	(0.06)	(0.07)

Dependent variable is protest campaign onset; RE probit estimator; NxT=4510; standard errors clustered on 279 regime-cases in 117 countries; 1946–2010. * $p < .05$.

Cox models

The analysis thus far employs binary dependent variable models. In this section, we report results from a Cox proportional hazard model with shared frailties for each regime, which is similar to a random effects (intercept) logistic model. This estimator is a semi-parametric model in which the hazard rate is estimated non-parametrically even though the covariates retain a functional form. We test the baseline specification reported in the main text: leader tenure (log), population size (log) and regional protests (ln), as well as the security personalization variable. This specification violates the proportional hazards assumption (that the hazard rate is constant over time), and analysis of the Schoenfeld residuals suggests that the leader tenure variable violates the PH assumption. To address this, we interact leader tenure with time since last onset (log) and re-analysis of the Schoenfeld residuals suggests that the PH assumption is no longer violated. We report the result from this specification, with *Leader tenure (log)* $\times$ *Time to event (log)* included. The estimated coefficient for security personalization is -0.289, and is statistically significant at the 0.076 level.

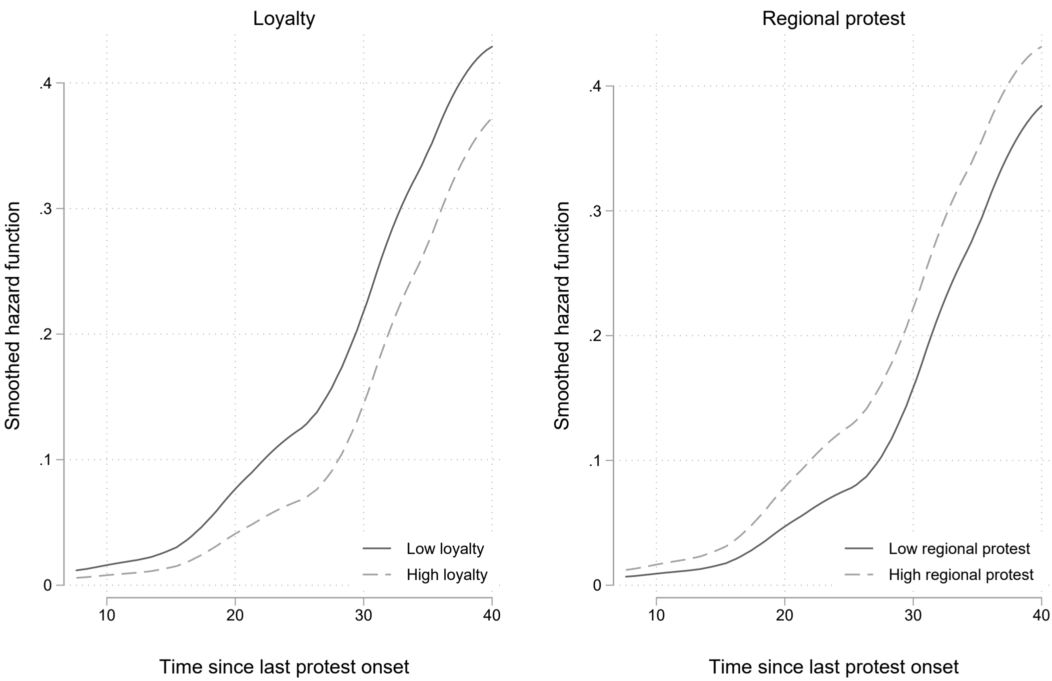

Figure B-8: Drop regions, one at a time

Figure B-8 reports the substantive effects for two variables: region protest and security personalization. The horizontal axis measure time since last protest onset and the vertical axis measures the smoothed hazard rate, which is the probability of protest onset at a given time conditional on its not having occurred before that time. The left panel shows that increasing loyalty from low to high levels lowers the hazard function.[59] For example, at time 20 this shift lowers the hazard rate from roughly 8 percent to roughly 4 percent. The right panel shows that increasing regional protest increases the hazard function. Visually, the two sets of curves suggest similar substantive effects – only in opposite directions. In short, employing a Cox model yields roughly the same results are the RE logit estimator.

[59]Both plots show a change from the 10th percentile of the distribution to the 90th percentile.

IV-2SLS models

This section addresses potential endogeneity, using both leader fixed effects estimators and IV-2SLS estimators that treats security force personalization as an endogenous variable. Endogeneity can arise from numerous sources, included mis-measurement, (unobserved) selection into treatment that is correlated with the outcome, and reverse causation. The measurement of the items that comprise the security personalization variable are coded for information observed on January 1 of each calendar year, which means that the event to which the information corresponds occurred in the prior calendar year, effectively lagging the variable by one year. This mitigates reverse causation because the information in the treatment variable chronologically precedes the realization of the outcome.

Unobserved selection into treatment is a greater threat to inference in this application because there may be some unobserved characteristic of the leader that prompts selection into personalization and also deters protest. Reputation or individual willingness to employ repression, for example, could make personalization easier and deter protest. We address possible endogeneity from (unobserved) leader-specific selection by testing estimators with leader-effects rather than regime-case-effects. When we test an RE-logit estimator (varying the intercept by leader) the estimated marginal effect is -1.6 percent – similar to that reported in the main text. In an FE-logit estimator (with leader-effects), the estimate is much larger (-8.4 percent), but the size estimate is biased because marginal effects do not draw inference from leaders with no protest in the FE-logit (Cook, Hays, and Franzese 2018).[60] Given these results from leader-effects models, it is unlikely that leader-specific selection upwardly biases estimates.

Next we address selection based on unobserved, potentially time-varying factors that induce treatment and may be related to protest by testing IV-2SLS estimators. Geddes, Wright, and Frantz (2018) argue that personalization arises when leaders have a strong bargaining position relative to their support coalition, which is more likely when that coalition is divided, as opposed to being unified. They posit that regimes that come to power with an inherited political party – one that chooses the leader rather than being created by the leader as an instrument to propel him to power – and regimes that come to power with a divided military are more likely to personalize. Both features of regimes – inherited party and divided military – are observed, in their framework, using information from prior to the regime seizing power. Thus this information is purely cross-sectional and we would like to identify plausibly exogenous information with some variation over time.

However, we can still use the theoretical approach in Geddes, Wright, and Frantz (2018) to identify within-regime information related to the divisiveness of the military. We therefore posit that officer rank – if the leader was in the military prior to becoming regime leader – is a plausibly exogenous proxy for a divided support coalition, in particular a divided military. Leaders who were junior officers prior to being leader must have sidelined more senior officers to ascend to power, which means the military must, at some level, have been divided between a junior officer who becomes the leader and at least some senior officers who are passed over. In unified militaries, in contrast, junior officers follow senior officers' orders. We thus use officer rank (prior to becoming leader), which is coded at five levels, with 0 being not in the military and four ordinal military ranks. Because we want to identify a prior division within the military, we also include the square of this term, allowing for the possibility that more senior officers have a lower personalization than junior officers (i.e., officer rank may have a non-linear 'encouragement'

[60]An FE-LPM with leader-specific time trends yields a significant estimate of -2.8 percent.

effect on personalization). We interpret this information as isolating the leader's bargaining power over the military.

The identifying information is therefore whether the leader faces a more or less unified military, based on information observed prior to the leader ascending to power. This information is not necessarily exogenous to unobserved factors that influence selection into leader but we addressed this selection issue earlier with leader fixed effects models. Instead, the identifying information in military rank is treated as plausibly exogenous to *unobserved, potentially time-varying, strategic behavior of the leader once in power* that vary across the leader's tenure and that may influence both selection into personalization attempts and protest onset.

TABLE B-3: Security personalization and NVC onset, IV-2SLS

	Random effects (1-2)		Fixed effects (3-5)		
					Lewbel instruments
	OLS	2SLS	OLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)
Security personalization	-0.0106*	-0.0289*	-0.0105	-0.0656*	-0.0257*
	(0.0034)	(0.0081)	(0.0063)	(0.035)	(0.0065)
Weak-ID F-stat				5.5	11.0

Dependent variable is protest campaign onset; all specifications adjust for leader tenure, time since last onset, and region onsets. NxT=4486; 1946–2010. Cluster-robust standard errors in parenthesis. * $p < 0.05$.

We test two IV-2SLS estimators that account for regime-case-level unit heterogeneity: a random intercept (RE) and a fixed effects (FE) estimator. Given the outside instrument and these estimators, the identifying information reduces to between-leader variation within autocratic regimes in the extent to which they face a unified or divided military.[61]

Table B-3 reports the results. The first two columns report RE models and the latter three FE models. The first column reports an OLS-RE model using the baseline specification: the estimate for security personalization is negative (-0.0106) and statistically significant at the 0.05 level. The next column reports the 2SLS-RE model and the estimated effect is negative and significant but much larger in (absolute) size, indicating that if the instrument is exogenous, the OLS estimator yield downwardly biased estimates of interest.[62]

The third column reports results from an OLS-FE estimator (recall that the cross-section unit is a regime-case). The estimate is almost identical to the RE estimate in column 1, but the variance estimate is larger, such that the coefficient is only significant at the 0.10 level. The fourth column reports results from a 2SLS-FE estimator with the excluded instrument set.[63] The estimate of interest is negative, large in absolute size, and significant . Note here that the instrument is relatively weak, with weak-id F-statistic of only 5.5, well below the critical value of 10.9. This indicates a weak instrument, which can lead to potentially unstable and inefficient estimates. To address this, we add Lewbel instruments, which are internal instruments based

[61]Because we test fixed effects estimators that rely on within-panel identifying information we drop all singleton cases.

[62]We use Baltagi's EC2SLS random-effects estimator Baltagi and Li (1992), which is more efficient than the G2SLS estimator (Baltagi and Liu 2009). A Wald-test of joint significance of the two external instruments in a first-stage equation has a p-value <0.04, providing little evidence of a weak instruments.

[63]Columns 4 and 5 report results from GMM-2SLS estimators.

on heteroskedasticity in the first stage equation (Lewbel 2000; Baum and Schaffer 2018). This approach yields a smaller estimate than that in column (4), with a much smaller error estimates such that the point estimate is statistically significant at the 0.01 level. The weak-id F-statistic of 11 is greater than the weak-id value – indicating a much stronger instrument when adding the Lewbel instruments.

The take-away from the 2SLS tests is that if we believe unobserved time-varying strategic behavior on the part of the leader induces both selection into personalization and influences protest, then our estimates suggests that naive models that do not account for endogeneity from unobserved selection might be biased towards zero; that is, the estimates reported throughout may be too conservative.

Cross-validation

This section reports results from two cross-validation exercises. One issue that arises in predictive models versus causal models that address confounding is model specification. Some cross-section variables, such as GDP per capita or urban population size, may be good predictors of campaign onset but may not be ideal variables for causal models using observational data because they are post-treatment. Further, predictive models in the vein of Chenoweth and Ulfelder (2017) largely eschew fixed effects estimators, which can be helpful in causal inference models with observational data. Indeed, our inference strategy for modeling campaign onset relies on directly modeling cross-section heterogeneity and thus drawing inference from within-variation.

In light of these issues, we conduct two cross-validation exercises and report them here. First, we use a k-fold cross-validation process with a linear probability model and regime-case fixed effects. This link function and modeling approach directly isolates within-dictatorship variation in the treatment variable. We choose the LPM because, as Timoneda (2021, 3) notes, "the LPMFE produces predicted probabilities much closer to the observed probability for a majority of the distribution" when the number of positive outcomes is less than 25 percent. The LPME not only easily accommodates fixed effects (unlike a non-linear link such as logit) but also yields realistic predicted probabilities. With this test, we employ the RMSE as metric for gauging predictive accuracy. We report the percent change in the RMSE when adding a variable to the baseline specification, which tells us the extent to which the added variable *reduces* the model error (i.e., *improves* the predictive power of the model).

The second approach is the one utilized by Chenoweth and Ulfelder (C-U): a k-fold cross-validation using a logit link function. This approach eschews modeling fixed effects and thus does *not* isolate the predictive accuracy using within variation in the treatment variables. Instead, this approach leverages the combined between- and within-variation in treatment variables to gauge predictive power. Similar to C-U, we report the change in the AUC for the out-of-sample predicted probabilities when adding a variable to the baseline specification.

Conceptually, the LMPFE approach tells us whether the added variable helps predict *when* protests are likely to start because this approach isolates within-dictatorship variation. In contrast, the logit cross-validation exercise can tell us whether the added variable improves prediction of both where and when protests are likely to start. However, this latter approach cannot tell us the relative weight of the *where* and the *when*. If structural variables that largely capture between-variation substantially improve predictive accuracy, this may simply result from added cross-sectional information and thus only helps us understand *where* protests are likely to emerge.

For each approach we begin with the baseline specification used throughout the analysis, with the following covariates: time since last event (i.e. duration dependence); leader tenure; and region protests. We then add one of 41 variables to this specification – including the main treatment variable of interest, *Security personalism* – and report the extent to which adding each of these variables (one at a time) to the specification changes the value of the predictive metric, either the change in the RMSE relative to the baseline (*decreases* mean more accurate) or the change in the AUC relative to the baseline (*increases* mean more accurate).[64] The reported estimates then tell us how well *Security personalism* fares in increasing (or decreasing) model accuracy, relative to the baseline specification, compared to the other 40 added variables.

[64]Because some added covariates have missing data, we estimate a baseline RMSE or AUC for the sample of observations without missing data on any one added covariate. For an added variable with no missing data, such as *Security personalism* or *Election year*, the baseline RMSEs and AUCs are the same. For added variables with missing data, such as *Trade*, the baseline RMSE and AUC is slightly different because it uses a slightly different set of observations.

Figure B-9 reports the first analysis (LPMFE), which shows that only a handful of added variables actually lower the RMSE. The two variables that improve accuracy (i.e., lower the RMSE) the most are *Elections* and *Security personalism*. Most added variables – including nearly all structural variables that move slowly over time such as population and GDPpc – actually decrease model accuracy (i.e. increase the RMSE). This latter finding shouldn't be surprising because the LPMFE model absorbs the between-variation that these structural variables capture. Indeed, only variables with non-trivial within-variation are likely to improve model accuracy in the first place. And the tests we conduct indicate that elections and security personalism – two variables with some within-variation – improve accuracy, though the improvements are substantively small. Even though small, these improvements in model accuracy by definition mean that security personalism is one of the few variables that improves prediction of *when* mass protest campaigns are likely to begin. This analysis also tells us that most added variables actually decrease predictive accuracy of *when* protests are likely to emerge.

The C-U approach to assessing predictive accuracy, a k-fold cross-validation with a logit model, leverages both within- and between-information. Here we should expect more predictive power from structural variables that capture mostly cross-section variation. The results reported in Figure B-10 indicate that most added variables improve accuracy (i.e., increase the AUC above the baseline AUC). The predictors that improve accuracy the most include structural variables such as *Population*, *Urban population*, and *GDP pc*. Prior pro-democracy mobilization, which captures low-level protest in prior years, improves the AUC the most; and *Military spending* and *Military size* also boost accuracy substantially. *Elections* and *Security personalism* are the next best variables for improving model accuracy.

Election events not only vary substantially across time but the timing of these events is also typically public knowledge and therefore can catalyze the mobilization of opponents in a civil resistance campaign. Prior mobilization, while an important predictor of mass protest mobilization, is also not particularly informative, insofar it is akin to a lagged outcome variable. *Military spending* and *personnel* capture government methods of potentially inducing military loyalty – similar to the theoretical mechanism we propose for interpreting the finding that *Security personalism* deters civil resistance campaigns from starting. That said, both military-related variables capture substantial variation in population size and may simply be proxies for population.[65] In short, even though security personalism is not the best predictor, it ranks among the best. And aside from standard structural variables such as GDPpc and population size, security personalism and military-related variables are the best predictors.

To sum, security personalism is one of two variables in this analysis that improves model accuracy when isolating within-variation. And it is also among the best predictors in a cross-validation test that combines cross-section and within variation.

[65]If we include population in the baseline specification, adding military spending increases the average out-of-bag AUC 1 percent and adding military personnel decreases the average out-of-bag AUC. Omitting population from the baseline, these changes are both slightly more than 6 percent. This suggests that the military variables only improve predictive accuracy because they pick up cross-section variation in country size.

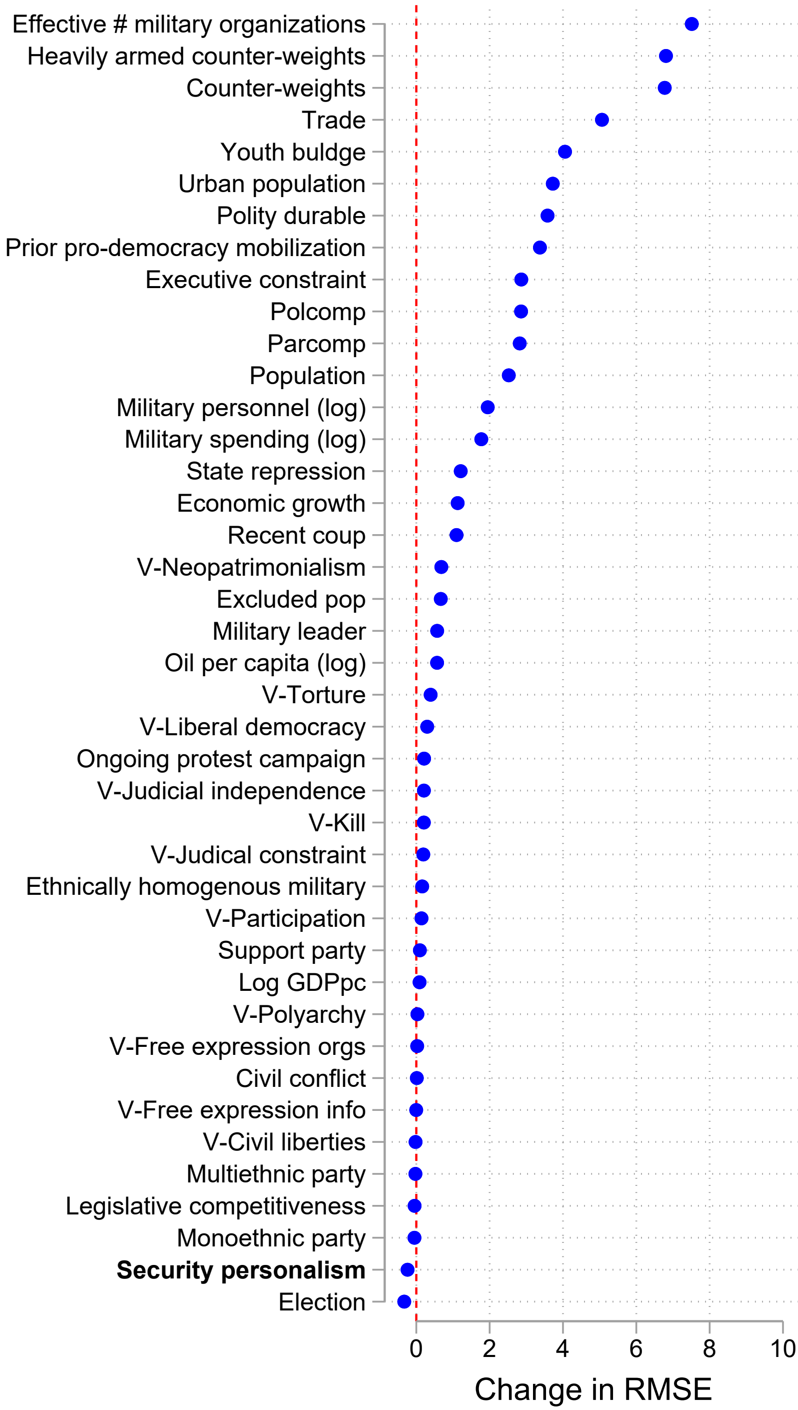

Figure B-9: LPMFE cross-validation, change in average out-of-bag RMSE

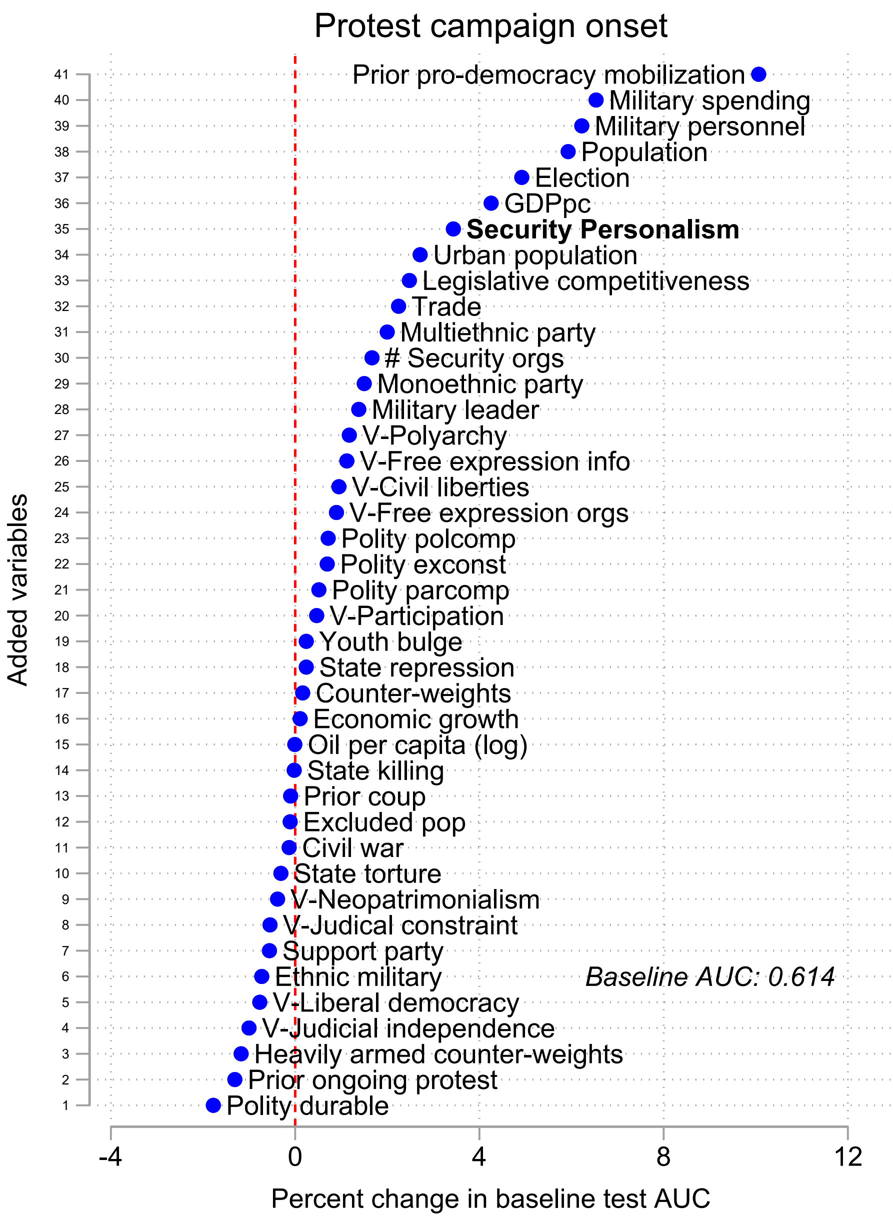

Figure B-10: Logit cross-validation, change in average out-of-bag AUC

Appendix C: Additional results for repression

Additional potential confounders

Figure C-1 shows the estimate of $\beta_{SecurityPers.}$ for the repression model in the main text (Table 1, column 2, depicted with the blue line at roughly 0.05) but with additional potential confounders, each added to the specification separately. The figure shows that the main result is fairly robust.

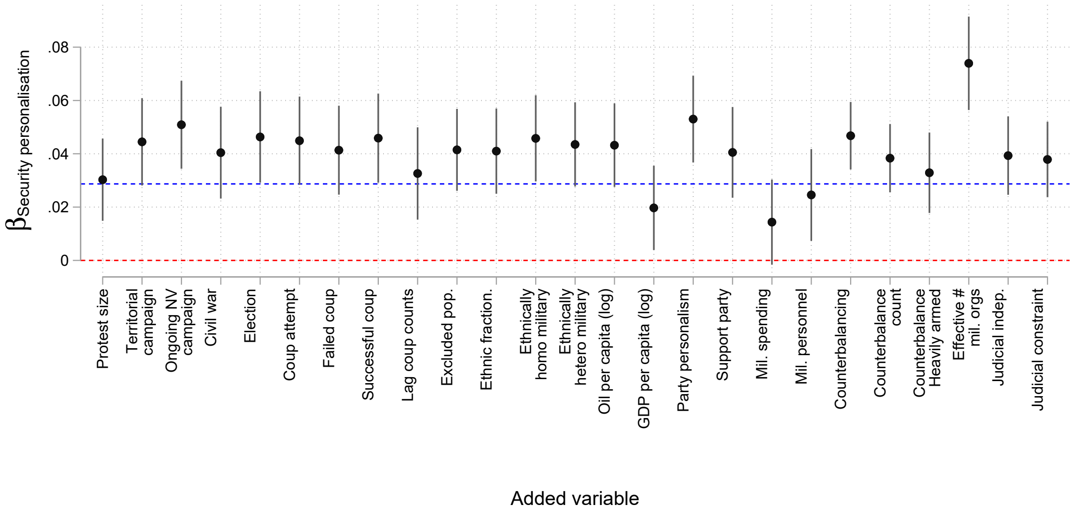

Figure C-1: Adding potential confounders to the repression model

Changes to the lag structure in repression models

This appendix examines the robustness of the finding that repression increases more during protest campaigns in regimes with a personalized security apparatus than in regimes that lack such loyalty mechanisms. All of the empirical approaches condition the estimates of personalism on lagged repression because regimes with high security personalization in onset years are likely to have higher levels of state-led repression prior to onset years than regimes with low security personalization. Figure C-2 shows the average level of repression *prior* to protest onset in regimes with high security personalization (top one-half of the in-sample distribution) and low personalization (bottom half). The first three plots show average levels 1 to 3 years prior to the onset; and the final plot shows the lagged moving average of these three years. In all plots, the average level of repression prior to onset is higher in regimes with high security personalization, although these differences are not statistically significant. Nonetheless, we conditional all estimates on the observed prior levels of state-led repression.

Next, we demonstrate that how we condition on past repression does not appreciably alter the reported results. Further, failing to account for prior repression likely biases the estimates upwards. Figure C-3 shows the results of a series of kernel regression models with the same specification as reported in the main text (Table 1, column 2, depicted with the blue line at roughly 0.05) – except we change the lagged repression variable. The first three estimates show results when lagging repression 1, 2, and 3 years prior to protest onset. The fourth estimate is from a specification that includes the average level of observed repression for all years the regime was in power. This is similar to a fixed effects "within" design that conditions estimates of interest on the unit means. The fifth estimate pursues a similar strategy but calculates the unit average of repression for all non-protest campaign years to focus the comparison on onset years relative to non-protest years. The final column reports estimates without conditioning on past repression or average regime-repression.

The results convey three pieces of information. First, all results for security personalization are positive and significant, suggesting robustness of the key finding. Second, omitting lagged repression yields the largest estimate for security personalization. This should not be surprising given the patterns shown in Figure C-2. Finally, conditioning on unit means yields similarly sized estimates as omitting lagged repression measures. This suggests that a traditional fixed effects estimator might yield upwardly biased estimates, particularly if repression just prior to an onset is typically higher than average regime-level repression, and more so for regimes with highly personalized security apparatus. (A parallel trends assumption is unlikely to be met.)

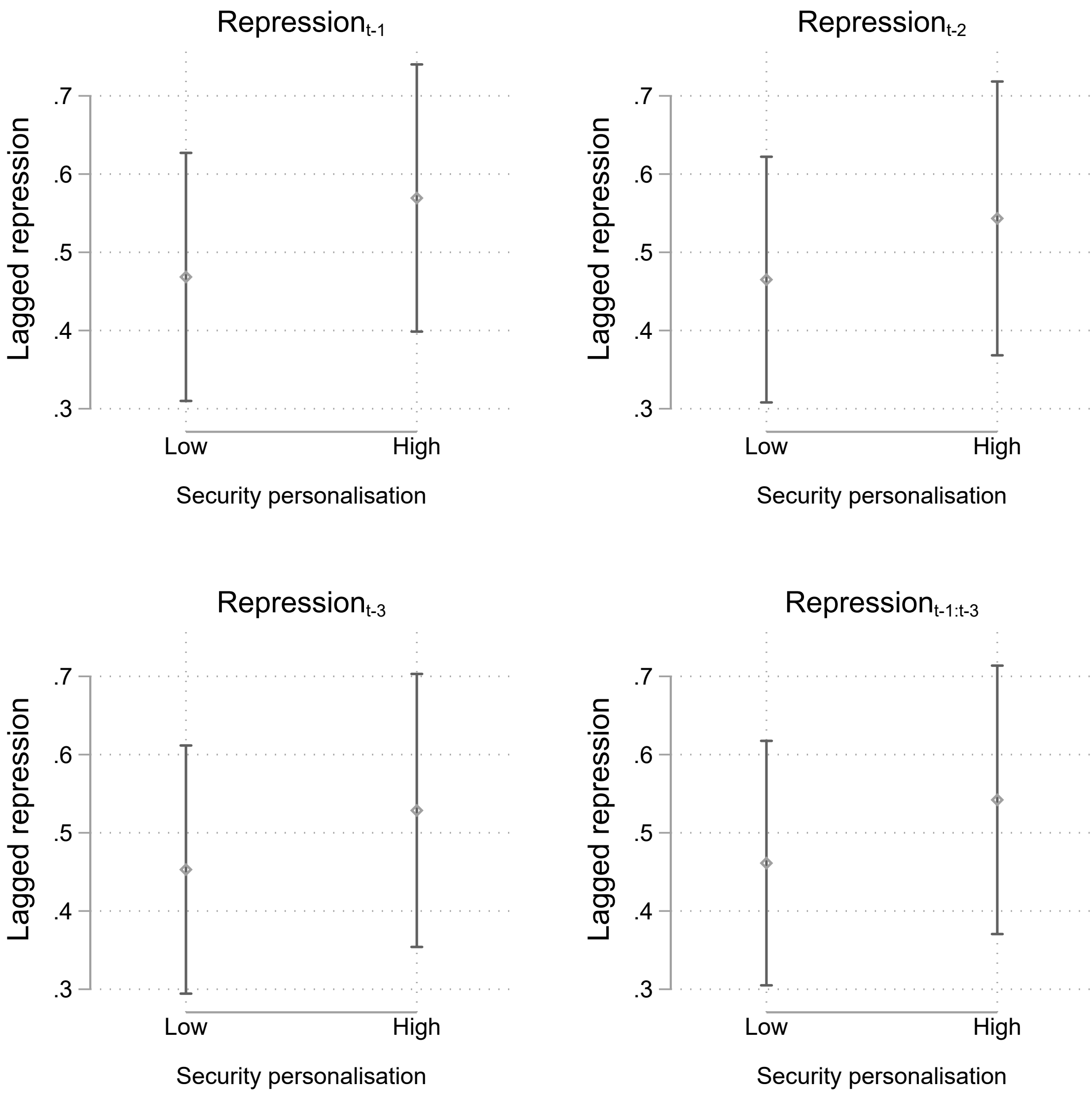

Figure C-2: Repression prior to protest onset, by personalism

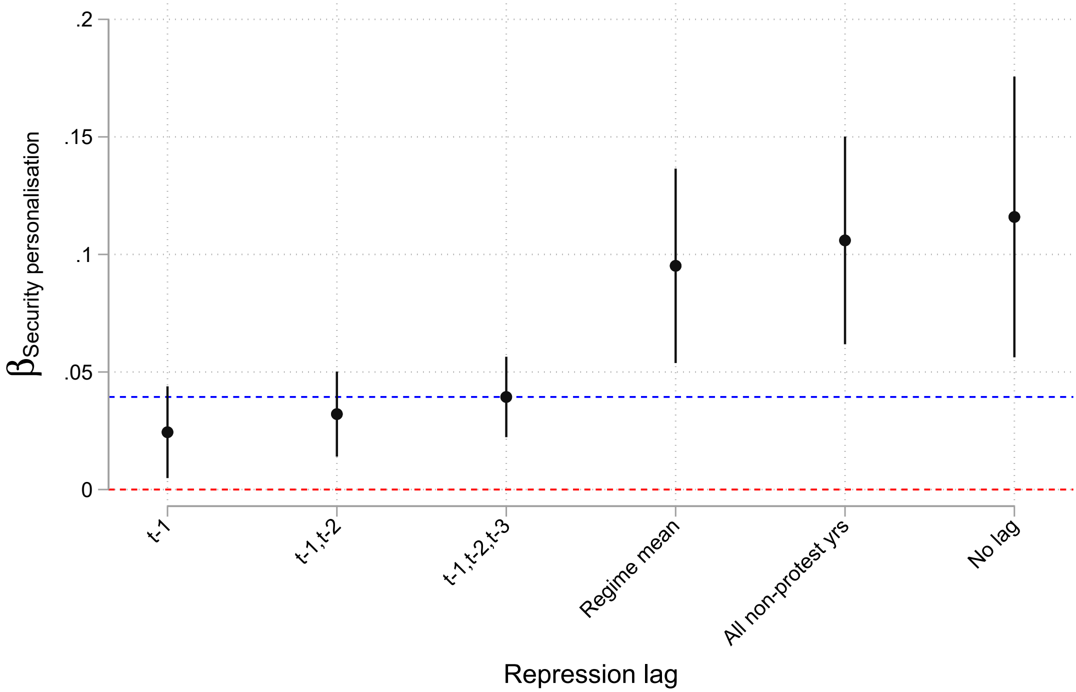

Figure C-3: Lags in the repression model

Dynamic repression models

Next, we test a series of models in differences to estimate the dynamic relationship between protest campaigns and repression. The estimating equation is the following, with a fixed parameters for calendar time period (ϕ_t) and leader time in power (τ_s) to account for common time trends:

$$R_{i,t} - R_{i,t-1} = \beta(NVC_{i,t} - NVC_{i,t-1}) + \psi(SFP_{i,t} - SFP_{i,t-1}) + \gamma(X_{i,t} - X_{i,t-1}) + \phi_t + \tau_s + \epsilon_{i,t} \quad (8)$$

First, we estimate the equation without an interaction between the differenced measures of NVC and SFP; then we estimate the equation with the interaction to assess whether the effect of protest campaigns changes as security personalism increases. The first two columns of Table C-1 show these results. The second two columns add control variables to account for potential (dynamic) confounding effects of elections, civil war, and coups. And the last two columns simply split the same into high/low security personalism regimes.

TABLE C-1: NVC and repression, difference models

	(1)	(2)	(3)	(4)	High SFP (5)	Low SFP (6)
NVC	0.0229*	0.0268*	0.0199	0.0240*	0.0411*	0.0089
	(0.010)	(0.011)	(0.011)	(0.011)	(0.018)	(0.012)
Security personalization	0.0097	0.0062	0.0120	0.0084		
	(0.007)	(0.007)	(0.007)	(0.008)		
NVC × Security pers.		0.0210		0.0213*		
		(0.011)		(0.011)		
Failed coup			0.0222*	0.0223*	0.0147	0.0275
			(0.010)	(0.010)	(0.011)	(0.019)
Successful coup			0.0457*	0.0456*	0.0535*	0.0407*
			(0.014)	(0.014)	(0.021)	(0.020)
Election year			-0.0019	-0.0020	0.0030	-0.0084
			(0.007)	(0.007)	(0.014)	(0.008)
Civil war			0.0275	0.0282	0.0369	0.0155
			(0.017)	(0.017)	(0.021)	(0.028)
Time since last onset (log)			0.0005	0.0003	0.0082	-0.0056
			(0.004)	(0.004)	(0.007)	(0.006)
$\beta_{NVC} + \beta_{NVC \times SFP} \times LowSFP$		-0.007		-0.010		
		(0.015)		(0.016)		
$\beta_{NVC} + \beta_{NVC \times SFP} \times HighSFP$		0.054*		0.052*		
		(0.021)		(0.021)		
Period effects	✓	✓	✓	✓	✓	✓

Dependent variable is differenced repression. N × T = 4135. 251 regimes; cluster robust standard errors in parentheses; 1946–2010. * $p < .05$.

The first column of Table C-1 shows that, on average, the short-term dynamic effect of protest campaign increases repression. The second column shows that security force personalization increases the size of this effect. The estimate of β_{NVC} at high levels of personalization (90th percentile) is over 0.05, while the same estimates at low personalism (10 percentile) are roughly zero. These results again suggest that protest onset increases short-term repression but only in highly personalized regimes.

Next, we fully embed the dynamic effect in an error-correction model that estimates short- and long-term effects of protest campaigns on repression simultaneously. The results, shown in Figure C-4,[66] indicate that in regimes with high levels of security personalization, protest campaigns have a large, positive, and statistically significant *short-tern* effect on repression. However, the long-run effect in highly personalized regimes is negligible, which means that regimes return to an equilibrium level of repression after initial protest onset. For regimes with low levels of security personalization, there is a negligible short-term effect, and although the long-run effect is positive, it is not statistically significant at conventional levels.

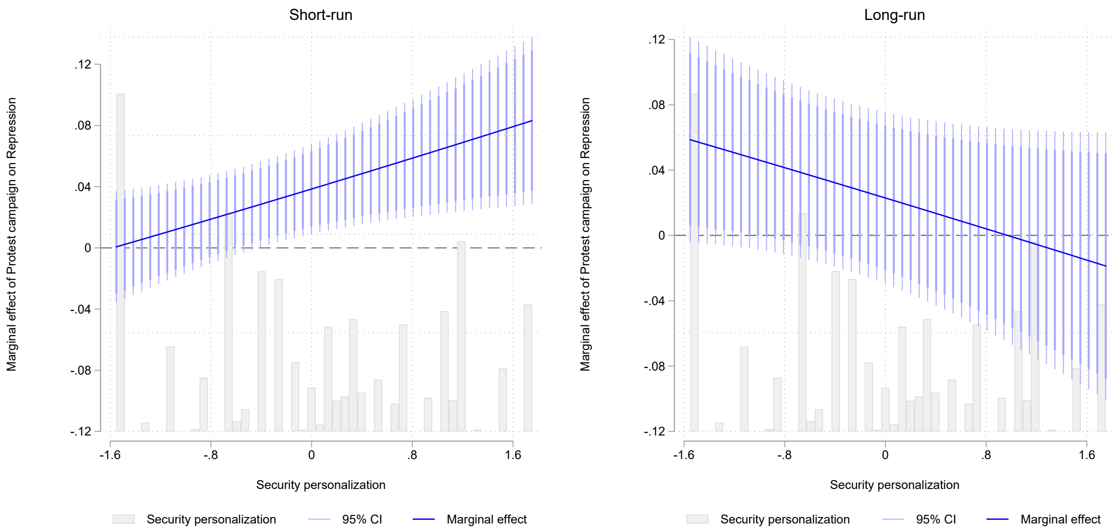

Figure C-4: Error-correction model results

[66]Plots of the interaction effect for short- and long-run effects calculating using kernel regression (Hainmueller, Mummolo, and Xu 2019).

An alternative measure of repression

Table C-2 shows the results for tests that use the Varieties of Democracy (VDEM) data on repression. Instead of using the Fariss (2014) measure of latent political repression, these tests use a linear index of two ratings – for political killings and political torture – of state-led violence. In dictatorships the Fariss' measure and the VDEM index are correlated at 0.45; within the sample of protest onset years analyzed in this table, the correlation is 0.41. One difference between the two measures relates to dynamic estimation of the latent value in the Fariss' approach; this causes the latent estimate for each country-year to be informed by the lagged and forward years of the estimate, which could bias estimates of repression that rely on sharp changes from year to year to be bias downwards. The VDEM data series are not estimated with a dynamic model, though it is derived from a latent model that combines information from multiple country experts.

TABLE C-2: Repression during protest campaigns: Varieties of Democracy data on repression

	Onset years only		All campaign years
	(1)	(2)	(3)
Security personalization	0.0423+	0.0336*	0.0334*
	(0.022)	(0.017)	(0.015)
Population (log)		0.0341*	0.0322*
		(0.011)	(0.010)
Region NVC onsets (log)		0.0059	-0.0031
		(0.014)	(0.012)
Leader tenure (log)		0.0182	0.0181
		(0.017)	(0.015)
Time since last onset		-0.0041	0.0172
		(0.015)	(0.011)
Campaign duration (log)			0.0108
			(0.021)
Lag repression	0.7689*	0.6803*	0.6589*
	(0.034)	(0.030)	(0.020)
Campaigns	182	182	182
Campaign-years	182	182	332

182 protest campaigns in 112 regimes in 81 countries, 1946–2010. Kernel least-squares estimator. Standard errors in parentheses; * $p < .05$.

Alternative measures of personalization

Table C-3 shows the results for tests that include various ways of measuring personalization. The first column reports results from a specification that includes a binary measure of personalist regime (with military junta, dominant party regime, and monarchy as the combined excluded category) (Geddes, Wright, and Frantz 2014). The estimate for this variable suggests no systematic difference between these regimes and other types of autocracies in the level of state-led repression in protest campaign onsets. The second column reports results when including a full measure of personalization that combines party and security items in the latent estimate (Geddes, Wright, and Frantz 2018; Wright 2021). The estimate is positive and statistically significant. The third column introduces a measure of party personalization only; the estimate is negative but not statistically significant. The fourth column is a specification that includes security personalization, which is the same specification and results as reported in column 2, Table 1 in the main text. Again, the estimate is positive and statistically significant. The final column reports results from a specification that includes both measures of personalization – party and security. Only the estimate of security personalization is positive and significant. These results suggest that security personalization – not party personalization – account for the main repression findings.

TABLE C-3: Repression during protest campaign onsets: alternative measures of personalization

	(1)	(2)	(3)	(4)	(5)
Personalist regime	0.0264				
	(0.025)				
personalization		0.0437*			
		(0.009)			
Party personalization			0.0080		-0.0079
			(0.009)		(0.008)
Security personalization				0.0394*	0.0530*
				(0.009)	(0.008)

182 protest campaigns in 112 regimes in 81 countries, 1946–2010. Covariates not reported: three lags of repression, population, time since last onset, leader tenure, region NVC onsets. Time period dummies included but not reported. Kernel least-squares estimator. Standard errors in parentheses; * $p < .05$.

Measuring repression using the NAVCO data

Theoretically, we posit that security personalism should increase state-led repression during protest campaigns. Indeed, one might argue that we should only examine repression targeting the protest campaign itself – and not a more general measure of protest, as we have done in the manuscript. There are three issues that arise when using the NAVCO data of repression targeting the campaign instead of a more broad-based measure of state-led repression.

First, the NAVCO data on repression are only available *during* the campaign and not prior to the campaign. This means we cannot compare repression during the campaign to prior repression; thus inference relies mostly on cross-section comparisons which we believe are more likely to be biased from confounding than our preferred design that compares repression during campaigns to prior levels of state-led repression.

Second, there is very little variation in the NAVCO measure of state-repression during campaigns: 91 percent of campaign-years are coded as state-repression. That is, the vast majority of campaigns are targeted with state-led repression. We are thus not confident that this measure of repression will yield meaningful inferences. Finally, the repression variable from NAVCO is binary, which does not distinguish between levels of state-repression targeting the campaign.

For these reasons, we find using the a more general measure of repression – which is still likely to capture increases in state-repression targeting nonviolent campaigns – more convincing. Nonetheless, here we test, using kernel estimator, whether security personalism explains repression during campaigns, employing the NAVCO data. We test two samples and three specifications for each sample. To start, we examine only the first year of each protest campaign, testing three specifications: bivariate; conditional on the following covariates: region protest, population size, campaign size, and tenure; conditional on the covariates plus lagged levels of broad-based state-led repression. These latter variables are potentially important because they capture the cross-section variation in how repressive different regimes are. The second sample is for all years, and we test each of the aforementioned specifications and only add campaign duration as a covariate in each.

In all six tests, we find that security personalism increases the incidence of any repression targeting the campaign by 3.6 percent in campaign onset years and by roughly 5 percent in any campaign year. Figure C-5 shows the results in two ways. The left plot shows the average marginal effect of security personalism on the incidence of state-led repression in the first year of each campaign througout the past six decades. The estimated effect of security personalism is substantially larger since 1990 than before. The right plot shows the marginal effect of security personalism across campaign duration, using the covariate specification for the tests using all campaign years. The effect of personalism on increasing repression is greatest in the first year of the campaign.

In short, consistent with the tests reported in the main text and with the theory, more personalized security forces are associated with more repression during campaigns. While we account for lagged levels of state-led repression using the Fariss measure in some specifications, we still believe these tests rely on relatively narrow cross-section variation and should be interpreted with appropriate caution.

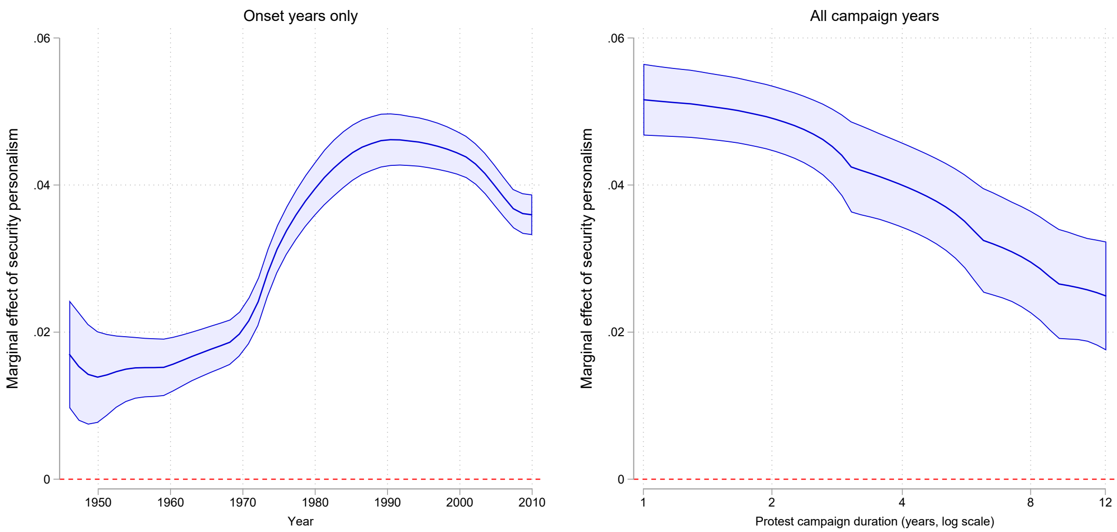

Figure C-5: Security personalism increases the incidence of repression targeting the protest campaign

Appendix D: Additional results for Democratization

Regime collapse placebo tests

This section reports tests of regime collapse where we treat all democratic transitions as right censored. This means that the outcome is a binary indicator of regime collapse than ends in either state collapse or a transition to a new autocratic regime. Examples include: the Iranian Revolution in 1979 where the theocratic regime replaced the Pahlavi monarchy and the 1997 collapse of the Mobutu regime in the former Zaire, when rebels led by Laurent Kabila took control of the capital city, instituting a new autocratic regime led by Kabila and, later, his son. While estimates for *Security personalism* in both models are positive, the marginal effects are small and statistically insignificant: 0.9 percent in the FE-LPM (-2.3 percent is the estimate for democratic transitions); and 0.9 percent in the CRE probit (-2.2 percent is the estimate for democratic transitions). Note that a postive estimate indicates that security personalism could *destabilize* dictatorships by increasing the risk of collapsing via mechanisms other than transitions to democracy.

TABLE D-1: Regime collapse, not democratization

	FE-LMP (1)	CRE probit (2)
Security personalization	0.0085	0.2563
	(0.006)	(0.141)
Log regime-case duration	0.0385*	2.8719*
	(0.005)	(1.156)
Region NVC onsets (log)	-0.0024	-0.0951
	(0.003)	(0.067)
Country effects	✓	
Year effects	✓	
Period effects		✓
Within transform		✓

N × T=4,535; 256 regimes in 117 countries, 1946–2010. Cluster-robust standard errors in parentheses; * $p < .05$.

Adjusting for additional covariates

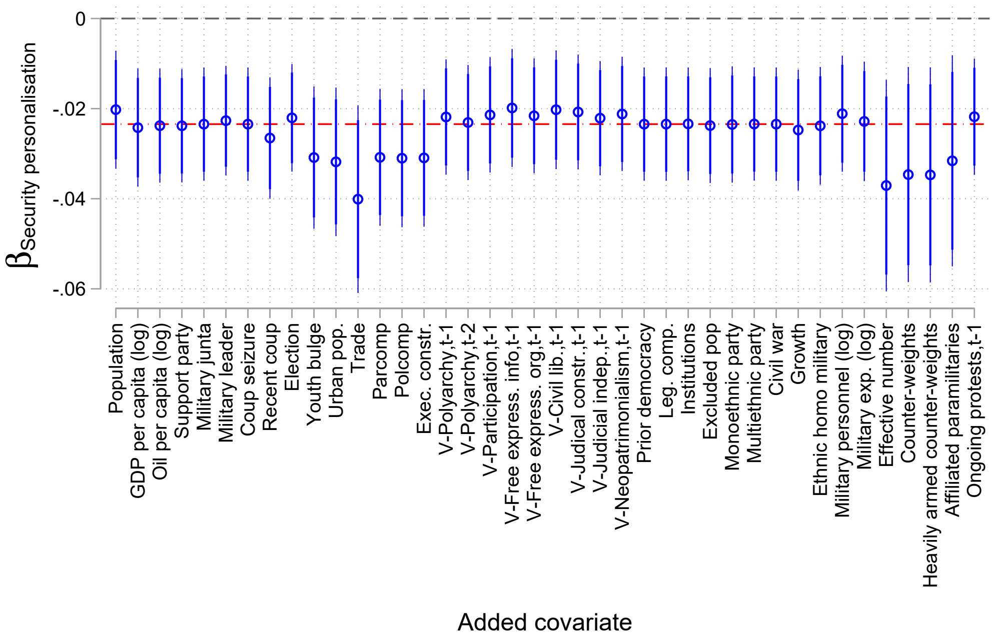

Figure D-1: Security personalization and democratic transition, additional covariates

Addressing time trends

Figure D-2 shows the results for democratic transitions when adjusting how we model the common time trend. Some of these models are reported in the main text.

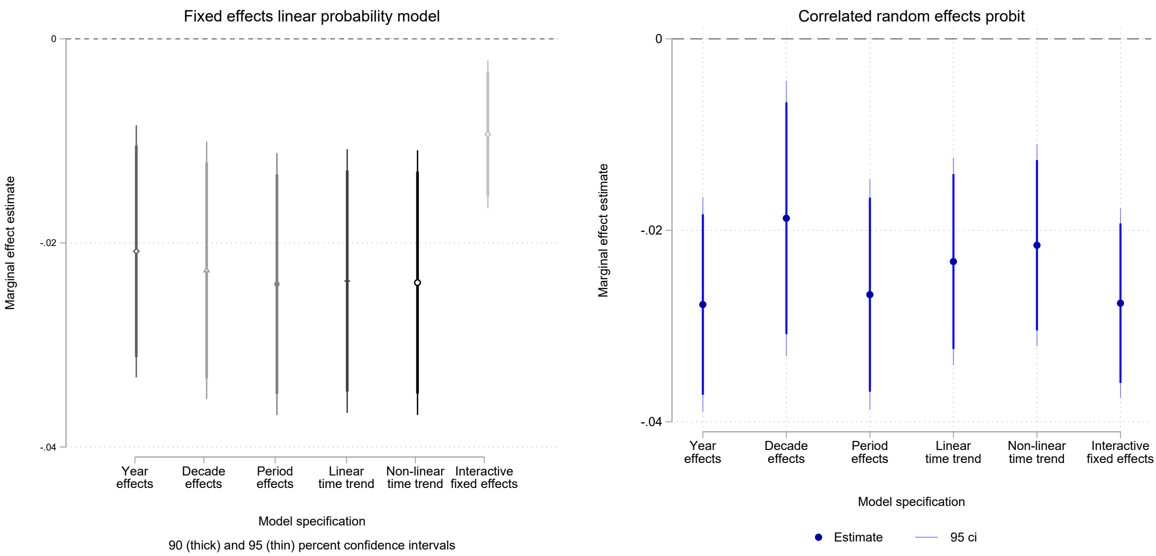

Figure D-2: Security personalization and democratic transition, alternative time specifications

Alternative FE-LPMs

Figure D-3 shows the results for democratic transitions when adjusting how we model the panel unit. Regime-cases – which are the panel units used throughout – are nested within countries, while individual leaders are nested with regime-case panel units. We show results for both RE and FE estimators with all three types of panel units: country, regime-case, and leader.

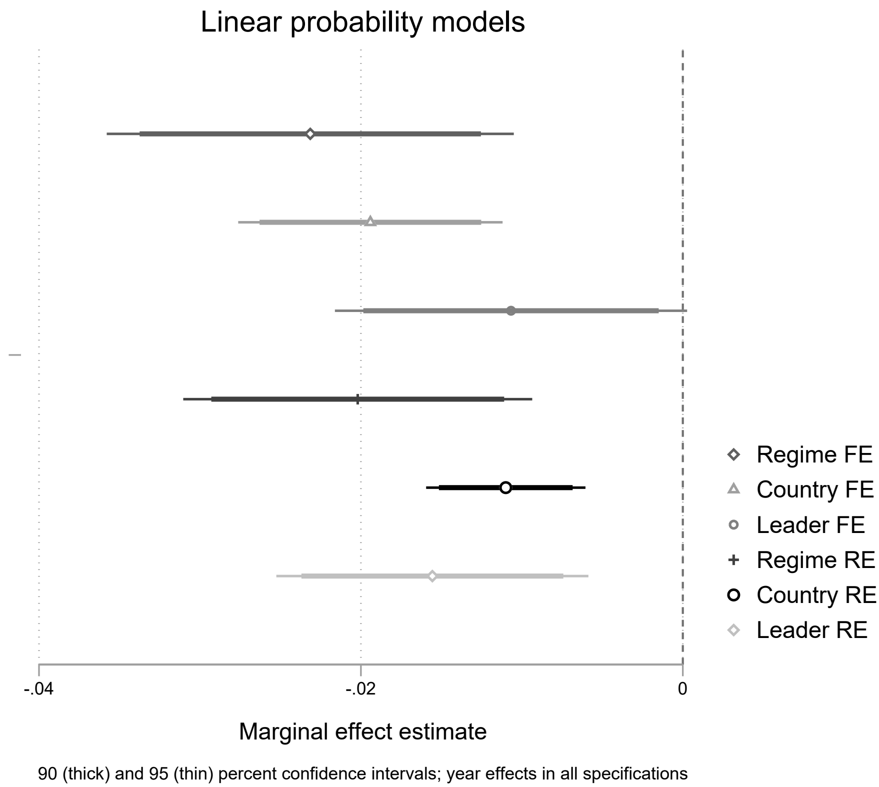

Figure D-3: Security personalization and democratic transition, alternative LPMs

Cox models

Figure D-3 shows the results for democratic transitions when using binary treatment variable and Kaplan-Meier survival plot. The treated cases show a lower survival probability for all duration years.

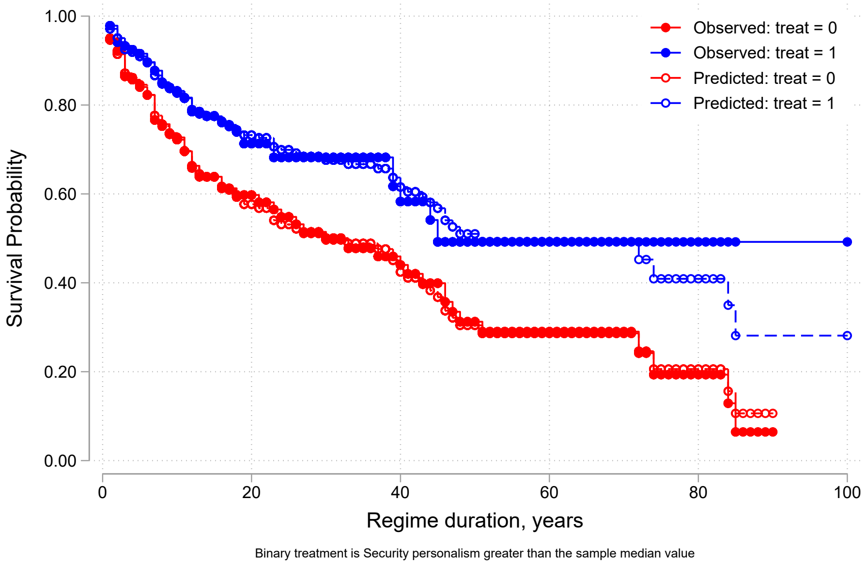

Figure D-4: Kaplan-Meier survival curves, by treatment status

Figure D-2 shows the results of Cox duration models that more flexibly model regime duration. The first column reports results with adjustment of panel-level heterogeneity. The second uses 'within' transformations of the explanatory variables to model time-varying changes in these factors (similar to a CRE probit). Column (3) reports estimates with shared frailties (similar to panel random effects); and Columns (4)-(6) show results from stratified Cox models (similar to FE models).

TABLE D-2: Cox duration models

	(1)	(2)	(3)	(4)	(5)	(6)
Security personalization	-0.4715*	-0.9511*	-0.4810*	-0.5176*	-0.8099*	-0.8523*
	(0.106)	(0.203)	(0.109)	(0.251)	(0.332)	(0.414)
Region NVC onsets (log)	0.2290*	0.0655	-0.5218*	-0.6780*	-0.7687*	-0.8874*
	(0.097)	(0.122)	(0.240)	(0.199)	(0.221)	(0.297)
Log GDP					-1.4205	-1.1945
					(0.770)	(0.770)
Population (log)					-4.7921	-5.6663
					(3.047)	(3.031)
Oil rents (log)					-4.9069	-4.8692*
					(2.755)	(2.376)
Democracy$_{t-1}$						6.5982
						(6.459)
Democracy$_{t-2}$						-7.9932
						(5.485)
N $\times$ T	4559	4559	4559	4559	4485	4473
Regimes	280	280		280	277	277
Period effects	✓	✓	✓	✓	✓	✓
Within transform		✓				
Regime-case frailties			✓			
Stratified				✓	✓	✓

Dependent variable is democratic transition. Cluster robust standard errors in parentheses; 1946–2010. * $p < .05$.

IV-2SLS tests for democratic transition

This section reports results from IV-2SLS tests for democratic transitions. As before, we use officer rank as an excluded instrument under the assumption that this characteristic of dictators determines selection into security personalism but does not influence democratic transition via other mechanisms that are pre-treatment to security personalism. Thus officer rank can still influence outcomes such as state-led repression, which may reduce the chances of democratization, that are themselves the outcome of security personalism (i.e., post-treatment from the perspective of security personalism). For example, Frantz et al. (2020) show that personalism increases state-led repression in all years and our analysis in this paper shows that security personalism influences state-led repression during nonviolent resistance campaigns. Thus officer rank can shape state-led repression provided it only does so via security personalism.

As throughout we estimate 2-way FE OLS models, but with excluded instruments to address endogeneity from unobserved, time-varying strategic behavior on the part of leaders. Column (1) of Table D-3 reports the IV estimate: -0.092. This estimate is much stronger than most reported in the main text but, because the excluded instrument is relatively weak the error estimate is also large and the coefficient estimate is not significant. The second column adds lagged democracy levels to the specification, yielding a strong negative result (-0.065). But, again, the instrument is relatively weak. The last two column report results when adding Lewbel instruments, which produce weaker results (1 percent reduction in democratic transition risk) but because the instrument is much stronger, these estimates are much more precise, reaching conventional statistical significance.

None of these results, however, cast doubt on the negative effect of security personalism on democratic transition. If anything, the upper bound of a reasonable estimate is about -1 percent; the lower bound could be as low as -9 percent.

TABLE D-3: IV-2SLS models of democratic transition

			Lewbel instruments	
	(1)	(2)	(3)	(4)
Security personalization	-0.0921	-0.0649	-0.0088*	-0.0139*
	(0.079)	(0.069)	(0.003)	(0.004)
Regime-case duration (log)	0.0543*	0.0489	0.0197*	0.0252*
	(0.026)	(0.027)	(0.004)	(0.006)
Region NVC onsets (log)	0.0042	0.0051	0.0011	0.0008
	(0.003)	(0.003)	(0.002)	(0.002)
Democracy_{t-1}		0.3820*		0.3356*
		(0.168)		(0.067)
Democracy_{t-2}		-0.0963		-0.0828
		(0.151)		(0.061)
N × T	4535	4004	4559	4020
Regimes	256	218	280	234
F-statistic	2.7	3.6	16.2	24.9
Stock-Yogo weak ID	19	19	11	11
Regime-case effects	✓	✓	✓	✓
Year effects	✓	✓	✓	✓

Dependent variable is democratic transition. 2-step GMM estimator. F-statistic is the Kleibergen-Paap rk Wald F statistic. Stock-Yogo weak ID is the test critical value for 10% maximal IV relative bias Cluster robust standard errors in parentheses; 1946–2010. * $p < .05$.

Appendix E: Security Force Defection

The Appendix examines whether security force personalism increases the risk of security force defection once protest campaigns have begun (i.e. conditional on observing protest campaign).

One potential cost of personalizing the security forces, and in our view the most important one during mass uprisings, is that personalization creates backlash within the military, as some officers lose power and prestige.[67] Case studies of mass uprisings in Asia (Lee 2014), Africa (Morency-Laflamme 2018), and the Arab Spring (Barany 2012) posit that military personalization prompts military defection.[68]

The dictator's order to repress realized protest may cause personalized security forces to splinter because the loyalty mechanisms in place create "winners" and "losers" within the security sector: dictators funnel greater privileges and resources to favored units, often outside the regular military command, comprised of security agents whose selection and promotion is based on loyalty. By ensuring the loyalty of preferred security agents ("winners"), the dictator may lose the loyalty of others ("losers"). Elites who have been sidelined under the personalist dictatorship thus feel less responsibility to the regime, and may seek an accommodation (pact) with the opposition leading the uprising in hopes that they may join the winning coalition should the uprising succeed (e.g. Lee 2014). That is, different organizations within the security apparatus may have different post-exit payoffs. The hand-picked troops of a special presidential guard, for example, will have a worse 'outside option' than the regular army. And those with less loyalty should have less to lose should the regime fall.

Dictators may try to mitigate such security force competition and produce a cleavage with the opposition by stacking the military with co-ethnics of the leader, thus tying security forces to the regime (Makara 2013). All else equal, greater social distance between security forces and the opposition makes it more difficult for the opposition to build and leverage channels of communication with security forces to court defections (Binnendijk 2009; Johnson 2017). As a result, ethnic minorities, especially those with territorial goals (Sutton, Butcher, and Svensson 2014), are less likely to initiate nonviolent uprisings than violent insurgencies in the first place (Thurber 2018).[69] Since mass uprisings are typically led by the ethnic majority, dictators can only occasionally play the ethnic card to limit defections. Thus while there may be some avenues for mitigating the risk of security force defection, the zero-sum logic of personalization of the security apparatus nonetheless creates internal divisions that raise the risk of security forces splitting when tasked with repressing nonviolent protest campaigns.

The NAVCO data distinguish between *state defections* and *security force defections*, where the former picks up major defections by civilian bureaucrats and the latter picks up major defections by the military or other security forces (Chenoweth and Lewis 2013). Research on revolution

[67]Security force personalization thus increases short-term coup risk (due to backlash) even if reducing long-term coup risk (Song 2022). While the opportunity to shape coercive forces tends to occur very early in a leader's tenure (Greitens 2016; Sudduth 2017), mass uprisings are more common later on; most of our empirical analyses account for leader duration.

[68]Lee (2014) analyzes four cases, two failed uprisings where there was a "low" degree of personalism and security forces defended the regime (e.g. China 1989 and Burma 2007) and two successful uprisings where there was a "high" degree of personalism and security forces defected from the regime (e.g. Philippines 1986 and Indonesia 1998). Morency-Laflamme (2018) analyzes two cases in Benin and Togo; Barany (2016) examines over half a dozen cases (Iran 1979, Burma 1988 and 2007, China and Eastern Europe 1989, and the Arab Spring in 2011). More general empirical analyses of counter-balancing – one element of security force personalization – provide mixed results. Dahl (2016) finds that counter-balancing increases the risk of security force defections, whereas Lutscher (2016) finds a U-shaped effect.

[69]We account for ethnic militaries in the empirical analysis. Adjusting for this factor produces *stronger* results than those reported below.

and civil resistance strongly points to security force defections as a leading cause of successful mass uprisings (e.g. Chorley 1953; Russell 1974; Katz 2004; Chenoweth and Stephan 2011). These defections come in many forms (Albrecht and Ohl 2016; Neu 2018), from the passive quartering or refusal to carry out orders to shoot by some (but not necessarily all) units or services (Pion-Berlin, Esparza, and Grisham 2014) to active participation of individual troops in protests (e.g. marching with or protecting protesters) to military commanders' leading regime change coups in support of civilian protests. The NEVER data generally follow the widely used NAVCO 2.0 coding (Sutton, Butcher, and Svensson 2014), but fill in missing observations in NAVCO coding.[70] Security force defection occurs in 39 percent of campaigns and 37 percent of campaign years.

Table 1 reports a series of tests using logit regression, though results remain using a kernel regression as well. The marginal effect of security force personalization is positive and statistically significant across all the models, from the barest bivariate model in column 1 to the fullest model in column 4. In the latter model, we adjust for the most prominent potential confounders discussed in the literature. The effect of these controls generally conforms with findings from prior literature. There is a positive and strongly significant effect of campaign size, or the (log) of campaign *membership*, confirming the insights of Chenoweth and Stephan (2011). Defections are less likely in more populous countries, repression increases the odds of defection, and *ethnically homogenous militaries* are less likely to defect.[71] This latter adjustment is important because we want to ensure that the defection effect results from security personalization and not the ethnic composition of the military, which is, as shown, a very strong predictor of *not* defecting.

[70]NAVCO 2.0 has missing data on security force defections for a third of campaign-years.
[71]This variable is from Geddes, Wright, and Frantz 2018.

TABLE 1: Military defection during protest campaigns

	(1)	(2)	(3)	(4)
Security personalization	0.493*	0.627*	0.534*	0.475*
	(0.192)	(0.222)	(0.240)	(0.230)
Leader tenure (log)		-0.252	-0.189	-0.123
		(0.195)	(0.201)	(0.204)
Population (log)			-0.534*	-0.589*
			(0.182)	(0.181)
Membership (log)			0.796*	0.782*
			(0.261)	(0.250)
Region NVC onsets (log)			0.014	-0.029
			(0.155)	(0.160)
Repression			0.422	0.729*
			(0.287)	(0.309)
Campaign duration (log)			-0.627*	-0.523
			(0.309)	(0.303)
Ethnically homogenous military				-1.654*
				(0.688)
(Intercept)	-0.629	-0.802	4.575*	6.182*
	(1.048)	(0.998)	(1.874)	(1.916)
N × T	316	316	311	311
# Campaigns	182	182	178	178

182 protest campaigns in 111 regimes in 81 countries, 1946–2010. All specifications include time period indicators (not reported). Logit regression with errors clustered by campaign; * $p < .05$.

Appendix F: Bibliography for Appendices

References

Albrecht, Holger, and Dorothy Ohl. 2016. "Exit, resistance, loyalty: military behavior during unrest in authoritarian regimes." *Perspectives on Politics* 14 (1): 38–52.

Baltagi, Badi H, Dong Li, et al. 2002. "Series Estimation of Partially Linear Panel Data Models with Fixed Effects." *Annals of Economics and Finance* 3 (1): 103–116.

Baltagi, Badi H, and Qi Li. 1992. "A note on the estimation of simultaneous equations with error components." *Econometric Theory* 8 (1): 113–119.

Baltagi, Badi H, and Long Liu. 2009. "A note on the application of EC2SLS and EC3SLS estimators in panel data models." *Statistics & Probability Letters* 79 (20): 2189–2192.

Barany, Zoltan. 2012. *The soldier and the changing state: building democratic armies in Africa, Asia, Europe, and the Americas.* Princeton, NJ: Princeton University Press.

———. 2016. *How armies respond to revolutions and why.* Princeton, NJ: Princeton University Press.

Baum, Christopher, and Mark Schaffer. 2018. "IVREG2H: Stata module to perform instrumental variables estimation using heteroskedasticity-based instruments."

Belkin, Aaron, and Evan Schofer. 2003. "Toward a Structural Understanding of Coup Risk." *Journal of Conflict Resolution* 47 (5): 594–620.

Binnendijk, Anika Locke. 2009. "Holding fire: Security force allegiance during nonviolent uprisings." PhD diss., Tufts University, August.

Chenoweth, Erica, and Orion A Lewis. 2013. "Unpacking nonviolent campaigns: Introducing the NAVCO 2.0 dataset." *Journal of Peace Research* 50 (3): 415–423.

Chenoweth, Erica, and Maria J Stephan. 2011. *Why civil resistance works: The strategic logic of nonviolent conflict.* New York: Columbia University Press.

Chenoweth, Erica, and Jay Ulfelder. 2017. "Can structural conditions explain the onset of nonviolent uprisings?" *Journal of Conflict Resolution* 61 (2): 298–324.

Chorley, Baroness Katharine Campbell Hopkinson Chorley. 1953. *Armies and the Art of Revolution.* Boston: Beacon Press.

Cook, Scott J, Jude C Hays, and Robert J Franzese. 2018. "Fixed effects in rare events data: a penalized maximum likelihood solution." *Political Science Research and Methods,* 1–14.

Dahl, Marianne. 2016. "Military Defection." *PRIO Policy Brief* (Oslo, Norway) 18:1–4.

De Bruin, Erica. 2018. "Preventing coups d'état: How counterbalancing works." *Journal of Conflict Resolution* 62 (7): 1433–1458.

———. 2020. "Mapping coercive institutions: The State Security Forces dataset, 1960–2010." *Journal of Peace Research,* 0022343320913089.

Fariss, Christopher J. 2014. "Respect for Human Rights has Improved Over Time: Modeling the Changing Standard of Accountability." *American Political Science Review* 108 (2): 297–318.

Frantz, Erica, Andrea Kendall-Taylor, Joseph Wright, and Xu Xu. 2020. "Personalization of power and repression in dictatorships." *The Journal of Politics* 82 (1): 372–377.

Geddes, Barbara. 2003. *Paradigms and Sand Castles: Theory Building and Research Design in Comparative Politics.* University of Michigan.

Geddes, Barbara, Joseph Wright, and Erica Frantz. 2014. "Autocratic Breakdown and Regime Transitions." *Perspectives on Politics* 12 (2): 313–331.

———. 2018. *How Dictatorships Work: Power, Personalization, and Collapse.* New York: Cambridge University Press.

Greitens, Sheena Chestnut. 2016. *Dictators and their secret police: Coercive institutions and state violence.* Cambridge University Press.

Hainmueller, Jens, Jonathan Mummolo, and Yiqing Xu. 2019. "How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice." *Political Analysis* 27 (2): 163–192.

Hamilton, James D. 2018. "Why you should never use the Hodrick-Prescott filter." *Review of Economics and Statistics* 100 (5): 831–843.

Imai, Kosuke, and In Song Kim. 2019. "When should we use unit fixed effects regression models for causal inference with longitudinal data?" *American Journal of Political Science* 63 (2): 467–490.

Johnson, Paul Lorenzo. 2017. "Virtuous Shirking: Social Identity, Military Recruitment, and Unwillingness to Repress." PhD diss., University of California, Davis.

Katz, Mark N. 2004. "Democratic revolutions: Why some succeed, why others fail." *World Affairs* 166 (3): 163–170.

Lee, Terence. 2014. *Defect or defend: military responses to popular protests in authoritarian Asia.* Baltimore, MD: John's Hopkins University Press.

Lewbel, Arthur. 2000. "Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables." *Journal of Econometrics* 97 (1): 145–177.

Li, Qi. 2000. "Efficient Estimation of Additive Partially Linear Models." *International Economic Review* 41 (4): 1073–1092.

Libois, François, and Vincenzo Verardi. 2013. "Semiparametric fixed-effects estimator." *The Stata Journal* 13 (2): 329–336.

Lutscher, Philipp M. 2016. "The More Fragmented the Better?—The Impact of Armed Forces Structure on Defection during Nonviolent Popular Uprisings." *International interactions* 42 (2): 350–375.

Makara, Michael. 2013. "Coup-proofing, military defection, and the Arab Spring." *Democracy and Security* 9 (4): 334–359.

Morency-Laflamme, Julien. 2018. "A question of trust: Military defection during regime crises in Benin and Togo." *Democratization* 25 (3): 464–480.

Neu, Kara Leigh Kingma. 2018. "Defections and Democracy: Explaining Military Loyalty Shifts and Their Impacts on Post-Protest Political Change." PhD diss., University of Denver.

Pemstein, Daniel, Kyle L Marquardt, Eitan Tzelgov, Yi-ting Wang, Joshua Krusell, and Farhad Miri. 2019. "The V-Dem measurement model: latent variable analysis for cross-national and cross-temporal expert-coded data." *V-Dem Working Paper* 21.

Pilster, Ulrich, and Tobias Böhmelt. 2011. "Coup-proofing and military effectiveness in interstate wars, 1967–99." *Conflict Management and Peace Science* 28 (4): 331–350.

———. 2012. "Do Democracies Engage Less in Coup-Proofing? On the Relationship between Regime Type and Civil-Military Relations." *Foreign Policy Analysis* 8:355–371.

Pion-Berlin, David, Diego Esparza, and Kevin Grisham. 2014. "Staying quartered: civilian uprisings and military disobedience in the twenty-first century." *Comparative Political Studies* 47 (2): 230–259.

Powell, Jonathan. 2012. "Determinants of the Attempting and Outcome of Coups d'etat." *Journal of Conflict Resolution* 56 (6): 1017–1040.

Quinlivan, James T. 1999. "Coup-proofing: Its Prace and Consequences in the Middle East." *International Security* 24 (2): 131–165.

Reise, Steven P., and Niels G. Waller. 2009. "Item Response Theory and Clinical Measurement." *Annual Review of Clinical Psychology* 85:27–48.

Russell, Diana EH. 1974. *Rebellion, Revolution, and Armed Force: A Comparative Study of Fifteen Countries with Special Emphasis on Cuba and South Afri.* New York: Academic Press.

Song, Wonjun. 2022. "Dictators, personalized security forces, and coups." *International Interactions,* 1–29.

Sudduth, Jun Koga. 2017. "Strategic Logic of Elite Purges in Dictatorships." *Comparative Political Studies,* 1–39.

———. 2021. "Purging militaries: introducing the Military Purges in Dictatorships (MPD) dataset." *Journal of Peace Research.*

Sutton, Jonathan, Charles R Butcher, and Isak Svensson. 2014. "Explaining political jiu-jitsu Institution-building and the outcomes of regime violence against unarmed protests." *Journal of Peace Research* 51 (5): 559–573.

Thurber, Ches. 2018. "Ethnic Barriers to Civil Resistance." *Journal of Global Security Studies* 3 (3): 255–270.

Timoneda, Joan C. 2021. "Estimating group fixed effects in panel data with a binary dependent variable: how the LPM outperforms logistic regression in rare events data." *Social Science Research* 93:102486.

Wright, Joseph. 2021. "The latent characteristics that structure autocratic rule." *Political Science Research and Methods* 9 (1): 1–19.