Online Appendix

This appendix contains data, statistics and additional tests that were not reported in the main text of the book. In section A1 we describe the coding procedure and list all cases included in the datasets used for empirical analysis. In section A2, we report descriptive statistics for all covariates included in the analysis. Section A3 covers the matching procedures used in the empirical analysis. In section A4 we report detailed results from the survival analysis in chapter 2. Section A5 describes the results related to the two-turnover test in chapter 2. In section A6, we report the detailed results for the regression models used in the analysis of democratic quality in chapter 2. Finally, section A7 provides additional results for the mechanisms discussed in chapter 4 of the book.

A1: Coding procedure

The data for the empirical analysis combines information about democratic transitions with information about the involvement of resistance campaigns in the transition process. We combined the data on democratic regimes provided by three datasets (Boix et al., 2013; Geddes et al., 2014; Ulfelder, 2012) with data on resistance campaigns from the NAVCO 2.0 dataset (Chenoweth & Lewis, 2013). For the datasets provided by Geddes et al. (2014) and Ulfelder (2012) we coded the transition process for all regimes that democratized after 1955. For the Boix et al. (2013) data we coded all democratic regimes that evolved after 1945. Given the coverage of the NAVCO 2.0 dataset, we consider only democratic regimes that originated before 2007 and measure indicators of democratic consolidation until 2011. Furthermore, we focus on resistance campaigns aimed at political change of the incumbent autocratic regime and did not consider campaigns where the goal was ‘territorial secession’, ‘greater autonomy’, or ‘anti-occupation’.

To ensure the validity of this coding, we inspected each case with respect to the question if the form of resistance was violent or nonviolent and checked whether there was indeed a causal link between the resistance campaign and the transition process. For each regime-dataset this coding procedure was conducted by at least two independent coders using historical and case-specific information. Cases where the coders disagreed entered a problem set, which was then evaluated by all authors to arrive at the final coding.[1] Table A1.1 lists all cases included in the analysis together with the final coding of resistance campaigns inducing the transition process for the regime dataset by Ulfelder (2012). The same information is given for the Geddes at al. (2014) data and the Boix et al. (2013) data in tables A1.2 and A1.3 respectively. Tables A1.4, A1.5, and A1.6 provide information about the distribution of transition events for each dataset.

[1] Some of this coding process was conducted while working on other articles (Bayer et al. 2016; Bethke and Pinckney 2017) and therefore involved coding decisions by other authors.

Table A1.1: Cases included in the analysis when using regime data by Ulfelder

Country	Transition year	Resistance campaign
Albania	1991	nonviolent
Albania	1997	none
Argentina	1958	none
Argentina	1963	none
Argentina	1973	none
Argentina	1983	nonviolent
Azerbaijan	1992	none
Bangladesh	1991	nonviolent
Benin	1991	nonviolent
Bolivia	1956	none
Bolivia	1982	nonviolent
Brazil	1985	nonviolent
Bulgaria	1990	nonviolent
Burkina Faso	1978	none
Burma	1960	none
Burundi	1993	violent
Burundi	2005	none
CAR	1993	nonviolent
Chile	1990	nonviolent
Colombia	1958	none
Comoros	2002	none
Congo-Brazzaville	1992	none
Croatia	2000	nonviolent
Cyprus	1983	none
Czechoslovakia	1990	nonviolent
Dominican Republic	1963	none
Dominican Republic	1978	none
Dominican Republic	1996	none
Ecuador	1966	none
Ecuador	1979	none
Ecuador	2003	none
El Salvador	1982	violent
Fiji	1992	none
Georgia	1992	none
Ghana	1969	none
Ghana	1979	none
Ghana	1992	none
Greece	1974	nonviolent
Guatemala	1986	violent
Guinea-Bissau	1994	none
Guinea-Bissau	2000	none
Guinea-Bissau	2005	none
Guyana	1992	nonviolent
Haiti	1991	none
Haiti	1994	none
Honduras	1957	none

Country	Transition year	Resistance campaign
Honduras	1971	none
Honduras	1982	none
Hungary	1990	nonviolent
India	1977	none
Indonesia	1955	violent
Indonesia	1999	nonviolent
Kenya	2002	none
Korea, South	1960	nonviolent
Korea, South	1963	none
Korea, South	1988	nonviolent
Lebanon	2005	nonviolent
Lesotho	1993	none
Lesotho	1994	none
Liberia	2006	none
Madagascar	1993	nonviolent
Malawi	1994	nonviolent
Mali	1992	nonviolent
Mexico	1997	none
Moldova	1997	none
Mongolia	1990	nonviolent
Mozambique	1994	none
Nepal	1959	none
Nepal	1991	nonviolent
Nicaragua	1984	violent
Niger	1993	none
Niger	1999	none
Nigeria	1979	none
Pakistan	1972	none
Pakistan	1988	none
Panama	1956	none
Panama	1989	nonviolent
Paraguay	1993	none
Peru	1956	none
Peru	1963	none
Peru	1980	none
Peru	1993	none
Philippines	1986	nonviolent
Poland	1989	nonviolent
Portugal	1976	nonviolent
Romania	1996	none
Russia	1993	none
Senegal	2000	nonviolent
Sierra Leone	1968	none
Sierra Leone	1996	violent
Sierra Leone	1998	none
South Africa	1994	nonviolent
Spain	1977	none

Country	Transition year	Resistance campaign
Sri Lanka	1994	none
Sudan	1965	violent
Sudan	1986	nonviolent
Syria	1961	none
Taiwan	1992	none
Tanzania	1995	nonviolent
Thailand	1975	violent
Thailand	1983	none
Thailand	1992	nonviolent
Turkey	1961	none
Turkey	1973	none
Turkey	1983	none
Ukraine	2005	nonviolent
Uruguay	1985	nonviolent
Venezuela	1959	nonviolent
Yugoslavia, Former	2000	nonviolent
Zambia	1991	nonviolent
Zambia	2006	none
Zimbabwe	1980	violent

Table A1.2: Cases included in the analysis when using regime data by Geddes et al.

Country	Transition year	Resistance campaign
Albania	1992	nonviolent
Argentina	1974	none
Argentina	1984	nonviolent
Azerbaijan	1993	none
Bangladesh	1991	nonviolent
Benin	1992	nonviolent
Bolivia	1983	nonviolent
Brazil	1986	nonviolent
Bulgaria	1991	nonviolent
Burundi	1994	violent
Burundi	2006	none
Cen African Rep	1994	nonviolent
Chile	1990	nonviolent
Colombia	1959	none
Congo-Brz	1993	none
Czechoslovakia	1990	nonviolent
Dominican Rep	1963	none
Dominican Rep	1979	none
Ecuador	1969	none
Ecuador	1980	none
El Salvador	1995	none
Georgia	2005	nonviolent
Ghana	1970	none
Ghana	1980	none
Ghana	2001	nonviolent
Greece	1975	nonviolent
Guatemala	1996	none
Guinea Bissau	2001	none
Guinea Bissau	2006	none
Haiti	1991	none
Haiti	1995	none
Honduras	1958	none
Honduras	1972	none
Honduras	1982	none
Hungary	1991	nonviolent
Indonesia	2000	nonviolent
Kenya	2003	none
Korea South	1961	nonviolent
Korea South	1988	nonviolent
Lebanon	2006	nonviolent
Lesotho	1994	none
Liberia	2006	none
Madagascar	1994	nonviolent
Malawi	1995	nonviolent
Mali	1993	nonviolent
Mexico	2001	nonviolent

Country	Transition year	Resistance campaign
Moldova	1992	none
Mongolia	1994	nonviolent
Myanmar	1961	none
Nepal	1992	nonviolent
Nicaragua	1991	violent
Niger	1994	nonviolent
Niger	2000	none
Nigeria	1980	none
Nigeria	2000	nonviolent
Pakistan	1972	none
Pakistan	1989	none
Panama	1956	none
Panama	1990	nonviolent
Paraguay	1994	none
Peru	1957	none
Peru	1964	none
Peru	1981	none
Peru	2002	nonviolent
Philippines	1987	nonviolent
Poland	1990	nonviolent
Portugal	1977	nonviolent
Romania	1991	violent
Russia	1992	none
Senegal	2001	nonviolent
Serbia	2001	nonviolent
Sierra Leone	1997	violent
Sierra Leone	1999	none
South Africa	1995	nonviolent
Spain	1978	none
Sri Lanka	1995	none
Sudan	1966	violent
Sudan	1987	nonviolent
Syria	1955	none
Syria	1962	none
Taiwan	2001	none
Thailand	1976	violent
Thailand	1989	none
Thailand	1993	nonviolent
Turkey	1962	none
Turkey	1984	none
Uruguay	1985	nonviolent
Venezuela	1959	nonviolent
Zambia	1992	nonviolent

Table A1.3: Cases included in the analysis when using regime data by Boix et al.

Country	Transition year	Resistance campaign
Albania	1992	nonviolent
Albania	1997	violent
Argentina	1958	none
Argentina	1963	none
Argentina	1973	none
Argentina	1983	nonviolent
Bangladesh	1991	nonviolent
Benin	1991	nonviolent
Bolivia	1979	none
Bolivia	1982	nonviolent
Brazil	1946	none
Brazil	1985	nonviolent
Bulgaria	1990	nonviolent
Burundi	2005	none
Cape Verde	1991	none
Central African Republic	1993	nonviolent
Chile	1990	nonviolent
Colombia	1958	none
Comoros	2006	none
Costa Rica	1949	violent
Croatia	2000	nonviolent
Cyprus	1977	none
Czechoslovakia	1990	nonviolent
Dominican Republic	1966	none
Ecuador	1948	none
Ecuador	1979	none
Ecuador	2003	none
El Salvador	1984	violent
Gambia	1972	none
Georgia	2004	nonviolent
Ghana	1970	none
Ghana	1979	none
Ghana	1997	none
Greece	1974	nonviolent
Guatemala	1945	nonviolent
Guatemala	1958	none
Guatemala	1966	none
Guatemala	1986	none
Guinea-Bissau	1994	none
Guyana	1992	nonviolent
Honduras	1957	none
Honduras	1971	none
Honduras	1982	none
Hungary	1990	nonviolent
Indonesia	1955	none
Indonesia	1999	nonviolent

Country	Transition year	Resistance campaign
Kenya	2002	none
Korea, South	1960	nonviolent
Korea, South	1988	nonviolent
Latvia	1993	nonviolent
Lebanon	1971	none
Lesotho	2002	nonviolent
Liberia	2006	nonviolent
Lithuania	1992	nonviolent
Madagascar	1993	nonviolent
Malawi	1994	nonviolent
Mali	1992	nonviolent
Mexico	2000	nonviolent
Mongolia	1990	nonviolent
Mozambique	1994	none
Myanmar	1960	none
Nepal	1991	nonviolent
Nicaragua	1984	none
Niger	1993	nonviolent
Niger	1999	none
Nigeria	1979	none
Pakistan	1950	none
Pakistan	1988	none
Panama	1950	none
Panama	1952	none
Panama	1991	none
Paraguay	2003	none
Peru	1956	none
Peru	1963	none
Peru	1980	none
Peru	2001	nonviolent
Philippines	1986	nonviolent
Poland	1989	nonviolent
Portugal	1976	nonviolent
Romania	1991	violent
Sao Tome	1991	none
Senegal	2000	none
Sierra Leone	2002	none
Solomon Islands	2006	none
South Africa	1994	nonviolent
Spain	1977	none
Sri Lanka	1991	none
Sudan	1965	nonviolent
Sudan	1986	nonviolent
Suriname	1988	none
Suriname	1991	none
Taiwan	1996	none
Thailand	1975	none

Country	Transition year	Resistance campaign
Thailand	1983	none
Thailand	1992	nonviolent
Turkey	1961	none
Turkey	1983	none
Uganda	1980	none
Uruguay	1985	nonviolent
Venezuela	1959	nonviolent
Yugoslavia	2000	nonviolent

Table A1.4: Categorical coding of resistance campaigns during transitions (Ulfelder data)

	Freq.	Percent
No Resistance Campaign	66	58.93
Violent Resistance Campaign	9	8.03
Nonviolent Resistance Campaign	37	33.04
Total	112	100

Table A1.5: Categorical coding of resistance campaigns during transitions (Geddes data)

	Freq.	Percent
No Resistance Campaign	44	49.44
Violent Resistance Campaign	39	43.82
Nonviolent Resistance Campaign	6	6.74
Total	89	100

Table A1.6: Categorical coding of resistance campaigns during transitions (Boix et al. data)

	Freq.	Percent
No Resistance Campaign	56	55.45
Violent Resistance Campaign	41	40.59
Nonviolent Resistance Campaign	4	3.96
Total	101	100

A2: Descriptive statistics

Table A2.1. Descriptive statistics for the main covariates (Ulfelder cross-section data)

Variable	Obs	Mean	Std. Dev.	Min	Max
GDP p.c.$_{(log, t-1)}$	112	7.79	0.94	5.75	9.65
Military Legacy	112	0.39	0.49	0	1
Previous Instability	112	2.86	2.19	0	13
Neighboring Democracies$_{(t-1)}$	112	35.50	22.59	0	90
Total Population$_{(log, t-1)}$	112	9.22	1.32	6.44	13.33
Urbanization$_{(t-1)}$	112	0.21	0.14	0.00	0.66

Table A2.2. Descriptive statistics for the main covariates (Ulfelder panel data)

Variable	Obs	Mean	Std. Dev.	Min	Max
GDP p.c.$_{(log, t-1)}$	1448	8.22	1.06	5.09	10.28
Military Legacy	1448	0.44	0.5	0	1
Previous Instability	1448	3.08	2.26	0	13
Neighboring Democracies$_{(t-1)}$	1448	56.84	27.02	0	100
Total Population$_{(log, t-1)}$	1448	9.5	1.42	6.33	13.97
Urbanization$_{(t-1)}$	1448	0.28	0.15	0	0.75

Table A2.3. Descriptive statistics for the main covariates (Boix cross-section data)

Variable	Obs	Mean	Std. Dev.	Min	Max
GDP p.c.$_{(log, t-1)}$	101	7.82	0.9	5.75	9.82
Military Legacy	101	0.38	0.49	0	1
Previous Instability	101	1.30	1.77	0	8
Neighboring Democracies$_{(t-1)}$	101	34.06	23.23	0	91.67
Total Population$_{(log, t-1)}$	101	8.99	1.51	4.74	12.24
Urbanization$_{(t-1)}$	101	0.21	0.15	0	0.66

Table A2.4. Descriptive statistics for the main covariates (Boix panel data)

Variable	Obs	Mean	Std. Dev.	Min	Max
GDP p.c.$_{(log,\ t-1)}$	1554	8.26	1.02	5.09	10.28
Military Legacy	1554	0.4	0.49	0	1
Previous Instability	1554	1.47	1.91	0	8
Neighboring Democracies$_{(t-1)}$	1554	53.84	26.4	0	100
Total Population$_{(log,\ t-1)}$	1554	9.12	1.50	4.74	12.35
Urbanization$_{(t-1)}$	1554	0.27	0.16	0	0.75

Table A2.5. Descriptive statistics for the main covariates (Geddes cross-section data)

Variable	Obs	Mean	Std. Dev.	Min	Max
GDP p.c.$_{(log,\ t-1)}$	89	7.84	0.97	5.84	10.06
Military Legacy	89	0.42	0.5	0	1
Previous Instability	89	3.65	2.3	0	9
Neighboring Democracies$_{(t-1)}$	89	33.65	24.36	0	95.45
Total Population$_{(log,\ t-1)}$	89	9.35	1.2	6.82	12.25
Urbanization$_{(t-1)}$	89	0.22	0.14	0	0.66

Table A2.6. Descriptive statistics for the main covariates (Geddes panel data)

Variable	Obs	Mean	Std. Dev.	Min	Max
GDP p.c.$_{(log,\ t-1)}$	1165	8.3	1.02	5.09	10.28
Military Legacy	1165	0.49	0.5	0	1
Previous Instability	1165	3.66	2.38	0	9
Neighboring Democracies$_{(t-1)}$	1165	57.75	27.63	0	100
Total Population$_{(log,\ t-1)}$	1165	9.57	1.15	6.82	12.35
Urbanization$_{(t-1)}$	1165	0.29	0.15	0	0.75

A3: Description of the matching procedures

The matching methods used for the empirical analysis build on the so called Neyman-Rubin potential outcome framework of causality, which relies on counterfactuals to identify causal effects of treatment variables on outcome variables (Rubin, 2005; Splawa-Neyman, 1990 [1923]). We consider the presence of an NVR campaign during democratic transition as the treatment and the outcome refers to different measures of democratic consolidation (i.e. democratic survival, two-turnovers of power, or democratic quality). Potential outcomes refer the outcome of each measure of democratic consolidation under each of the two treatment conditions, i.e. transition was induced by an NVR campaign or not. However, if we observe a regime's outcome when the transition process was shaped by an NVR campaign, we cannot observe the outcome of this regime without NVR. For instance, our main dataset records an NVR induced transition in Mali in 1992 and subsequent democratic survival until 1997, when a democratic breakdown occurred. We cannot observe how long democracy would have survived in Mali if it experienced a transition by other means than NVR. This problem is often referred to as the 'fundamental problem of causal inference' (Holland, 1986: 947). Therefore, we have to settle for the estimation of an average treatment effect (ATE). The ATE is derived by comparing the average outcome of subjects that received the treatment with the average outcome of other subjects that did not receive the treatment (Guo & Fraser, 2010: 25). For our purposes, we compare the average of the respective outcome measure of democratic consolidation of regimes with NVR with the average outcome of those regimes without NVR during the transition process. Both of these quantities can be observed and thus compared.

However, in order to calculate an unbiased estimate of the ATE, the conditional independence assumption has to be met. This assumption requires that assignment to treatment is independent of the outcome. In observational studies, such as this one, where randomization of treatment is not possible but where the assignment instead follows a natural process of (self-)selection, the assumption is likely to be violated (Guo & Fraser, 2010: 30-35). With regard to the presence of NVR during the transition process, which is our treatment of interest, we are concerned that certain characteristics of a country affect opportunities and motivations for employing NVR while also influencing democratic consolidation of the subsequent democratic regime. In order to address this problem of selection/confounding, we apply different matching procedures.

With matching methods, treatment and control subjects are matched on a set of baseline characteristics with the goal of compiling a balanced sample where groups are as comparable as possible. The basic idea is that if two subjects are sufficiently similar on observed covariates but differ in terms of treatment assignment, then the selection process of treatment assignment is 'as good as random' (Sekhon, 2009: 495). In the case of propensity score matching, the method is implemented in three steps: (1) estimating the probability of receiving the treatment (the propensity score) using a set of observed covariates; (2) matching subjects on the propensity score to create a new dataset consisting of matched pairs of treated and control subjects; (3) estimating the effect of treatment on outcome using a matched sample.

For this study, the propensity score refers to the probability of NVR presence during the transition process. The propensity score essentially summarizes all observed factors that influence the probability of treatment assignment (Guo & Fraser, 2010: 132-135). We use a logistic regression model to estimate the conditional probability of NVR assignment given a set of observed covariates. As discussed in the main text, we rely on a set of confounding and prognostic factors for the estimation of the propensity score, namely: GDP per capita, military legacy, previous instability, proportion of neighboring democracies, total population and urbanization.

In order to construct the matched sample, we pair each case where the transition process was shaped by NVR with one case where the transition process was not influenced by NVR whose propensity score is as similar as possible. In order to ensure the robustness of the results, we follow the suggestion by Austin (2014: 1245-47) to apply multiple different matching schemes, namely

(1) greedy nearest neighbor matching, (2) greedy nearest neighbor matching within a caliper and (3) optimal pair matching. These schemes use different procedures to construct the matches. However, with all of them we implement one-to-one matching, which means that each treated observation is matched to only one control observation. Furthermore, we employ the restriction of no replacement, which means that control observation may not be used multiple times as a match for different treated observations.

Greedy nearest neighbor matching identifies matches for each treated observation by looking among the controls for the observation with the smallest difference in the propensity score. Once the control observation with the smallest difference is found, a pair is formed and both observations are removed from the sample and enter the new matched sample. This procedure is repeated until all treated units are matched to a control unit. However, we apply a common support restriction, which means to discard observations with extreme propensity score values. For instance, those treated subjects for which there are no control subjects with propensity score values as large or as small are discarded (Stuart & Rubin, 2008: 168).

Greedy nearest neighbor matching within a caliper restricts the matching of subjects to a caliper distance, i.e. threshold level for differences in propensity scores. The caliper distance is essentially a benchmark score that excludes those pairs of treated and control cases for which the difference in propensity score does not lie within the specified range (Austin, 2011: 406). The selection of caliper size creates a tradeoff between bias reduction and precision. If no restrictions are imposed with regard to the quality of the match, bias reduction for the observed covariates may be less effective. However, if the caliper is too tight, fewer treated units can be matched with controls and the effect estimates become less generalizable and are subject to another form of selection bias induced by the caliper restriction. As noted by Lunt: ‘It is no longer the effect of treatment for the treated subjects that is being estimated, but the effect of treatment for those treated subjects for whom we can find controls' (Lunt, 2014: 232). In line with previous research, we applied caliper sizes of 25% of the standard deviation of the sample estimated propensity scores (Austin, 2011: 407; Guo & Fraser, 2010: 147).

With optimal pair matching, each treated observation is matched to a different control observation in order to minimize the overall distance (i.e. propensity score) among matched pairs (Guo & Fraser, 2010: 149-153). This means that matches are created with the goal to minimize the sum of the distance between all pairs. Thus, while in greedy matching, the sample is constructed by sequentially forming one pair after another, optimal matching evaluates potential matches jointly and the overall quality of the sample is of key interest. Correspondingly, optimal matching always creates a complete sample of matched pairs (i.e. each treated subject is paired with at least one control subject).[2]

Because the different matching procedures only match one control case to each treatment case, some of the control observations are discarded and not part of the matched samples. Therefore, any empirical analysis that uses a matched same does not estimate an ATE but an average treatment effect for the treated (ATT) cases. Thus, the estimated effect of NVR on the respective outcome can only be generalized for the population of treated cases. Accordingly, we cannot tell how all control cases (i.e. democratic transitions without NVR) would have developed if they had received the treatment.

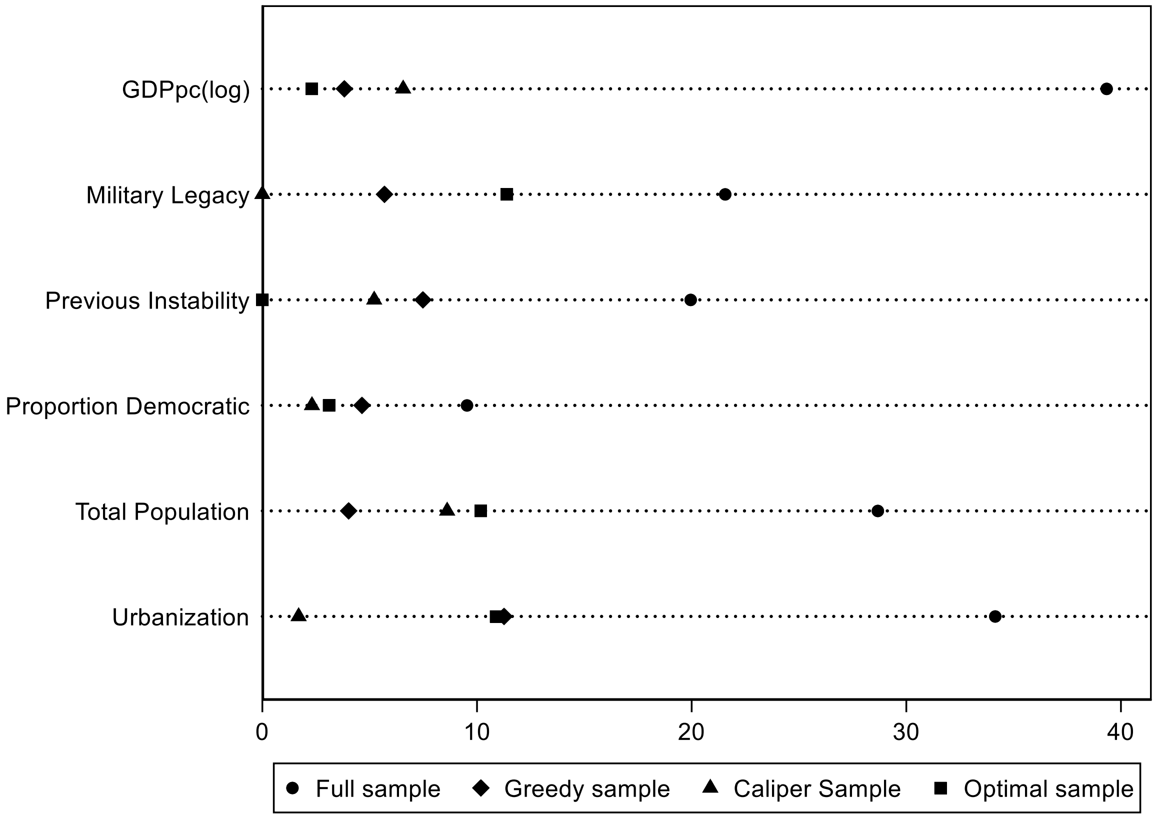
Covariate balance

Next, we describe covariant balance for all matched samples that were created by the respective procedures described above. Thus, we provide information of how well the different matching schemes achieved their goal of creating samples where cases where transition was induced by NVR and cases without this feature are as similar as possible on observed covariates. We report

[2] For a more detailed discussion of optimal matching see Hansen & Klopfer (2006).

standardized differences of covariate means along with p-values from two-sample t-tests and bootstrapped Kolmogorov-Smirnov (KS) tests of equality of distributions. Figure A3.1 describes standardized differences of means between treatment and control group for the full sample and the matched samples for the Ulfelder dataset.

Figure A3.1. Standardized differences of means (Ulfelder data)



As shown in A3.1, all matching schemes substantially reduced bias among the covariates, when compared to the full sample. Especially, the covariates measuring GDP per capita and urbanization substantially differ between treatment and control group and therefore are particularly likely to confound the estimates of the effect of NVR on democratic consolidation. The mean bias among covariates in terms of standardized differences is 25.5% in the full sample. With the different matching procedures, the average bias was reduced to 6.1% in the greedy sample, 4.1% in the caliper sample, and 6.3% in the sample created with the optimal matching scheme.

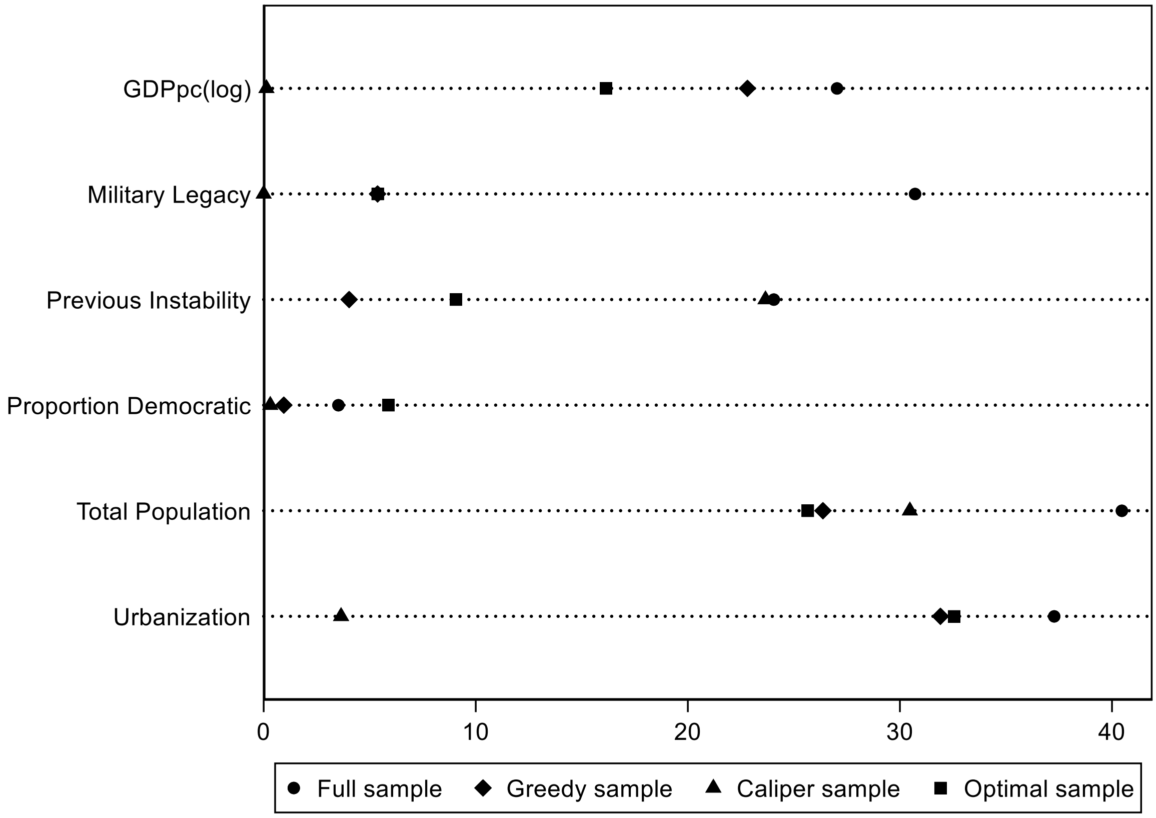
Table A3.1 reports p-values for two sample t-tests and bootstrapped KS tests. As shown in Table A3.1, in the full sample there are significant difference in the means and distribution of the covariates GDP p.c.$_{(\log,\ t-1)}$ and Urbanization$_{(t-1)}$ between treatment and control group. However, the different matching schemes are able to reduce bias quite effectively. None of the matched samples shows a significant difference in means or with regard to the distribution of the covariates between treatment and control group.

Table A3.1. Balancing statistics (Ulfelder data)

	Full Sample		Greedy Sample		Caliper Sample		Optimal Sample	
	t-test	KS-test	t-test	KS-test	t-test	KS-test	t-test	KS-test
GDP p.c.$_{(\log,\ t-1)}$	0.05	0	0.87	0.32	0.78	0.82	0.92	0.32
Military Legacy	0.29	.	0.81	.	1	.	0.63	.
Previous Instability	0.33	0.21	0.73	0.91	0.82	0.99	1	0.92
Neighboring Democracies$_{(t-1)}$	0.63	0.37	0.85	1	0.93	0.94	0.9	1
Total Population$_{(\log,\ t-1)}$	0.17	0.13	0.87	0.69	.75	0.82	0.69	0.86
Urbanization$_{(t-1)}$	0.09	0.1	0.63	0.68	0.95	0.96	0.65	0.68

Now figure A3.2 shows standardized differences for the full sample and the matched samples for the Geddes dataset.

Figure A3.2. Standardized differences of means (Geddes data)



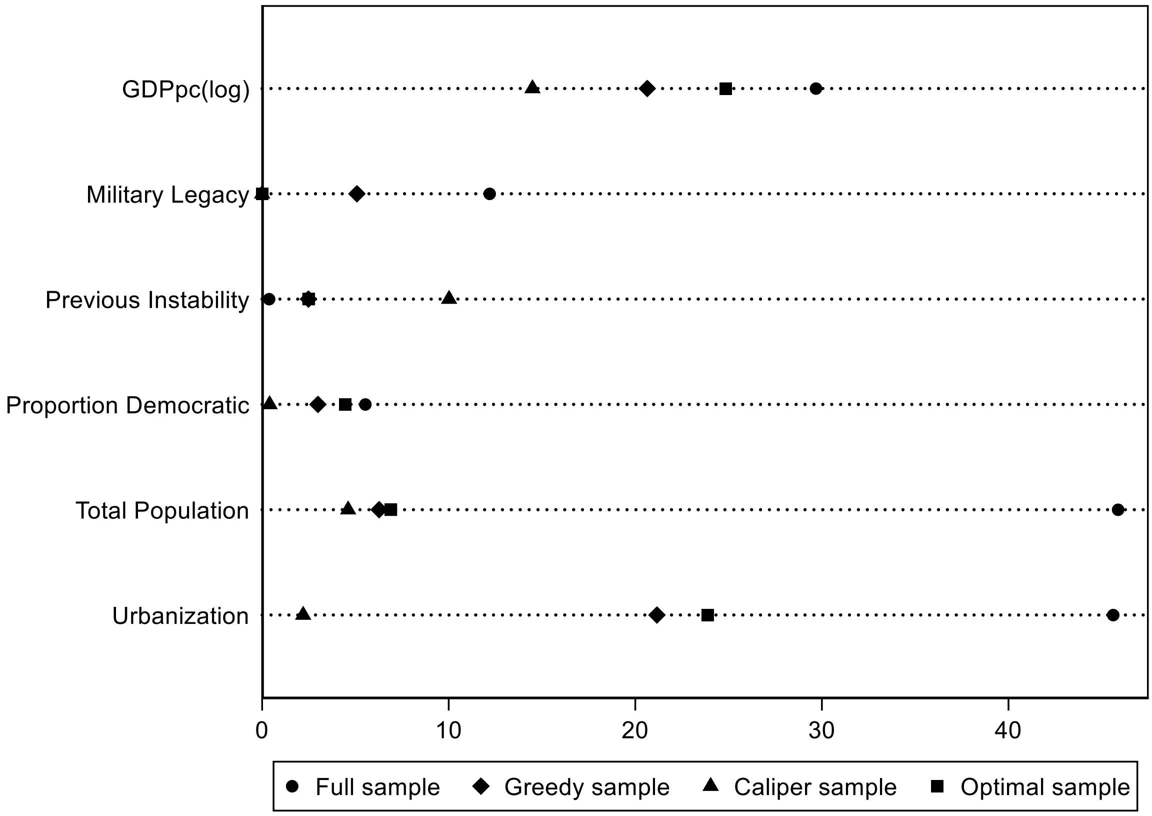
All matching schemes reduced bias among the covariates, when compared to the full sample but the improvement is not as good as with the Ulfelder (2012) data. The mean bias among covariates in terms of standardized differences is 27.2% in the full sample. With the different matching procedures, the average bias is reduced to 15.2% in the greedy sample, 9.7% in the caliper sample and 15.8% in the sample created with the optimal matching scheme. We consider only the sample produced by the caliper matching scheme managed as sufficiently balanced in terms of standardized difference of the potential confounding variables. Table A3.2 reports p-values from two-sample t-tests and bootstrapped Kolmogorov-Smirnov tests of equality of distributions.

Table A3.2. Balancing statistics (Geddes data)

	Full Sample		Greedy Sample		Caliper Sample		Optimal Sample	
	t-test	KS-test	t-test	KS-test	t-test	KS-test	t-test	KS-test
GDP p.c.$_{(log, t-1)}$	0.2	0.02	0.32	0.04	1	0.33	0.47	0.07
Military Legacy	0.16	.	0.81	.	1	.	0.81	.
Previous Instability	0.23	0.02	0.84	0.28	0.3	0.32	0.66	0.18
Neighboring Democracies$_{(t-1)}$	0.86	0.86	0.96	0.96	0.99	0.99	0.78	0.95
Total Population$_{(log, t-1)}$	0.06	0.12	0.26	0.52	0.31	0.52	0.27	0.52
Urbanization$_{(t-1)}$	0.07	0.09	0.15	0.22	0.9	0.74	0.14	0.22

As shown in table A3.2, again the full sample shows significant difference in means and distribution for some covariates. With the three matched samples none of the covariates shows significant differences in the means across treatment status. However, for the distribution of the of GDP p.c.$_{(log, t-1)}$ across regimes with and without NVR the KS-tests reveal significant differences in the samples created by greedy and optimal matching respectively. Only the caliper sample manages to adjust all potential confounding variables sufficiently. Finally, standardized differences for the covariates in the Boix data are reported in figure A3.3.

Figure A3.3. Standardized differences of means (Boix data)



As shown in figure A3.3, the different matching schemes are able to reduce covariate imbalance in the Boix data. Especially the covariates measuring total population and urbanization, which appear to be severely biased in the full sample, show substantially lower imbalance in the matched samples. Only the covariate balance for previous instability is worse in the matched samples than in the full sample. However, in none of the different samples this covariate appears to be substantially biased (i.e. >10% standardized difference). The average bias for covariates is 23.2% in the full sample. For the greedy, caliper, and optimal sample average bias is 9.8%, 5.3%, and 10.4% respectively. While the average covariate bias appears negligible in all three matched samples, none of the matching schemes is able to reduce covariate imbalance to below 10% for all covariates. Again, we further investigate potential problems with regard to covariate imbalance with two sample t-tests and bootstrapped KS-tests. The results are reported in table 3.3.

Table A3.3. Balancing statistics (Boix data)

	Full Sample		Greedy Sample		Caliper Sample		Optimal Sample	
	t-test	KS-test	t-test	KS-test	t-test	KS-test	t-test	KS-test
GDP p.c.$_{(\log,\ t\text{-}1)}$	0.12	0.02	0.31	0.1	0.53	0.35	0.23	0.1
Military Legacy	0.55	.	0.82	.	1	.	1	.
Previous Instability	0.98	0.68	0.9	0.73	0.65	.25	0.9	0.72
Neighboring Democracies$_{(t\text{-}1)}$	0.78	0.35	0.89	0.32	0.99	0.7	0.83	0.32
Total Population$_{(\log,\ t\text{-}1)}$	0.04	0.26	0.79	0.73	0.88	0.92	0.77	0.9
Urbanization$_{(t\text{-}1)}$	0.02	0.03	0.32	0.55	0.93	.75	0.27	0.38

Table A3.3 shows that in the full sample significant difference in means are present for the covariates total population and urbanization. Moreover, significant differences exist across regimes with and without NVR in the distribution of the covariates measuring GDP per capita and urbanization. The different matching schemes are able to create samples where none of the covariates shows any significant difference in means and distribution. The only exception is the variable GDP per capita, which still has a significant difference in the distribution in the sample created by optimal matching.

In sum, the different matching schemes are able to substantially reduce covariate imbalance in all three datasets used in the analysis. However, only for the Ulfelder data the matching procedures are able to balance all covariates to the point where the remaining bias is neglectable (i.e. <10% standardized difference for all covariates). For the Boix data, the matching schemes covariate balance is only slightly worse, but for the Geddes data, even after matching some of the covariates show substantial bias with regard to standardized differences, differences in means or difference in distribution of covariates between treatment and control group.

A4: Results of the survival analysis

In this section we report detailed results for the survival analysis. We begin by extending the comparison of Kaplan-Meier survival function between regimes that came about by means of NVR and regimes without this feature across different datasets. We report Kaplan-Meier survivor functions for the full sample and the samples generate by propensity score matching, which are balanced on potential confounding variables. For the matched samples, we compare democratic survival of all cases where the transition process was induced by an NVR campaign to a sub-sample of most similar cases where the transition occurred without NVR influence. Figure A4.1 describes the results for the Ulfelder data. In figure A4.1 and the following figures, the dashed line indicates the treated group of regimes that experienced NVR during the transition phase and the solid line indicates untreated regimes that did not experience NVR.

Figure A4.1. Kaplan-Meier survivor functions using different samples (Ulfelder data)

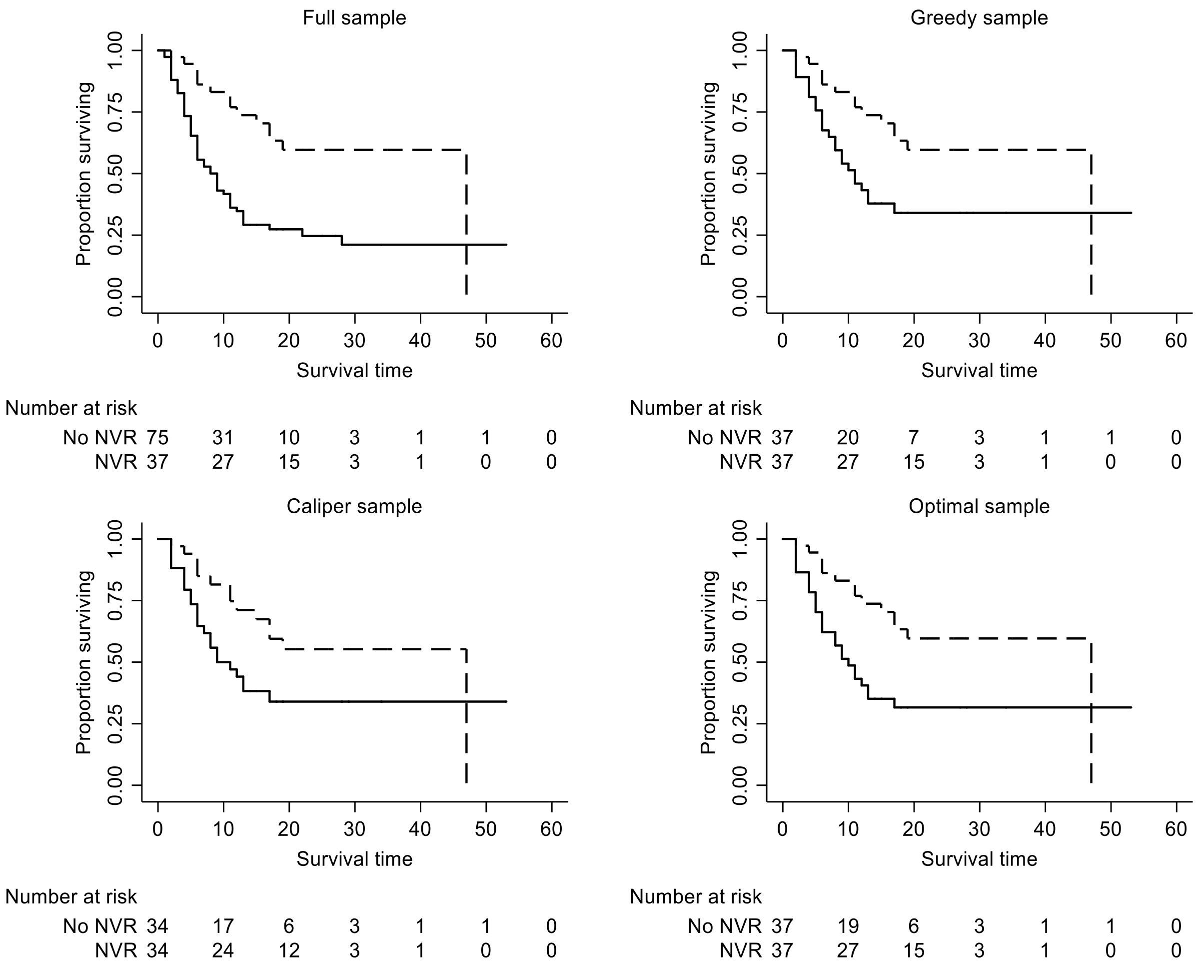
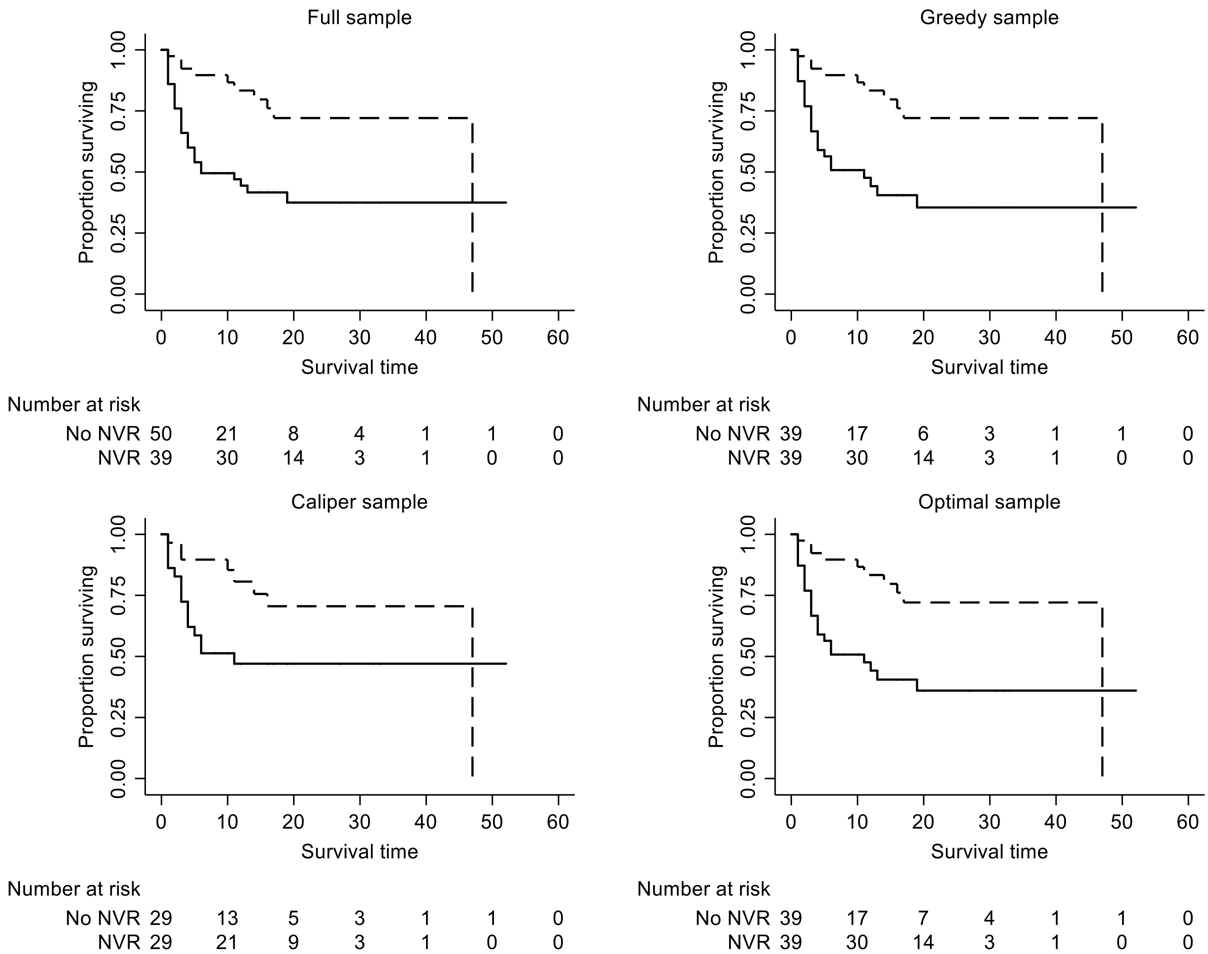


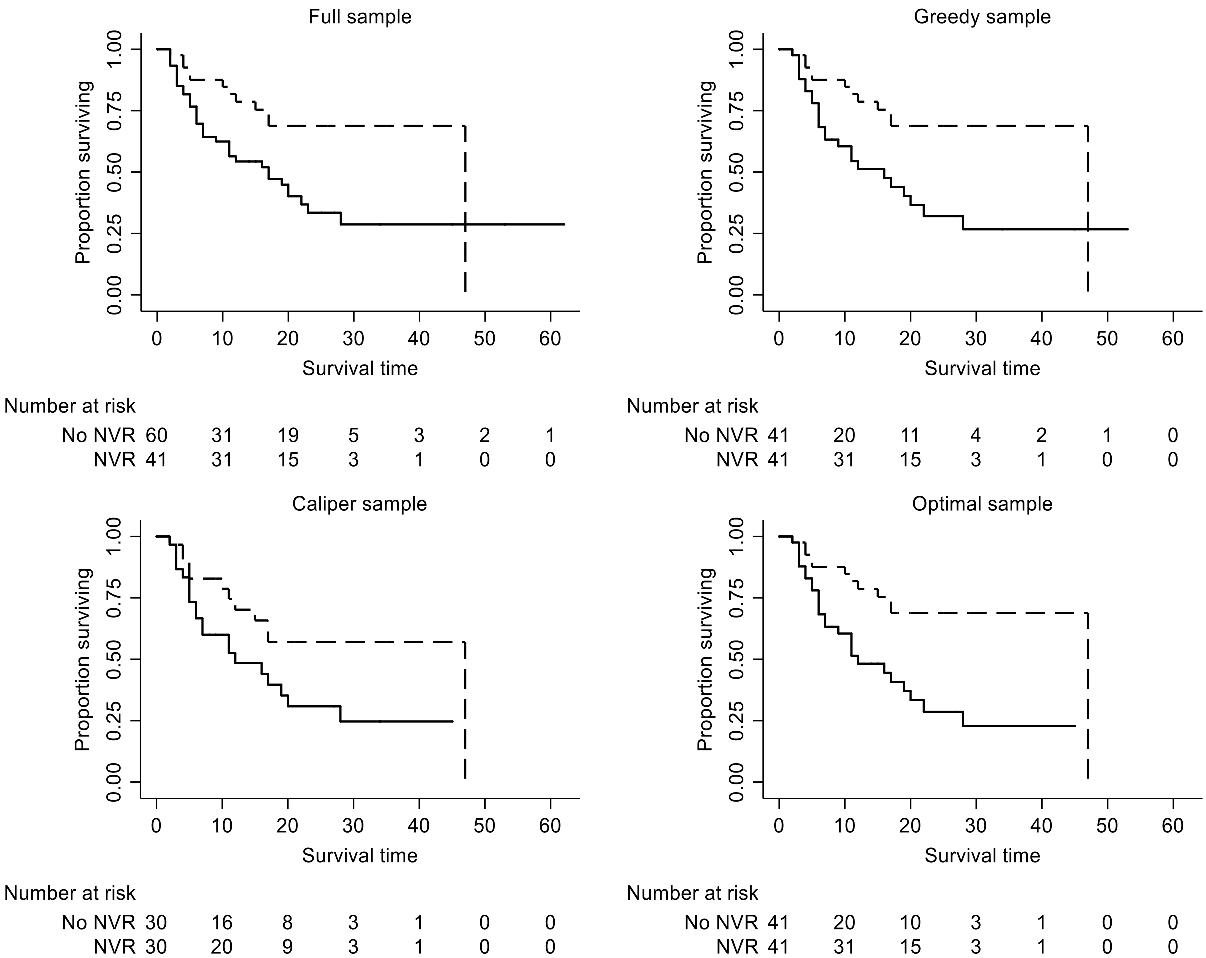
Figure A4.1 highlights that regardless of the sample used for the analysis, there always appears to be a substantial difference between the survival curves of democratic regimes that came about by means of NVR and those without this feature. Regimes with NVR induced transition have a higher median survival rate than regimes without this feature. This pattern is also confirmed by a log-rank test for equality of survivor functions. For all samples, figure A4.1 shows that the effect of NVR on democratic survival is mostly restricted to the time period of 30 years after transition. Afterwards, there are not enough cases that where observed for a longer time period to derive meaningful results. This means, we can be confident that the difference between survival functions exists during 30 years after transition, but we do not know if there is an effect after this time period. Figure A4.2 describes the results for Geddes data.

Figure A4.2. Kaplan-Meier survivor functions using different samples (Geddes data)



Again, across all samples, regimes with NVR induced transitions on average survive longer than regimes without NVR. This is also confirmed by a log rank test. Next, figure A4.3 describes the results for the three matching schemes using the Boix data.

Figure A4.3. Kaplan-Meier survivor functions using different samples (Boix data)



Also, the results from the Boix data show that across all samples, regimes that came about by means of NVR survive substantially longer relative to regimes without this feature. This is also confirmed by a log rank test stratified on matched pairs.

The effect of NVR seems to be most pronounced when using the Geddes data and least pronounced in the Boix data. However, all in general, the findings are robust, which means we always find a substantial and statistically significant difference in the survival functions of regimes with and without NVR, indicating that the former survive longer than the latter. Next, we analyze the effect of NVR on democratic survival with Cox-proportional hazard models again using the different samples and datasets.

Cox-regression models

The Cox proportional hazards estimator, which we use to estimate the effect of NVR-induced transitions on the hazard of democratic breakdown can be described by the following equation:

$$\log h_i(t) = \log h_0(t) + \beta_1 X_{1i} + \beta_2 X_{2it} + \ldots + \beta_N X_N$$

The expression h(t) denotes the hazard rate. The expression $h_0(t)$ represents the baseline hazard which is defined as the hazard for the occurrence of democratic breakdown if all covariates are zero. The values of the covariates are denoted as $X_1, X_2 \ldots, X_N$ and the expressions $\beta_1, \beta_2 \ldots, \beta_N$ stand for the coefficients which are meant to be estimated by the model. Whereas X_i represent the values of a time-fixed variable (e.g. our measures of NVR campaigns) for a specific regime included in the sample (denoted by i), X_{it}, represents the values of a time-varying variable for a regime included in the sample at a specific point in time (denoted by it). The hazard of democratic breakdown thus depends on the values of time-fixed and time-varying covariates. Therefore, we use a pooled data structure. Regime histories are broken down into discrete time intervals (regime-years) which are treated as different observations. Whereas our key independent variable measuring the presence of NVR campaigns during the transition period is time-constant, some of the control variables vary across time periods and thus warrant this data structure. *GDP per capita*, *total population*, *urbanization* and, *neighboring democracies* are measured as time-varying covariates. All time-varying covariates are lagged one year to address problems of reverse causality. Table A4.1 reports the results of the Cox models using the Ulfelder data for the full sample and the three matched samples.

Table A4.1. Cox-regression models (Ulfelder data)

	Full sample	Full sample	Greedy sample	Caliper sample	Optimal sample
GDP p.c.$_{(log,\ t-1)}$	0.68**	0.71**	0.82	0.84	0.61**
	(0.10)	(0.11)	(0.17)	(0.18)	(0.15)
Military Legacy	1.67*	1.53	1.42	1.36	1.49
	(0.44)	(0.42)	(0.53)	(0.52)	(0.50)
Previous Instability	1.02	0.99	0.83	0.83	0.86
	(0.08)	(0.08)	(0.10)	(0.10)	(0.13)
Neighboring Democracies$_{(t-1)}$	0.98**	0.98**	0.98**	0.98**	0.98**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Total Population$_{(log,\ t-1)}$	0.89	0.92	0.87	0.88	0.86
	(0.08)	(0.09)	(0.11)	(0.11)	(0.11)
Urbanization$_{(t-1)}$	0.10*	0.12	0.01**	0.01**	0.04
	(0.13)	(0.16)	(0.02)	(0.01)	(0.10)
NVR		0.46**	0.46**	0.45**	0.39**
		(0.12)	(0.15)	(0.14)	(0.13)
Observations	1448	1448	1133	1019	1110
Regimes	112	112	74	68	74
Democratic Breakdowns	69	69	38	36	39
AIC	553.30	548.17	265.66	247.14	272.25
BIC	584.97	585.12	300.89	281.63	307.34
Log lik.	-270.65	-267.09	-125.83	-116.57	-129.13
Chi-squared	39.62	53.83	56.55	53.49	65.55

*Hazard Ratios, robust standard errors clustered by regime spell in parentheses; ** p<0.05, * p<0.1*

As shown in table A4.1, NVR has a significant and substantial negative effect on the hazard of democratic breakdown. Depending on the sample, democratic regimes where the transition process was induced by an NVR campaign are at least 54% and at most 61% less likely to break down. A crucial requirement for the validity of the results of the Cox-model is the proportional hazard assumption. The assumption implies that covariate effects are constant over time, i.e. a constant relative hazard. To detect potential violations of the proportional hazard assumption scholars usually employ a so called Grambsch-Therneau test (Grambsch and Therneau, 1994). The test is based on the Schoenfeld residuals, which are calculated from the Cox-model. For each covariate as well as the full model, the procedure tests the null hypothesis of zero slope, i.e. if the log hazard-ratio function is constant over time. To address concerns about the arbitrary use of time transformations for Grambsch-Therneau tests (Park & Hendry, 2015), we conducted tests with four different transformations of time: a linear untransformed version, a logarithmic transformation, a left-continuous transformation of the Kaplan-Meier survival curve and a transformation using the observed event times placed in integer-rank order. For the models reported in table A4.1 we did not find any violation of the proportional hazard assumption.[3] Now table A4.2 reports the Cox models for the Geddes data.

Table A4.2. Cox-regression models (Geddes data)

	Full sample	Full sample	Greedy sample	Caliper sample	Optimal sample
GDP p.c.$_{(log, t-1)}$	0.78	0.76	0.77	0.85	0.69
	(0.15)	(0.16)	(0.18)	(0.22)	(0.18)
Military Legacy	1.80	1.28	1.25	1.31	1.33
	(0.65)	(0.49)	(0.52)	(0.64)	(0.57)
Previous Instability	0.99	0.98	0.96	0.90	0.93
	(0.07)	(0.08)	(0.09)	(0.10)	(0.09)
Neighboring Democracies$_{(t-1)}$	0.98**	0.98**	0.98**	0.98**	0.98**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Total Population$_{(log, t-1)}$	1.01	1.12	1.02	1.17	1.03
	(0.14)	(0.16)	(0.17)	(0.27)	(0.17)
Urbanization$_{(t-1)}$	0.06	0.16	0.13	0.20	0.21
	(0.10)	(0.28)	(0.25)	(0.52)	(0.41)
NVR		0.13**	0.15**	0.13**	0.15**
		(0.09)	(0.09)	(0.13)	(0.09)
NVR*time		1.14**	1.12**	1.22*	1.11**
		(0.07)	(0.06)	(0.14)	(0.06)
Observations	1165	1165	1051	799	1066
Regimes	89	89	78	58	78
Democratic Breakdowns	39	39	33	23	33
AIC	310.90	301.93	244.20	162.41	244.18
BIC	341.26	342.41	283.86	199.87	283.96
Log lik.	-149.45	-142.96	-114.10	-73.20	-114.09
Chi-squared	17.88	43.76	50.85	34.57	50.54

*Hazard Ratios, robust standard errors clustered by regime spell in parentheses; ** p<0.05, * p<0.1*

As shown in table A4.2, we estimated all models with the treatment variable indicating the presence of NVR and an interaction term between NVR and survival. Our tests yielded a potential

[3] The detailed results of the Grambsch-Therneau tests are reported in the accompanying log file.

violation of the proportional hazard assumption for the treatment variable when using the Geddes data. To address this issue, we estimated the models with NVR as a time-dependent covariate, i.e. including an interaction term between NVR and survival time in the model. The results indicate a significant and substantial treatment effect across all models, which is however, diminishing over time. For instance, the results for the full sample indicate that immediately after the transition, NVR reduces the risk of democratic breakdown by 87%. However, the significant interaction term implies that each following year this effect is reduced by 14%. Next, we report the results for the Cox-models using the Boix data in table A4.3. As shown in table A4.3, we estimated all models with total population and urbanization as time-dependent covariates, because the respective tests indicated a potential violation of the proportional hazard assumption for these variables. The effect of NVR is again significant and substantial across all models, indicating the democratic regimes that resulted from NVR induced transition are between 53% and 72% less likely to experience democratic breakdown relative to regimes without this feature.

Table A4.3. Cox-regression models (Boix data)

	Full sample	Full sample	Greedy sample	Caliper sample	Optimal sample
GDP p.c.$_{(log,\ t-1)}$	0.58**	0.57**	0.60**	0.54**	0.59**
	(0.11)	(0.11)	(0.14)	(0.13)	(0.14)
Military Legacy	2.81**	2.63**	3.57**	7.08**	3.81**
	(0.92)	(0.86)	(1.44)	(3.38)	(1.51)
Previous Instability	0.97	0.95	0.95	0.90	0.94
	(0.10)	(0.10)	(0.14)	(0.13)	(0.13)
Neighboring Democracies$_{(t-1)}$	0.99	0.99	0.99	0.99	0.99
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Total Population$_{(log,\ t-1)}$	0.92	0.93	1.03	0.74	1.00
	(0.19)	(0.19)	(0.31)	(0.26)	(0.31)
Total Population$_{(log,\ t-1)}$*time	1.03	1.03	0.99	1.00	1.00
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Urbanization$_{(t-1)}$	1.23	1.89	2.30	2.68	2.04
	(2.65)	(4.34)	(6.82)	(11.23)	(6.39)
Urbanization$_{(t-1)*time}$	0.76*	0.74*	0.58**	0.45**	0.62*
	(0.12)	(0.12)	(0.15)	(0.17)	(0.18)
NVR		0.47**	0.34**	0.28**	0.32**
		(0.15)	(0.12)	(0.12)	(0.11)
Observations	1554	1554	1264	901	1222
Regimes	101.00	101.00	82.00	60.00	82.00
Democratic Breakdowns	47.00	47.00	37.00	32.00	38.00
AIC	360.88	357.99	258.66	199.71	262.55
BIC	403.67	406.13	304.94	242.94	308.53
Log lik.	-172.44	-170.00	-120.33	-90.85	-122.28
Chi-squared	64.00	60.81	54.72	43.32	58.37

*Hazard Ratios, robust standard errors clustered by regime spell in parentheses; ** p<0.05, * p<0.1*

As an additional robustness test that addresses potential omitted variable bias, we estimated Cox proportional hazards models with time-varying covariates and shared regime frailties. These are essentially random effects regression models for survival data. Across regimes, the frailties are assumed to be gamma-distributed latent random effects that affect the hazard. The results are reported in table A4.4.

Table A4.4. Cox- frailty models

	Ulfelder	Geddes	Boix
GDP p.c.$_{(log, t-1)}$	0.55^{**}	0.76	0.57^{**}
	(0.12)	(0.17)	(0.11)
Military Legacy	1.60	1.28	2.63^{**}
	(0.59)	(0.48)	(0.93)
Previous Instability	1.03	0.98	0.95
	(0.10)	(0.08)	(0.11)
Neighboring Democracies$_{(t-1)}$	0.98^{**}	0.98^{**}	0.99
	(0.01)	(0.01)	(0.01)
Total Population$_{(log, t-1)}$	0.92	1.12	0.93
	(0.12)	(0.17)	(0.17)
Urbanization$_{(t-1)}$	0.14	0.16	1.89
	(0.24)	(0.30)	(4.21)
NVR	0.29^{**}	0.13^{**}	0.47^{**}
	(0.11)	(0.09)	(0.17)
NVR*time		1.14^{*}	
		(0.08)	
Total Population$_{(log, t-1)}$*time			1.03^{*}
			(0.02)
Urbanization$_{(t-1)}$*time			0.74^{*}
			(0.12)
Observations	1448	1165	1554
Regimes	112	89	101
Democratic Breakdowns	69	39	47
AIC	539.06	301.93	357.99
BIC	576.01	342.41	406.13
Log lik.	-262.53	-142.96	-170.00
Chi-squared	41.33	24.01	35.26

*Hazard Ratios, robust standard errors clustered by regime spell in parentheses; ** $p<0.05$, * $p<0.1$*

As shown in table A4.4, the estimated hazard ratios are similar to those reported above for the Cox-models without regime frailties. Again, the results indicate that NVR induced transitions reduce the subsequent hazard of democratic breakdown substantially.

A5: Analysis of peaceful turnovers of power

We also analyzed the occurrence of peaceful turnovers of power using three different regime data sets and the matched samples. For comparison, figures A5.1 and A5.2 reproduce the figures from the main text that utilize the full sample of the Ulfelder data. In figures A5.3 – A5.8, we describe the results for the matched samples. Because we implemented 1 to 1 matching, where each treated case is matched to one control case, the results for the regimes where transition was induced by NVR are often the same. While the group of non-treated cases without NVR always changes across the different samples, changes to the group of NVR-cases occur only if for a given matched sample, some of the treated units fall outside the region of common support or if treated cases are dropped due to the restriction of caliper matching. For the Ulfelder data changes did not occur due to lack of common support, but in the sample generated via caliper matching three treated cases where dropped because of the caliper restriction.

As shown in figures A5.3 – A5.8, with regard to the first turnover there appears to be no substantial difference between regimes that where induced by NVR and those without this feature. However, across all samples, regimes that achieved the first turnover more frequently achieve the second turnover, if they were induced by NVR. Whereas the chances are basically a coinflip for regimes without NVR, almost all NVR cases achieve the second turnover.

Next, we report the same analysis describing frequencies of peaceful turnover of power for the Geddes data in figures A5.9-A5.16. Using the Geddes data, we also find, that across all samples there appears to be no systematic difference between regimes that where induced by NVR and those without this feature with regard to the frequency of the achieving the first peaceful turnover of power. However, again there appears to be a difference related to the second peaceful turnover. Whereas regimes without NVR induced transition that achieved the first turnover very frequently also achieve the second turnover of power, the chances for regimes without much lower, i.e. about 50% in most samples

Finally, we report the analysis of turnover frequencies for the Boix data in figures A5.17-A5.24. For the Boix, data we also find no substantial effect of NVR on the first turnover, but again those NVR induced regimes that achieved the first turnover are more likely to achieve the second turnover compared to regimes that came about without the help of NVR. However, with the Boix data, the difference between NVR and non-NVR regimes is less pronounced than when using the other datasets. For example, in the full sample 14 out of 24 regimes without NVR (58%) and 19 out of 22 regimes with NVR (86%) achieved the second turnover. Thus, regimes with NVR have a 28% higher chance to achieve the second turnover than regimes without NVR. Now, when we consider the sample created by greedy matching, 12 out of 17 regimes without NVR (70%) and 19 out of 22 regimes with NVR (86%) achieved the second turnover. Now, the difference is much smaller, i.e. 16%.

Figure A5.1: Turnovers for regimes without NVR (Ulfelder, full sample)

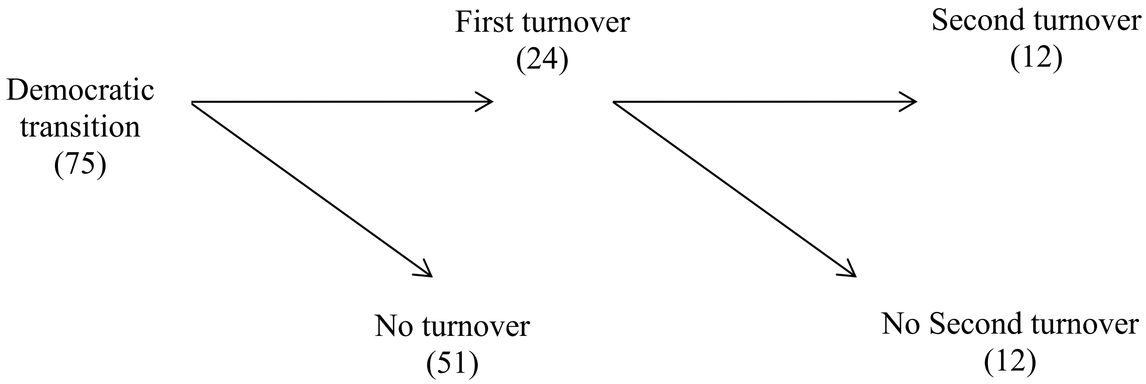


Figure A5.2: Turnovers for regimes with NVR (Ulfelder, full sample)

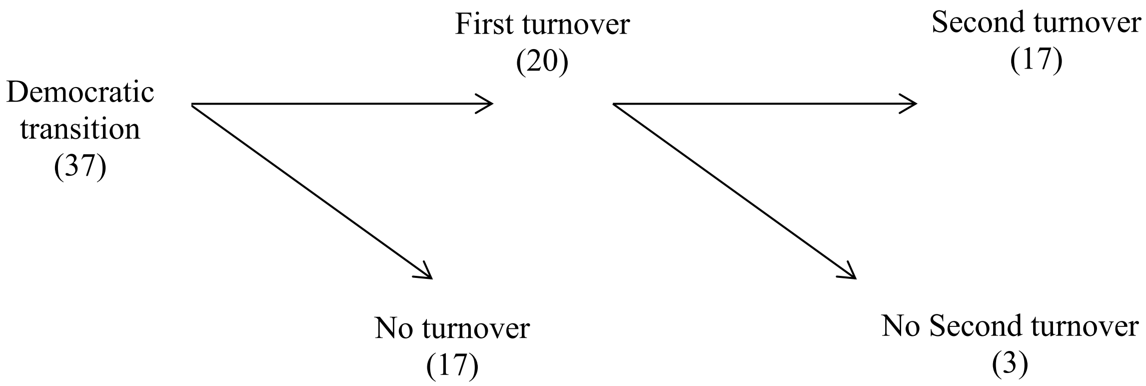


Figure A5.3: Turnovers for regimes without NVR (Ulfelder, greedy sample)

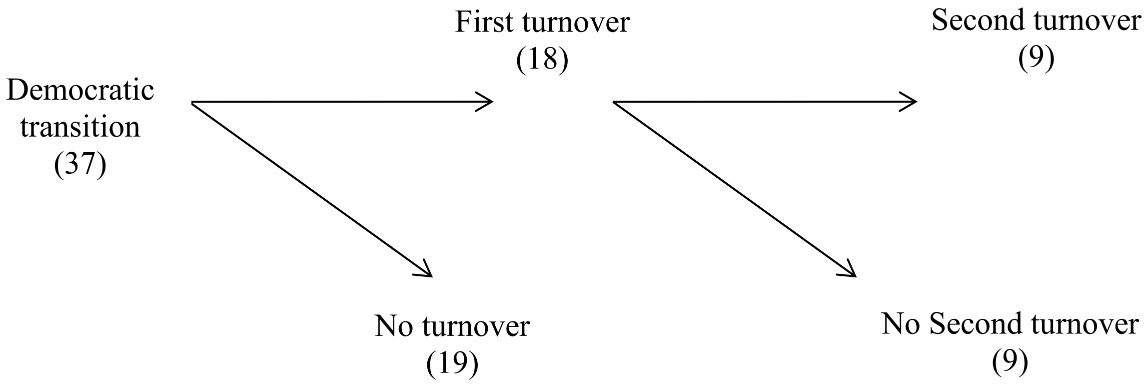


Figure A5.4: Turnovers for regimes with NVR (Ulfelder, greedy sample)

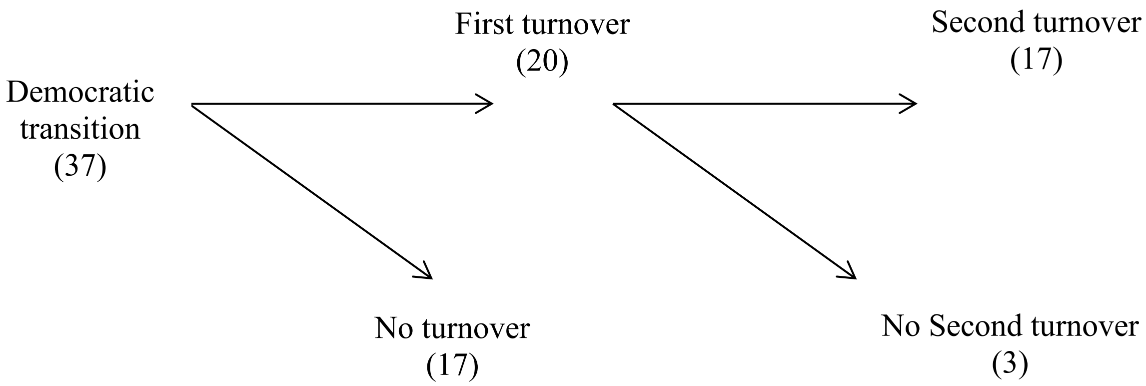


Figure A5.5: Turnovers for regimes without NVR (Ulfelder, caliper sample)

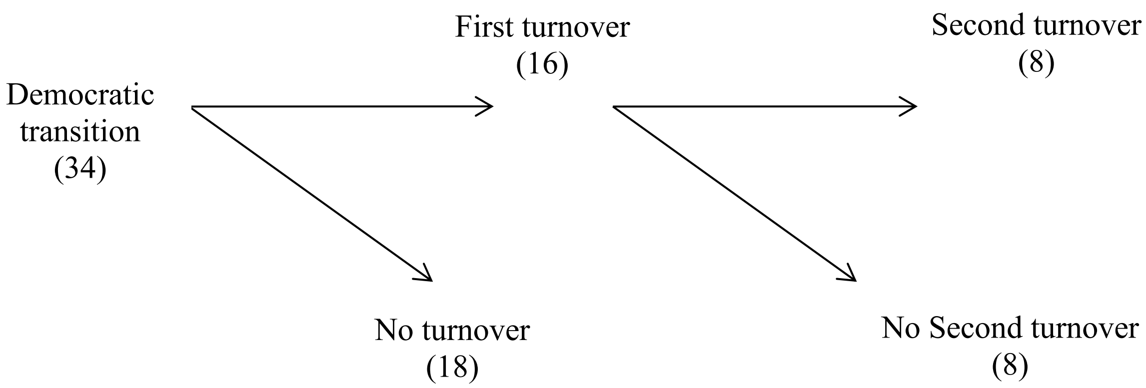


Figure A5.6: Turnovers for regimes with NVR (Ulfelder, caliper sample)

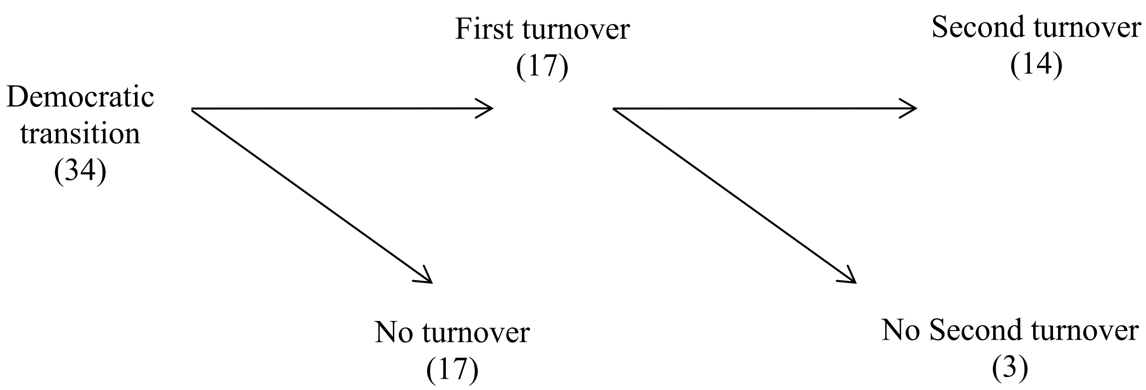


Figure A5.7: Turnovers for regimes without NVR (Ulfelder, optimal sample)

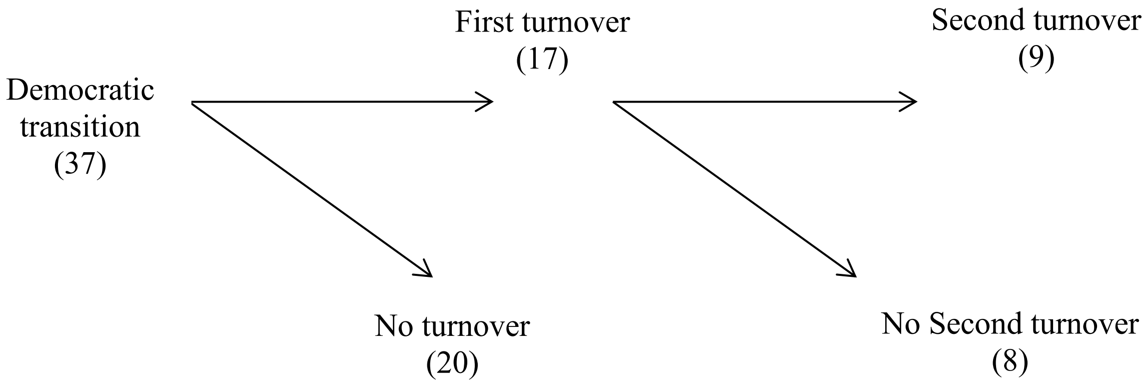


Figure A5.8: Turnovers for regimes with NVR (Ulfelder, optimal sample)

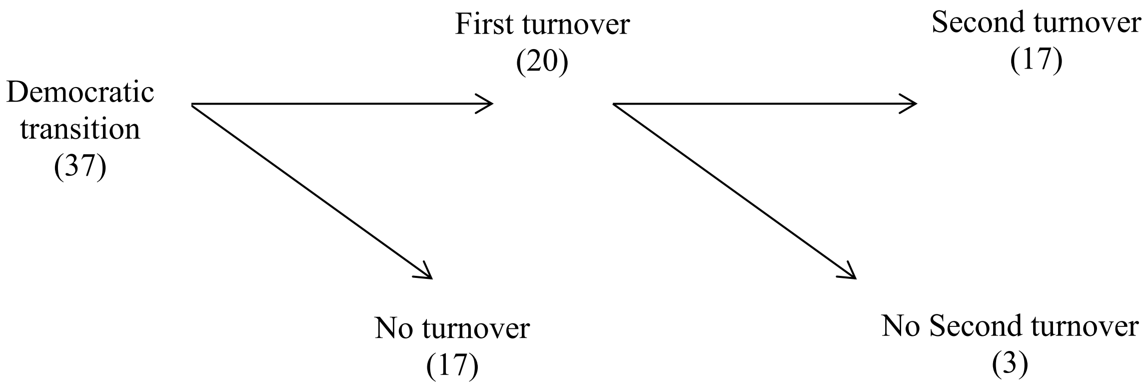


Figure A5.9: Turnovers for regimes without NVR (Geddes, full sample)

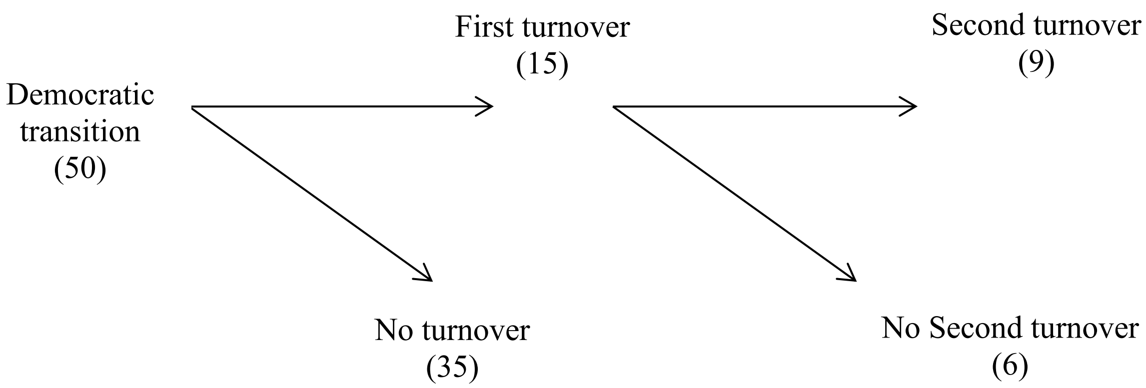


Figure A5.10: Turnovers for regimes with NVR (Geddes, full sample)

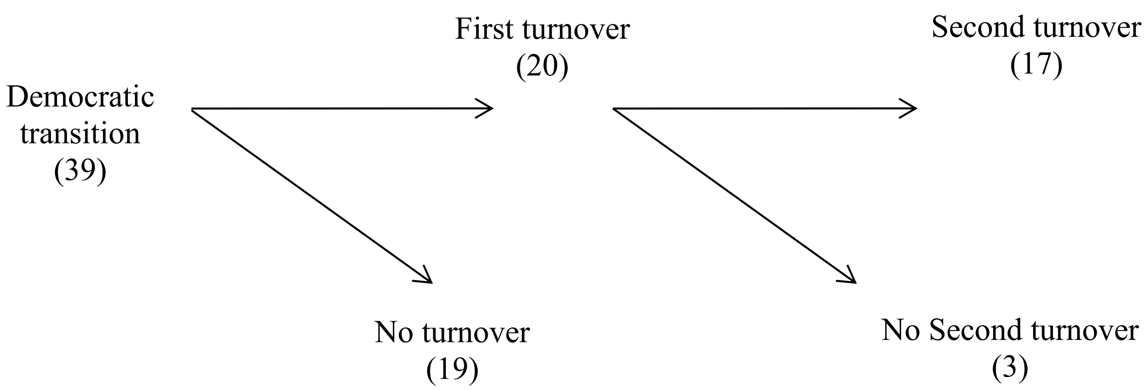


Figure A5.11: Turnovers for regimes without NVR (Geddes, greedy sample)

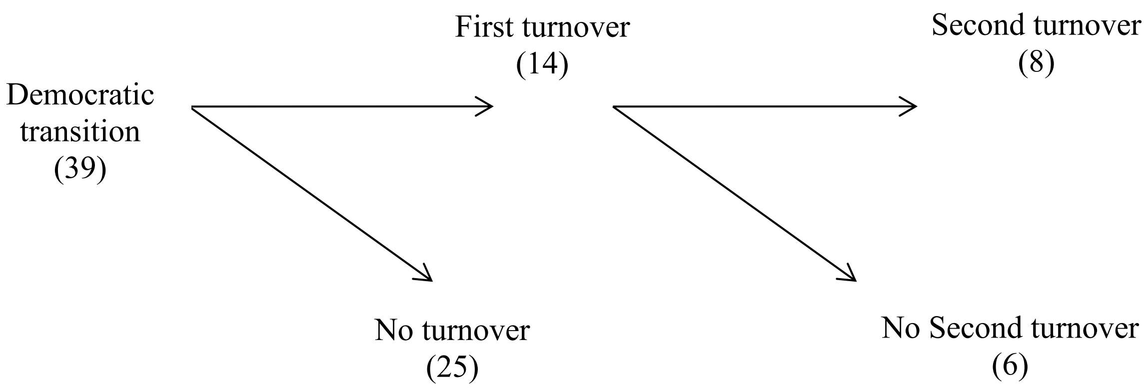


Figure A5.12: Turnovers for regimes with NVR (Geddes, greedy sample)

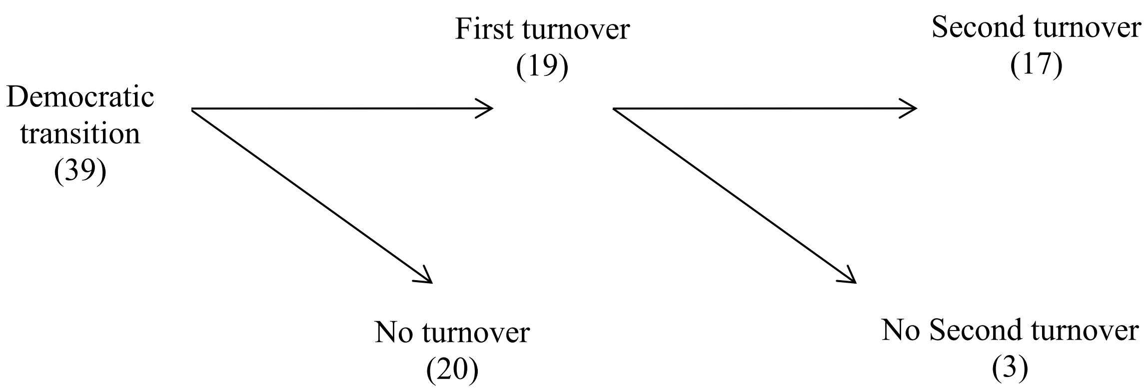


Figure A5.13: Turnovers for regimes without NVR (Geddes, caliper sample)

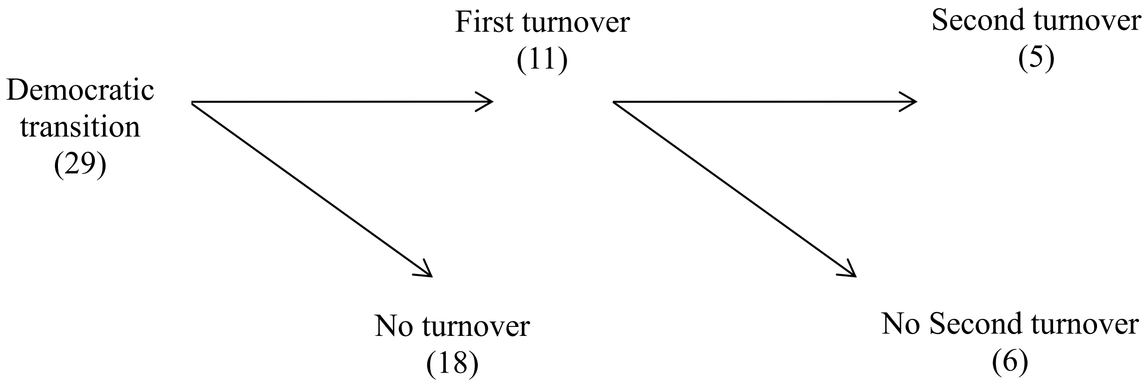


Figure A5.14: Turnovers for regimes with NVR (Geddes, caliper sample)

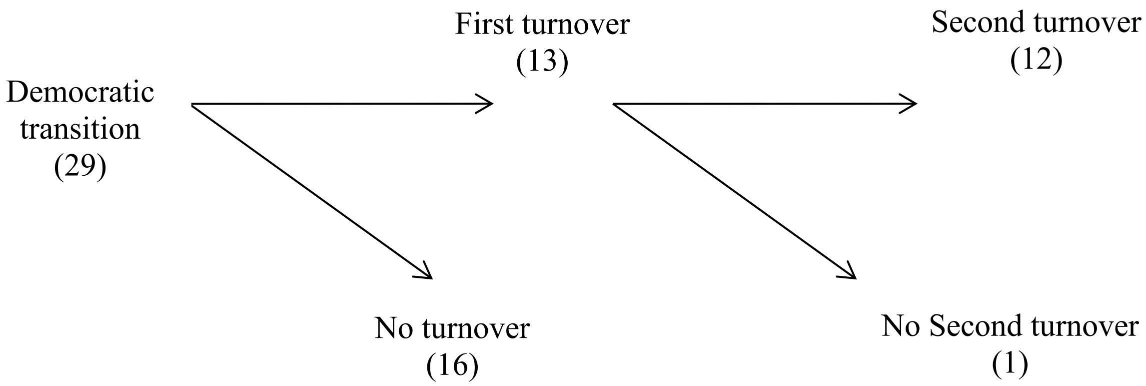


Figure A5.15: Turnovers for regimes without NVR (Geddes, optimal sample)

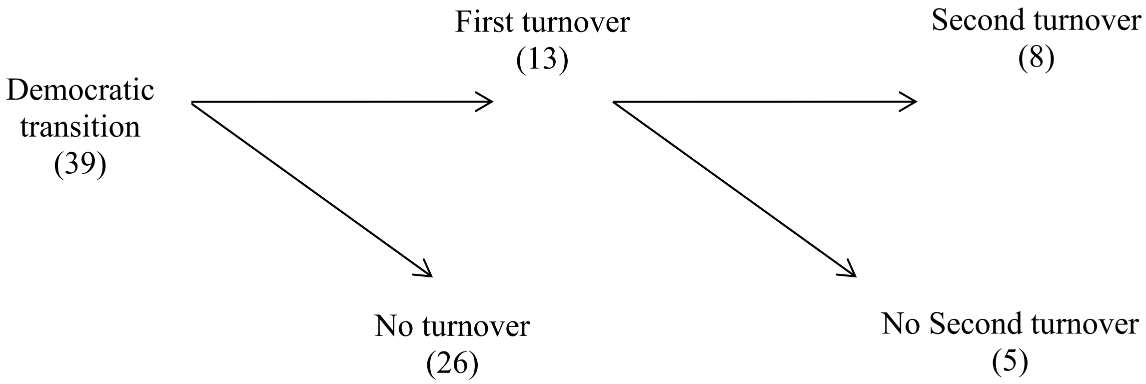


Figure A5.16: Turnovers for regimes with NVR (Geddes, optimal sample)

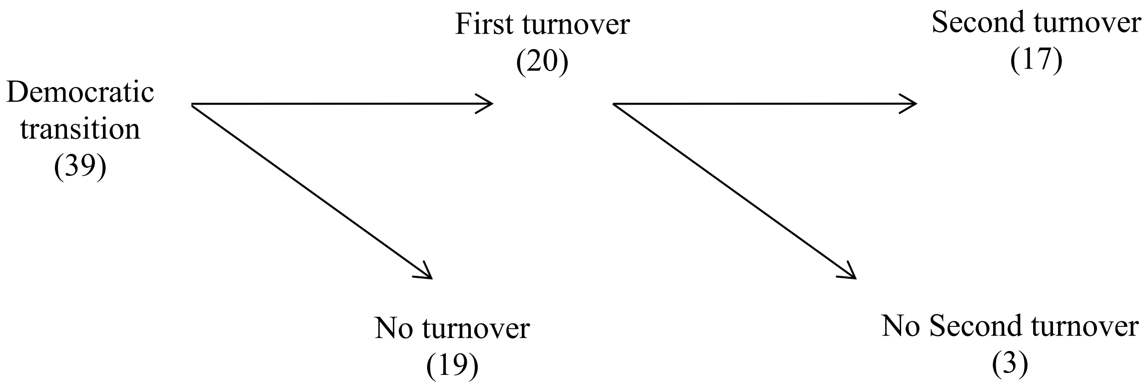


Figure A5.17: Turnovers for regimes without NVR (Boix, full sample)

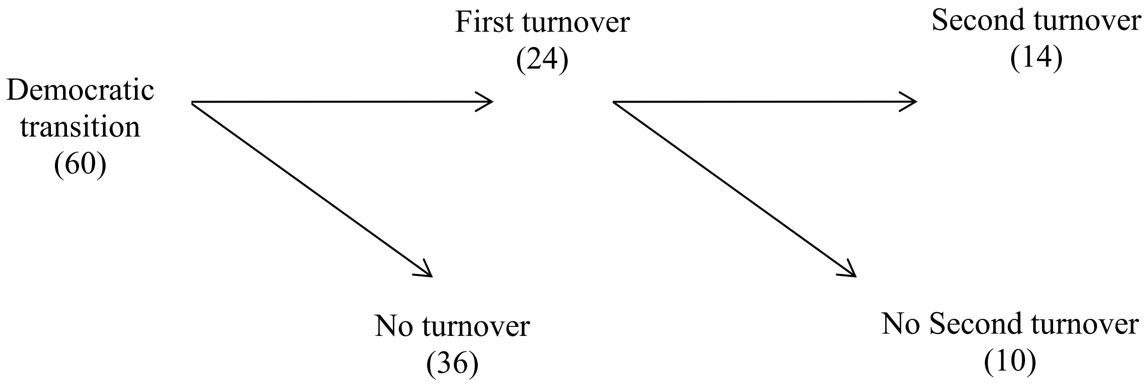


Figure A5.18: Turnovers for regimes with NVR (Boix, full sample)

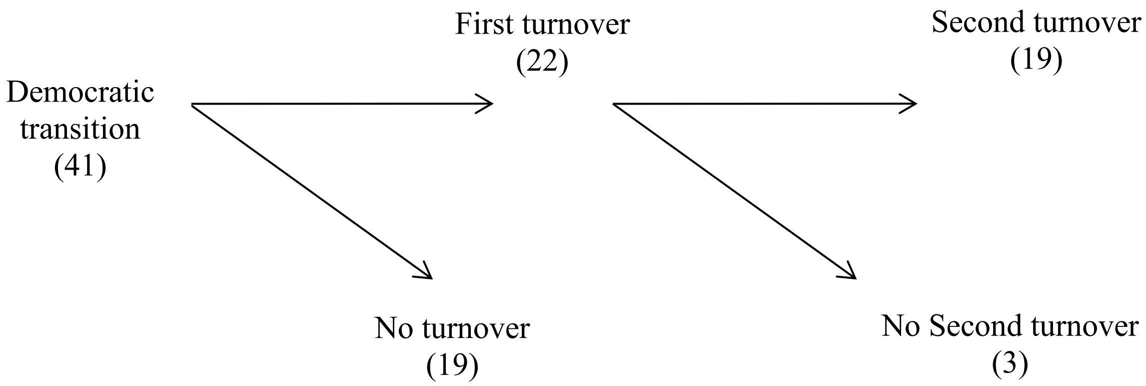


Figure A5.19: Turnovers for regimes without NVR (Boix, greedy sample)

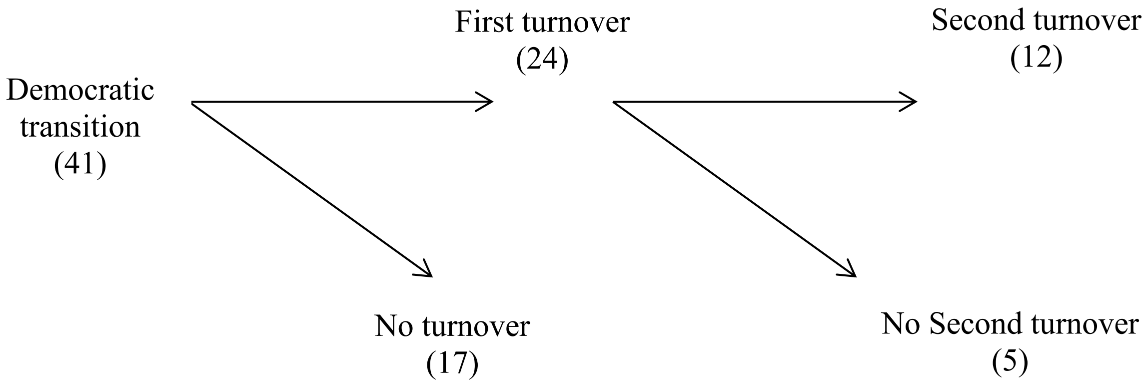


Figure A5.20: Turnovers for regimes with NVR (Boix, greedy sample)

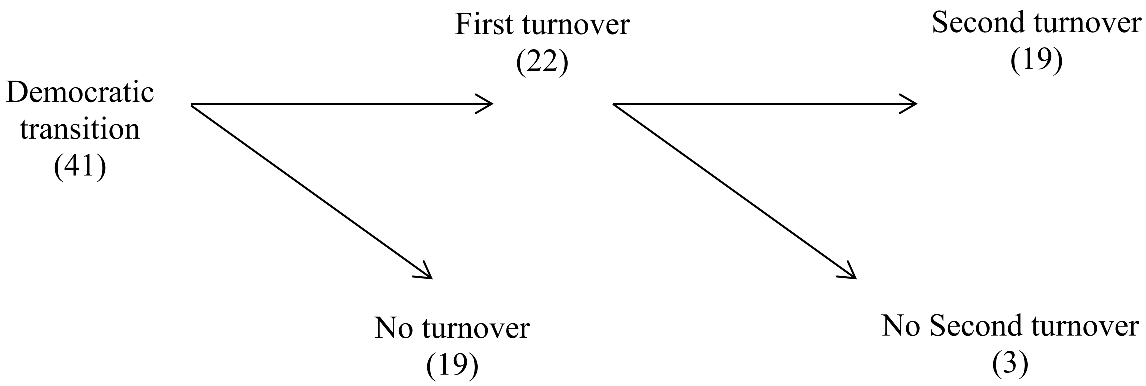


Figure A5.21: Turnovers for regimes without NVR (Boix, caliper sample)

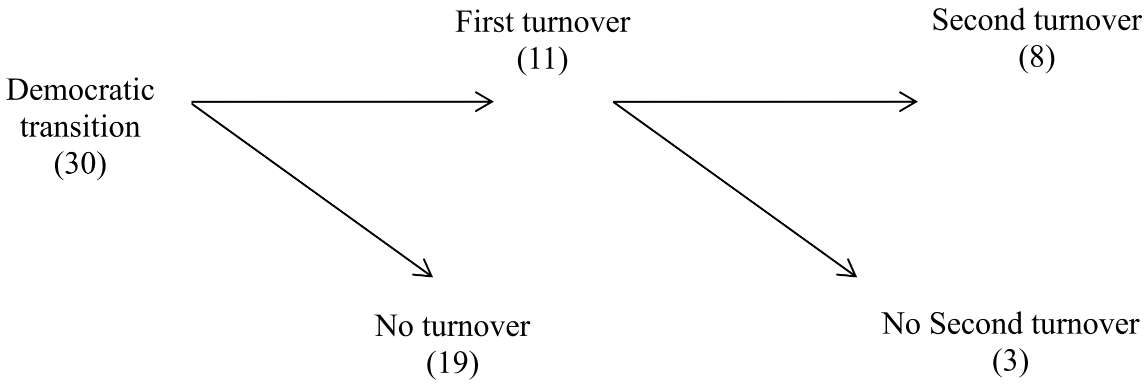


Figure A5.22: Turnovers for regimes with NVR (Boix, caliper sample)

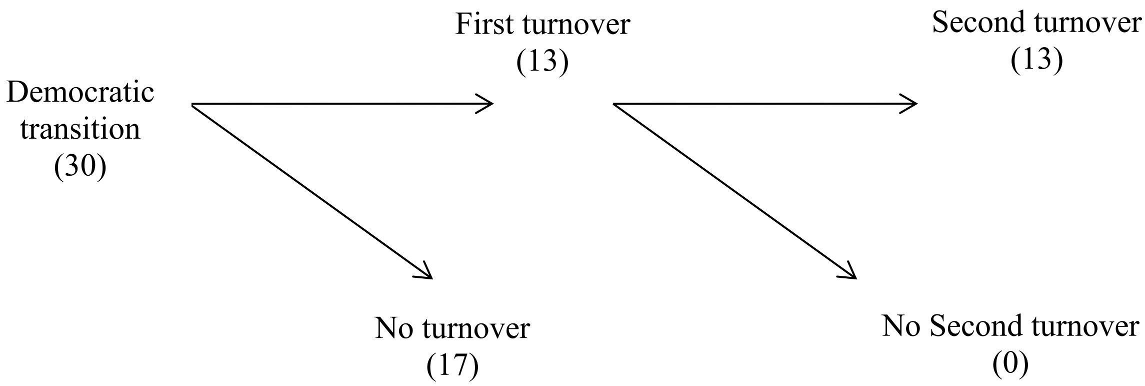


Figure A5.23: Turnovers for regimes without NVR (Boix, optimal sample)

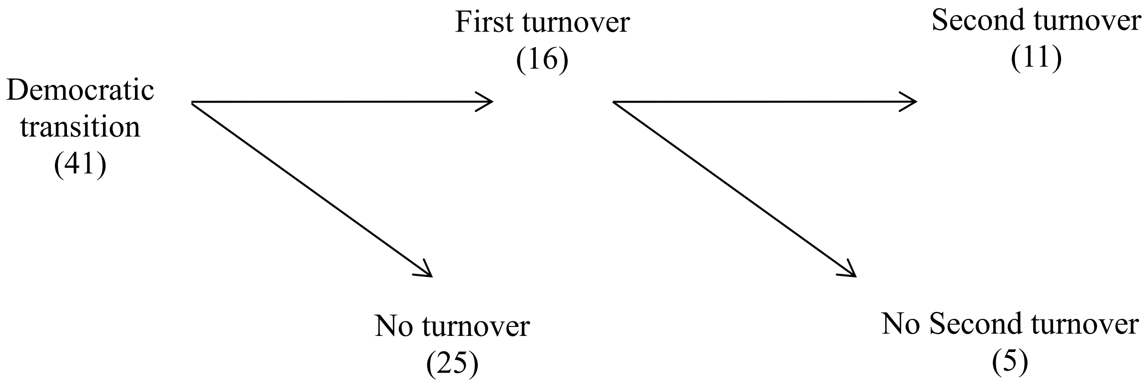
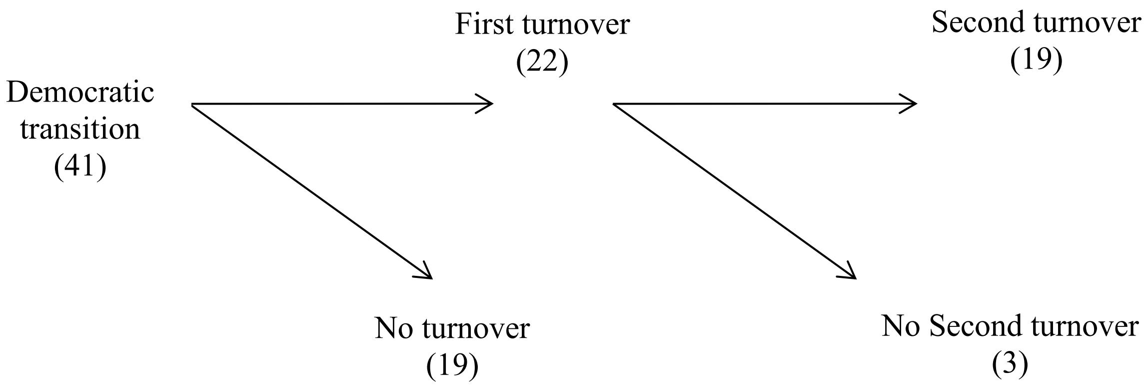


Figure A5.24: Turnovers for regimes with NVR (Boix, optimal sample)



Sequential logit models

To account for alternative explanations and confounding factors, we analyze the effect of NVR on peaceful turnovers of power with a sequential logit model (Buis 2013). The regression model accounts for the sequential nature of the dependent variable by estimating a separate logit for the first and second turnover. We implement the sequential logit model with a cross-sectional data structure and covariates measured at baseline, i.e. one year before the transition event occurred. The results of the sequential logit model using the Ulfelder data are reported in table A5.1. The upper part of the table explains the first and the lower part the second turnover of power, respectively.

Table A5.1. Sequential logit models (Ulfelder data)

	Full sample	Greedy sample	Caliper sample	Optimal sample
First turnover				
Neighboring Democracies$_{(t-1)}$	-0.01	-0.00	0.00	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)
GDP p.c.$_{(log, t-1)}$	0.85**	0.36	0.16	0.78**
	(0.31)	(0.37)	(0.40)	(0.37)
Urbanization$_{(t-1)}$	1.55	4.55*	5.26**	2.91
	(1.92)	(2.42)	(2.66)	(2.05)
Total Population$_{(log, t-1)}$	0.17	0.18	0.21	0.23
	(0.18)	(0.22)	(0.23)	(0.24)
Previous Instability	0.09	0.33*	0.33*	0.24
	(0.13)	(0.20)	(0.19)	(0.17)
Military Legacy	-0.15	0.05	0.17	-0.04
	(0.52)	(0.62)	(0.68)	(0.54)
NVR	0.64	0.20	0.23	0.41
	(0.43)	(0.50)	(0.53)	(0.51)
Constant	-9.26**	-6.40*	-5.55	-9.70**
	(2.90)	(3.71)	(3.85)	(4.07)
Second turnover				
Neighboring Democracies$_{(t-1)}$	-0.01	-0.01	-0.01	-0.01
	(0.02)	(0.02)	(0.02)	(0.02)
GDP p.c.$_{(log, t-1)}$	0.70	1.04**	1.12**	0.89*
	(0.48)	(0.50)	(0.52)	(0.54)
Urbanization$_{(t-1)}$	1.50	-0.66	-2.16	-0.60
	(2.99)	(2.66)	(3.46)	(2.67)
Total Population$_{(log, t-1)}$	-0.19	-0.12	-0.08	-0.17
	(0.34)	(0.34)	(0.33)	(0.33)
Previous Instability	0.24	0.27	0.31*	0.36*
	(0.18)	(0.17)	(0.18)	(0.19)
Military Legacy	-0.62	-0.93	-0.63	-1.33
	(0.98)	(1.06)	(1.09)	(1.15)
NVR	1.55**	1.64*	1.55*	1.61*
	(0.77)	(0.86)	(0.93)	(0.88)
Constant	-4.25	-7.26*	-8.28*	-5.65
	(3.52)	(4.05)	(4.34)	(4.38)
Observations	112	74	68	74

Raw coefficients with standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$

The results show that NVR during democratic transition is positively associated with achieving the first turnover of power, but the estimates of the effect are not statistically significant. However, the estimates for the effect of NVR on the second turnover of power are significant indicating that, given a successful first turnover, NVR improves the probability of accomplishing a second turnover of power. Using the estimates from the full sample, we calculated the marginal effect of NVR on achieving the second turnover. The results indicate that achieving democratic transition by means of NVR improves the probability of a second turnover by 38% relative to regimes that democratized without NVR. However, as shown in table A5.1, the effect estimates are not robust across samples, i.e. with the matched samples, the effect is only significant at $p < 0.1$. However, the results for the matched samples should be treated with caution, because due to the sequential nature of the data, the samples on which the estimates for the second turnover are based, are very small. Now table A5.2 reports the results of the sequential logit model using the Geddes data.

Table A5.2. Sequential logit models (Geddes data)

	Full sample	Greedy sample	Caliper sample	Optimal sample
First turnover				
Neighboring Democracies$_{(t-1)}$	0.01	0.01	0.02	0.00
	(0.01)	(0.01)	(0.01)	(0.01)
GDP p.c.$_{(log, t-1)}$	0.54	0.57*	0.70*	0.76**
	(0.35)	(0.33)	(0.39)	(0.35)
Urbanization$_{(t-1)}$	4.31*	3.60	3.12	2.79
	(2.29)	(2.21)	(2.82)	(2.13)
Total Population$_{(log, t-1)}$	0.10	0.03	-0.34	0.07
	(0.24)	(0.28)	(0.35)	(0.27)
Previous Instability	-0.01	0.04	0.13	0.04
	(0.13)	(0.13)	(0.18)	(0.13)
Military Legacy	0.42	0.74	0.82	0.47
	(0.57)	(0.59)	(0.74)	(0.56)
NVR	0.70	0.49	0.13	0.64
	(0.55)	(0.56)	(0.61)	(0.57)
Constant	-7.28**	-6.91**	-4.77	-8.53**
	(3.28)	(3.19)	(3.84)	(3.37)
Second turnover				
Neighboring Democracies$_{(t-1)}$	-0.02	-0.02	-0.03	-0.02
	(0.02)	(0.02)	(0.03)	(0.02)
GDP p.c.$_{(log, t-1)}$	1.48**	1.44**	2.76**	1.26*
	(0.60)	(0.59)	(1.21)	(0.76)
Urbanization$_{(t-1)}$	-6.76*	-6.46	-7.03	-5.95
	(4.03)	(3.98)	(6.08)	(4.13)
Total Population$_{(log, t-1)}$	-0.53	-0.50	-1.09	-0.48
	(0.43)	(0.44)	(1.10)	(0.43)
Previous Instability	0.14	0.12	0.14	0.11
	(0.18)	(0.17)	(0.25)	(0.19)
Military Legacy	1.25	1.15	1.99	1.21
	(1.16)	(1.19)	(1.79)	(1.22)
NVR	1.94**	1.94**	3.28**	1.84**
	(0.95)	(0.92)	(1.11)	(0.93)
Constant	-5.24	-5.21	-10.98	-3.98
	(5.91)	(5.81)	(7.27)	(7.10)
Observations	89	78	58	78

Raw coefficients with standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$

As shown in table A5.2, using the Geddes data, we obtain similar results as with the Ulfelder data. None of the models estimates a significant effect of NVR on the first turnover. However, the effect of NVR on the second turnover is significant across all samples. Marginal effect estimates for the full sample indicate that NVR improves the probability of a second turnover by 34%. Finally, table A5.3 reports the results for the Boix data.

Table A5.3. Sequential logit models (Boix data)

	Full sample	Greedy sample	Caliper sample	Optimal sample
First turnover				
Neighboring Democracies$_{(t-1)}$	-0.02	-0.01	-0.02	-0.01
	(0.01)	(0.01)	(0.02)	(0.01)
GDP p.c.$_{(log, t-1)}$	1.04**	1.27**	1.48**	1.28**
	(0.42)	(0.43)	(0.53)	(0.42)
Urbanization$_{(t-1)}$	3.00	1.77	0.16	1.81
	(2.19)	(2.35)	(2.87)	(2.30)
Total Population$_{(log, t-1)}$	-0.07	0.06	0.12	0.06
	(0.19)	(0.21)	(0.28)	(0.21)
Previous Instability	0.07	0.24	0.24	0.24
	(0.18)	(0.21)	(0.22)	(0.21)
Military Legacy	-0.13	-0.16	-0.38	-0.26
	(0.65)	(0.66)	(0.82)	(0.67)
NVR	0.19	0.33	0.09	0.40
	(0.49)	(0.57)	(0.64)	(0.57)
Constant	-7.86**	-11.13**	-12.62**	-11.22**
	(3.17)	(3.55)	(4.13)	(3.50)
Second turnover				
Neighboring Democracies$_{(t-1)}$	-0.02	-0.02	2.44**	-0.02
	(0.02)	(0.02)	(0.11)	(0.02)
GDP p.c.$_{(log, t-1)}$	1.66**	1.64**	74.50**	1.67**
	(0.52)	(0.80)	(3.78)	(0.79)
Urbanization$_{(t-1)}$	1.23	-1.59	-42.12**	-1.64
	(3.56)	(4.30)	(11.54)	(4.17)
Total Population$_{(log, t-1)}$	-0.33	-0.84**	-7.70**	-0.83**
	(0.31)	(0.40)	(0.58)	(0.40)
Previous Instability	-0.27	-0.32	-13.13**	-0.31
	(0.20)	(0.23)	(1.07)	(0.23)
Military Legacy	0.51	0.80	-38.66**	0.60
	(1.04)	(1.05)	(1.56)	(1.14)
NVR	1.63*	1.70*	43.59**	1.81*
	(0.88)	(1.02)	(1.44)	(1.05)
Constant	-9.33**	-3.11	-495.49**	-3.39
	(3.50)	(5.82)	(25.58)	(5.70)
Observations	101	82	60	82

Raw coefficients with standard errors in parentheses, * *p < 0.10,* ** *p < 0.05*

With the Boix data, we also find no significant effect of NVR on the first turnover. The effect of NVR on the second turnover is only significant at $p < 0.5$ when using the sample obtained via caliber matching. For this sample coefficient estimates for NVR are unusually large compared to the estimates from the other samples. This is due to the fact that, as shown in figure A5.22, all regimes that where induced by NVR and achieved the first turnover also achieved the second turnover.

In sum, the results from the empirical analysis of the effect of NVR on achieving peaceful turnovers of power generated two main insights. First, our results reveal that there is most likely no substantial effect of NVR on achieving the first turnover of power. Second, our results indicate that given a successful first turnover, NVR induced regimes more frequently also achieve the second turnover, relative to regimes that democratized without NVR. However, due to the fact that this second finding is not robust across different datasets and matching procedures, we remain skeptical with regard to the causal interpretation of this effect.

A6: Analysis of democratic quality

Finally, we evaluate if and how democratic regimes that evolved from NVR induced transitions and those without this feature differ in their development of democratic quality after transition. Specifically, we analyze the level of democratic quality that regimes achieve after transition and the degree of improvement from pre- to post transition until up to 10 years after the transition occurred. We consider only regimes that survived until the respective point in time, i.e. avoided democratic breakdown.

To measure democratic quality, we use a polyarchy index compiled by Teorell et al. (2016).[4] This measure of democratic quality is based on expert coding by more than 2,600 country experts,

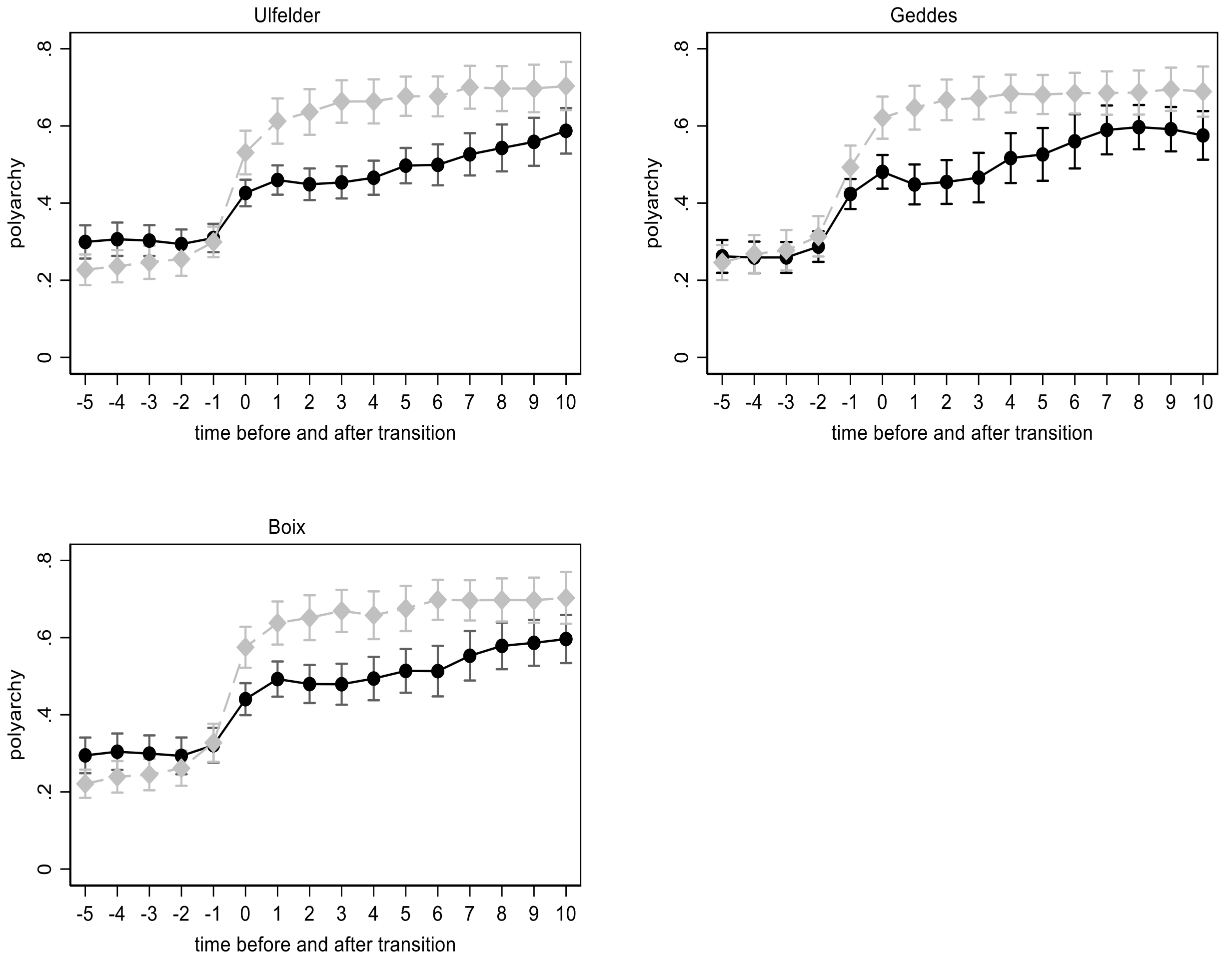
[4] Specifically, we use the variable “v2x_polyarchy” from the Varieties of Democracy (VDEM) dataset version 6.2. For details see Coppedge et al. (2016).

and gathers hundreds of indicators for almost every country in the world. The polyarchy index is based on the following five components: (1) elected officials, (2) free and fair elections, (3) freedom of expression, (4) associational autonomy, and (5) inclusive citizenship.

The first dimension, "elected officials," evaluates how the chief executive is elected. Depending on the system of government it also uses information on other political institutions such as the proportion of legislators that is elected. The second dimension, "free and fair elections," addresses whether elections can be considered free and fair, which refers to an absence of registration fraud, systematic irregularities, government intimidation of the opposition, vote buying, and election violence. The third dimension, "freedom of expression", addresses to what extent a government respects press and media freedom, the freedom of expression for ordinary citizens, as well as the freedom of academic and cultural expression. The fourth dimension, "freedom of organization," measures freedom of association for political parties and civil society organizations". Finally, the fifth dimension, "inclusive citizenship", relates to suffrage and captures the share of adult citizens that have the legal right to vote in national elections.

Expert ratings for these dimensions are combined to an index that ranges from zero to one, with higher values indicating a higher quality of democracy. Using the full sample of the three regime datasets, figure A6.1 reports the average pre- and post-transition trends in polyarchy levels for the regimes with NVR and without NVR.

Figure A6.1. Average levels of democracy before and after transition



The gray line and markers represent the level the polyarchy index for regimes where the transition was induced by NVR and the black line and markers refer to regimes without NVR induced transition.

The time trend in the average level of polyarchy appears to be similar in the three datasets. Before the accomplishment of the democratic transition, regimes that eventually democratized by means of NVR and regimes without this characteristic do not significantly differ in their average polyarchy scores. However, after the transition event, regimes that were induced by NVR achieve a much higher level of democratic quality than regimes without this feature. Depending on the dataset, a significant and substantial difference in average levels of democratic quality can be observed for at least six years and at most nine years after transition. Additionally, figure A6.1 highlights that across all three datasets, the average polyarchy score improves more substantially from pre- to post transition for regimes that where induced by NVR than for regimes without this feature. Finally, the data also indicates that starting at about six years after transition, regimes without NVR begin to catch up in average polyarchy scores. However, it is important to consider that we only analyze regimes that survived until the respective point in time, which means we compare successful cases of democratization induced by NVR to successful cases without NVR. This implies that even the successful cases without NVR induced transitions on average do not attain levels of democratic quality as high as the regimes that were induced by NVR.

While the general trends described above appear to be the same across all three datasets, the Geddes data in some ways slightly deviates from the other two. Most importantly, in the Geddes data, a substantial increase of the polyarchy score occurs one year earlier than in Ulfelder and Boix data (i.e. starting at two years before the transition event instead of one year before the event). This difference may be caused by the more demanding transition coding of the Geddes dataset. Whereas the Boix and the Ulfelder data only require regimes to meet the conditions of a minimum degree of inclusive suffrage and the executive being chosen via free and fair elections for coding the occurrence of democratic transition, the Geddes coding also requires party competition. Due to this difference in coding transition events, the Geddes data codes transitions later than the other two datasets, and therefore substantial improvements in democratic quality occur before the transition event.

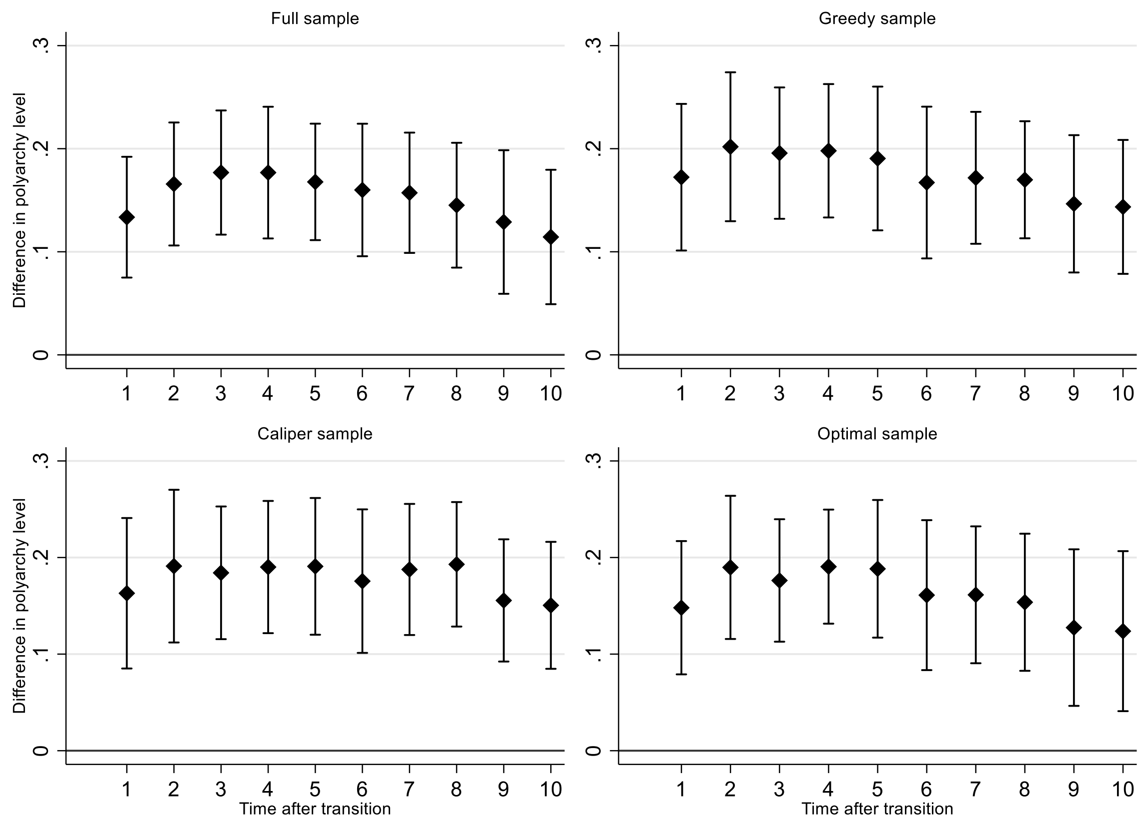
Regression analysis

To account for confounding factors and alternative explanations, we also evaluate the patterns described above with linear regression models. We analyze the effect of NVR on post-transition levels of democratic quality and on pre- to post-transition improvements of democratic quality. Thus, we utilize two outcome variables, one referring to the average level of polyarchy at a specific point in time after the transition and the other to the difference between the level of polyarchy at one year before the transition occurred and a specific point in time after the transition.

We estimate the effect of NVR induced transition on these outcome variables using a linear regression model with a cross-sectional data structure and covariates measured at baseline, i.e. one year before the transition event occurred. Because of the large number of specifications, we only the report the main effects of NVR in the subsequent analysis instead of the full regression tables.[5] Next, we describe the effect of NVR on post-transition polyarchy levels for the full and matched samples using the Ulfelder data. Figure A6.2 reports point estimates of the difference in polyarchy between NVR and non-NVR induced regimes at 1 to 10 years after transition along with 95% confidence intervals. For example, the estimate for the full sample displayed in figure A6.2 indicate that one year after the transition on average NVR induced regimes have a 0.13 higher polyarchy score than regimes without this feature. The confidence intervals suggest that with a probability of 95% the point estimate falls within the interval of 0.07 and 0.19.

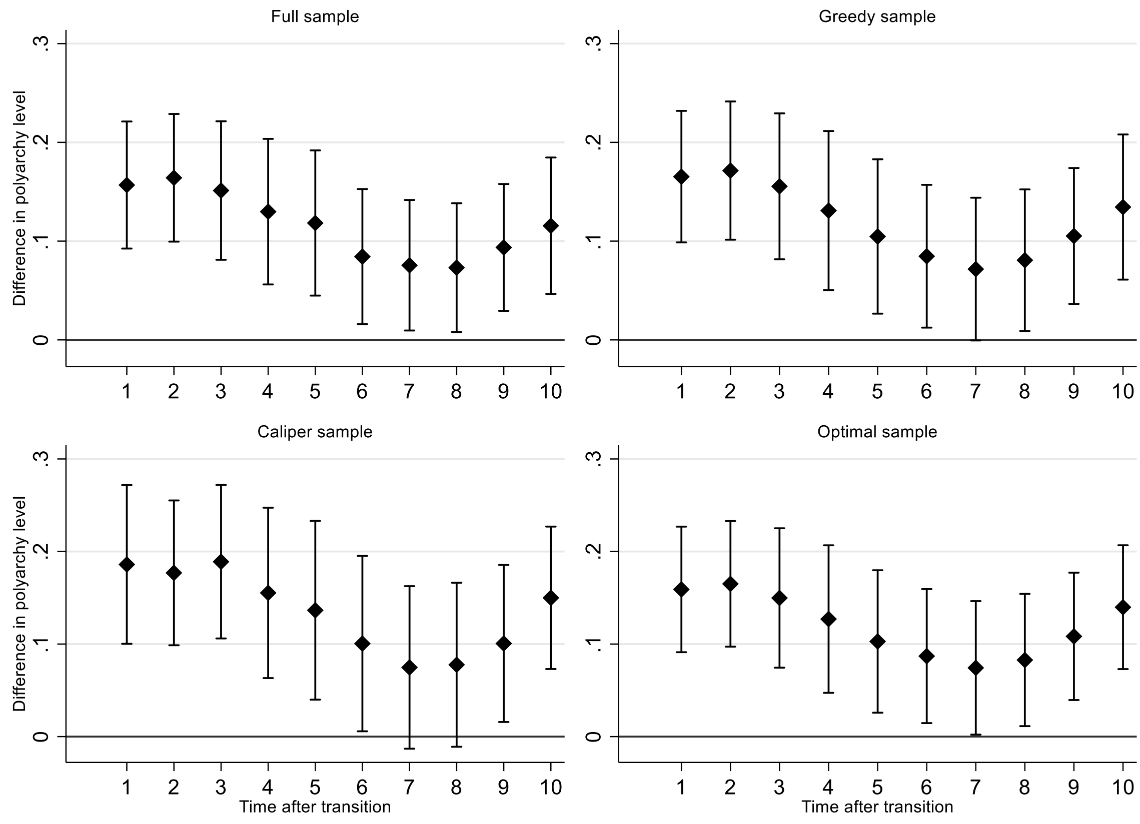
[5] The full results are reported in the accompanying log file included among the replication files.

Figure A6.2. Effect of NVR on post-transition polyarchy levels (Ulfelder data)



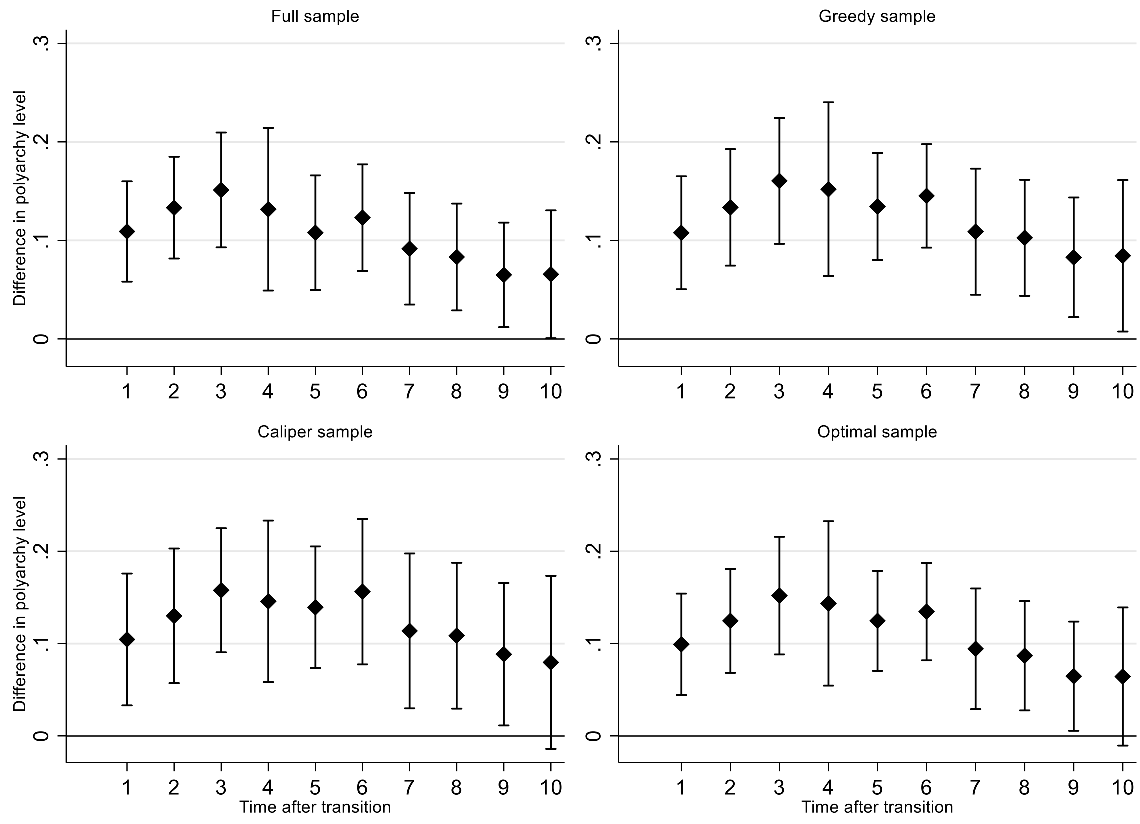
Considering all estimates reported in figure 6.2., we find a substantial and significant effect of NVR on democratic quality. The estimates are significant across all time periods in the full and matched samples. Depending on timing of measurement and sample used in the analysis, the point estimates range between 0.11 and 0.2, which we consider substantial given that the polyarchy index range from 0 to 1. With regard to the temporal dynamics of the effect of NVR, the results suggest a curve-linear trend, i.e. effect size increases during the early years after transition and then diminishes. This result is in line with the previous finding that after some time regimes that where not induced by NVR begin to catch up in terms of democratic quality. Now figure A6.3 reports the results for the Geddes data.

Figure A6.3. Effect of NVR on post-transition polyarchy levels (Geddes data)



As shown in figure A6.3, there appears to be a significant and substantial effect of NVR in most specification. However, for a one time point in the greedy sample and two time points in the caliper sample, the estimated effect is not significant, as indicated by the confidence intervals crossing zero. Additionally, the results from the Geddes sample also do not indicate a diminishing effect of NVR over time. Although after about four years the effect size becomes smaller, it increases again after about 8 years after transition. Finally, we report the results for the Boix data in figure A6.4.

Figure A6.4. Effect of NVR on post-transition polyarchy levels (Boix data)



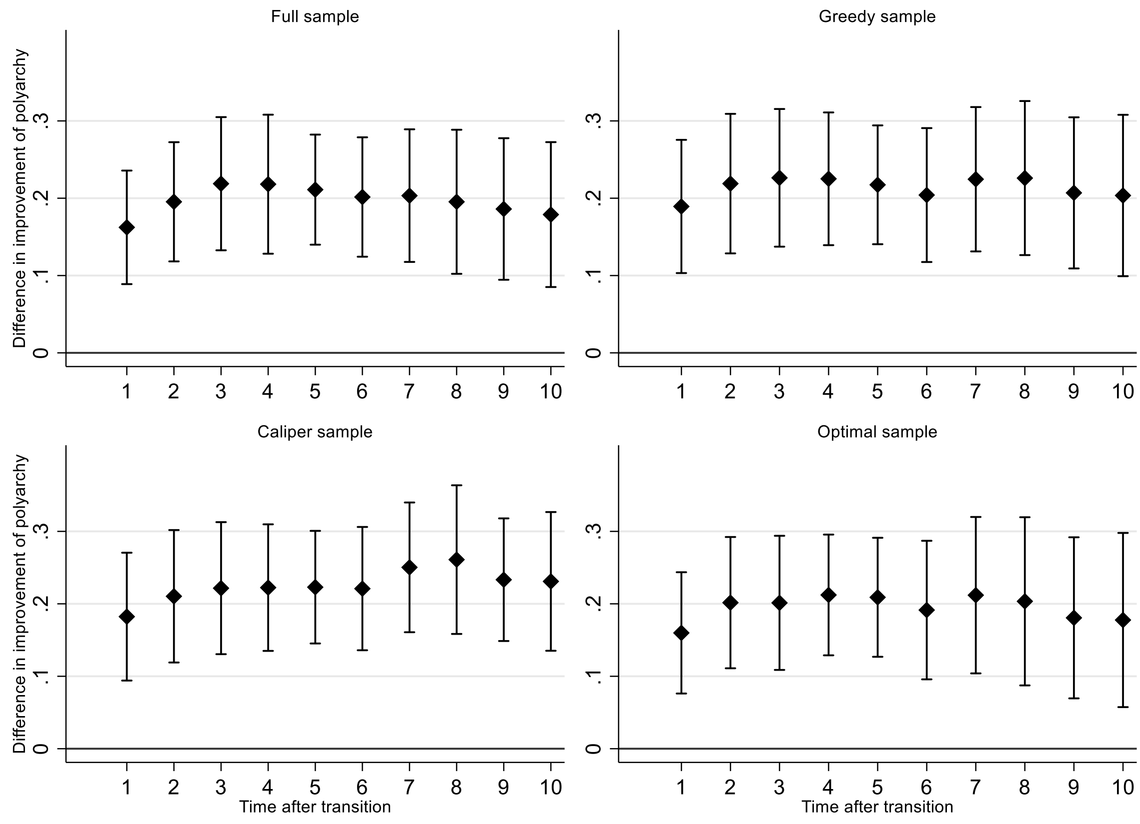
As shown in figure A6.4, the results from the Boix data largely mimic those from the Ulfelder data, although the effect size is generally smaller but still substantial. The estimated effect of NVR is significant in almost all specifications. Only in some samples at 10 years after transition, the confidence intervals cross zero, which indicates that the point estimates are not statistically significant. The results for the Boix data also suggest that the effect of NVR on levels of polyarchy after transition diminishes after about six years.

In sum, the results from the regression analysis suggest that there is robust evidence for a positive effect of NVR on post-transition levels of democratic quality. For the first five years after transition, the difference in democratic quality between NVR induced regimes and regimes without this feature is significant and substantial across all specifications. Moreover, for most specifications the findings also suggest a substantial long-term effect of NVR on democratic quality until up to 10 years after transition. However, the results are inconclusive about the temporal development of the effect of NVR. While the Ulfelder and the Boix data suggest a diminishing effect, the Geddes data does not reveal this pattern.

Next, we evaluate the results form the difference-in-differences analysis (DiD), i.e. how NVR affects the degree of improvement in democratic quality from pre- to post transition. In the DiD set up we observe the outcome for two groups at two points in time. Our groups are regimes where democratization was induced by an NVR campaign and regimes without this characteristic. For these two groups, we observe democratic quality before and after transition. We consider the transition process itself as an intervention at which regimes in the treatment group experience NVR and regimes in the control group do not. To obtain the DiD effect we simply subtract the average change from pre- to post transition democratic quality in the non-NVR group from the average change in the NVR group. DiD estimation accounts for unobservable but time-invariant differences between regimes that experienced NVR induced transitions and regimes without that feature.[6] For instance, one potential unmeasured factor, which we can account for with this procedure is democratic political culture or more specifically the differences in the predisposition of elites and the population towards democratic values. Countries with positive attitudes towards democracy among elites and the population should be more likely to experience the occurrences of NVR induced transitions and at the same time more likely to improve in democratic quality after transition. By looking at difference in improvements instead of difference in levels of democratic quality after transition, the DiD approach accounts for differences in democratic predisposition across countries. Figure A6.5 reports the results for the Ulfelder data.

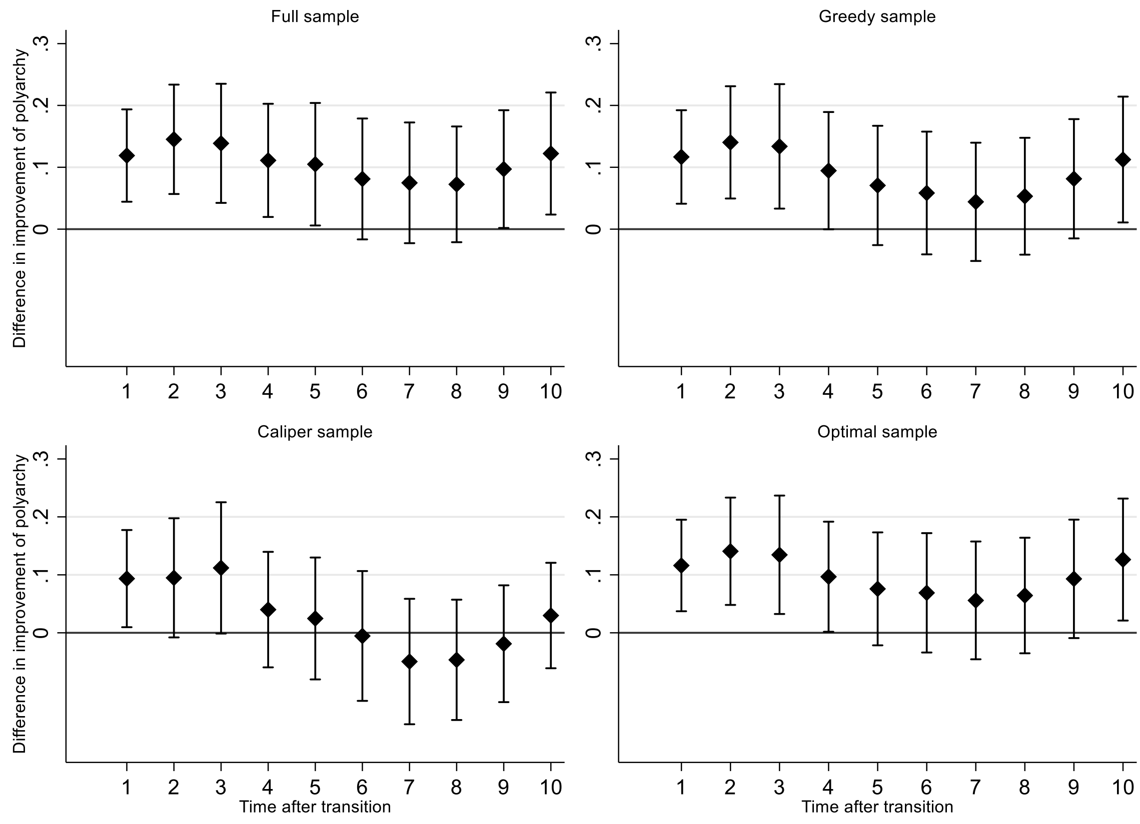
[6] DiD assumes common time trends for treatment and control group. The validity of this assumption is confirmed by the pre-treatment data on democratic quality reported in figure A6.1.

Figure A6.5. Effect of NVR on improvement of polyarchy (Ulfelder data)



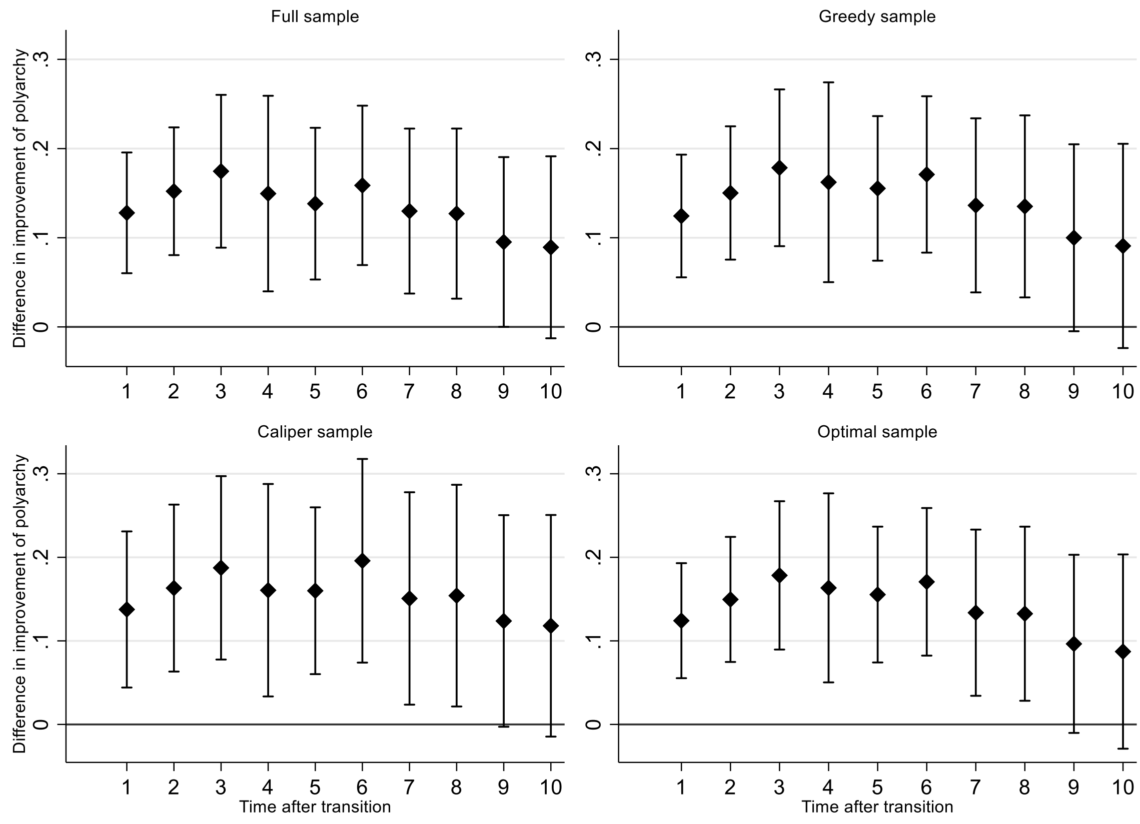
As discussed above, we analyze changes in the respective outcome variable from the year before the transition until up to ten years after the transition. Depending on the timing of measurement, the average improvement in democratic quality is estimated between 0.16 to 0.26 units higher for regimes with transitions induced by NVR compared to regimes without this characteristic. Again, we consider this is a substantial effect, given that the scale for change in the polyarchy score ranges from -1 to 1. Figure A6.6 reports the results for the Geddes data.

Figure A6.6. Effect of NVR on improvement of polyarchy (Geddes data)



For the Geddes data, the effect of NVR on improvements in democratic quality is only significant and substantial for the time period between one and three years after transition. When considering the difference between democratic quality from pre- to post transition at later time points, the effect of NVR is not significant. These findings are largely a result of the coding rules that the Geddes data employs to specify the timing of democratic transitions. As described above, due to a more demanding coding of transition events the Geddes data often records these one year later than in the other two datasets. Given that we measure improvement of democratic quality as the difference in the level of democratic quality from one year before the transition and the respective point in time after the transition, the Geddes data simply misses substantial changes of democratic quality that already occurred earlier. This pattern is also described in figure A6.1. Now figure A6.7 reports the results for the Boix data.

Figure A6.7. Effect of NVR on improvement of polyarchy (Boix data)



The effect of NVR over time follows a similar pattern as in the Ulfelder data, although generally the effect size is a bit smaller, but still substantial. For the Boix data, the effect of NVR is significant and substantial across all samples for the time period from one until 8 years after the transition. However, in all samples, the estimates for the effect of NVR on improvements in democratic quality are not significant at nine and ten years after transition.

In sum, the results from the DiD suggest that NVR induced regimes improve more substantially in democratic quality than regimes without this feature. This also corroborates the causal interpretation of the effect of NVR on democratic quality. The DiD effect of NVR robust and substantial for the early time periods after transition but our results are inconclusive for the later time periods. This again indicates that those regimes which democratized without NVR but managed to survive catch up in terms of democratic quality.

A7: Analysis of mechanisms

The following section provides additional results on the mechanisms described in chapter 4 of the book. Again, we mainly use linear regression to estimate treatment effects of NVR on different outcome variables.

Legislative constrains of the executive

In figure A7.1, A7.2, and A7.3 we describe the effect of NVR on post-transition levels of legislative constraints of the executive for the full and matched samples using the three different regime datasets. The figures report point estimates for the difference in legislative constrains between NVR and non-NVR induced regimes at 1 to 10 years after transition along with 95% confidence intervals.

The results generally corroborate the argument that NVR advances a leveling of the political playing-field by means of constraining the executive through parliament. The point estimates are positive and substantial for all time periods and across different datasets and matching procedures. Moreover, for the Ulfelder and the Geddes data (figure A7.1 and A7.2), the effects are also very robust, as the estimates are statistically significant for all but one specification. However, the results for the Boix data, described in figure A7.3, indicate that the effect of NVR on legislative constrains is not entirely robust. Although point estimates are always positive and substantial, some effects are not statistically significant as shown by the low bound of the confidence interval crossing zero.

Figure A7.1. Effect of NVR on post-transition legislative constraints (Ulfelder data)

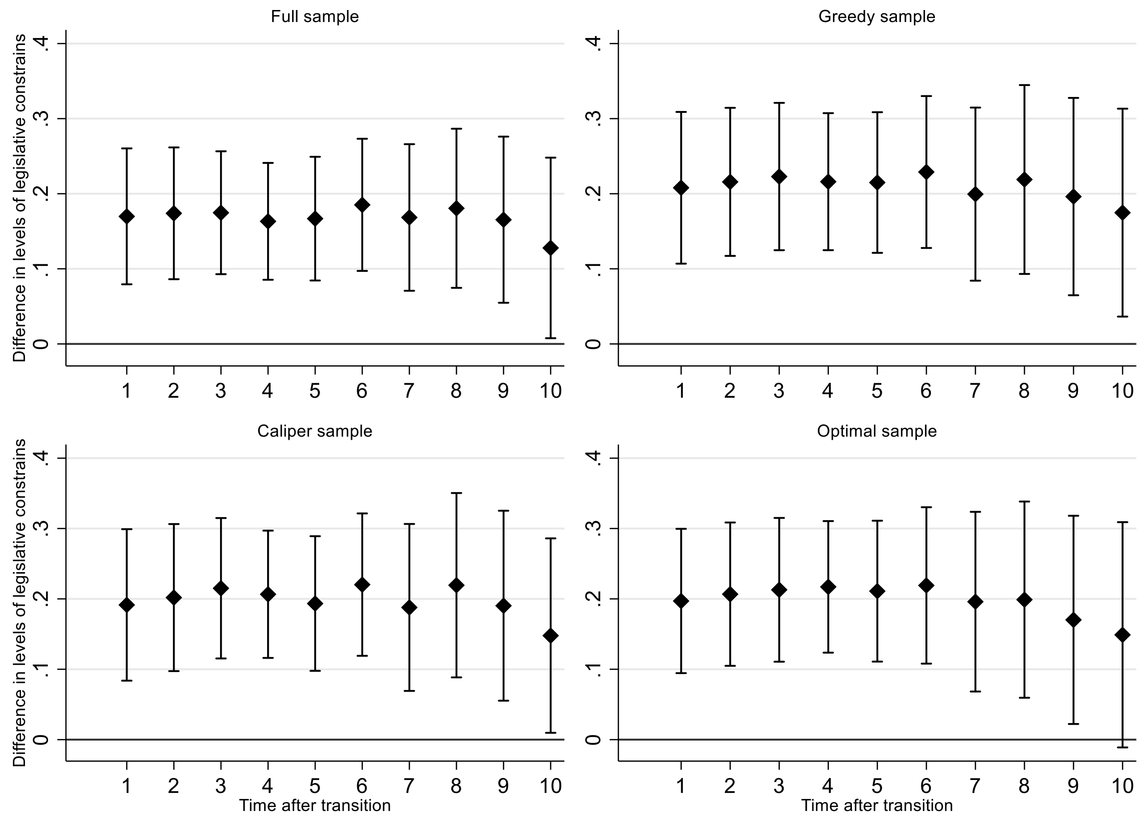


Figure A7.2. Effect of NVR on post-transition legislative constraints (Geddes data)

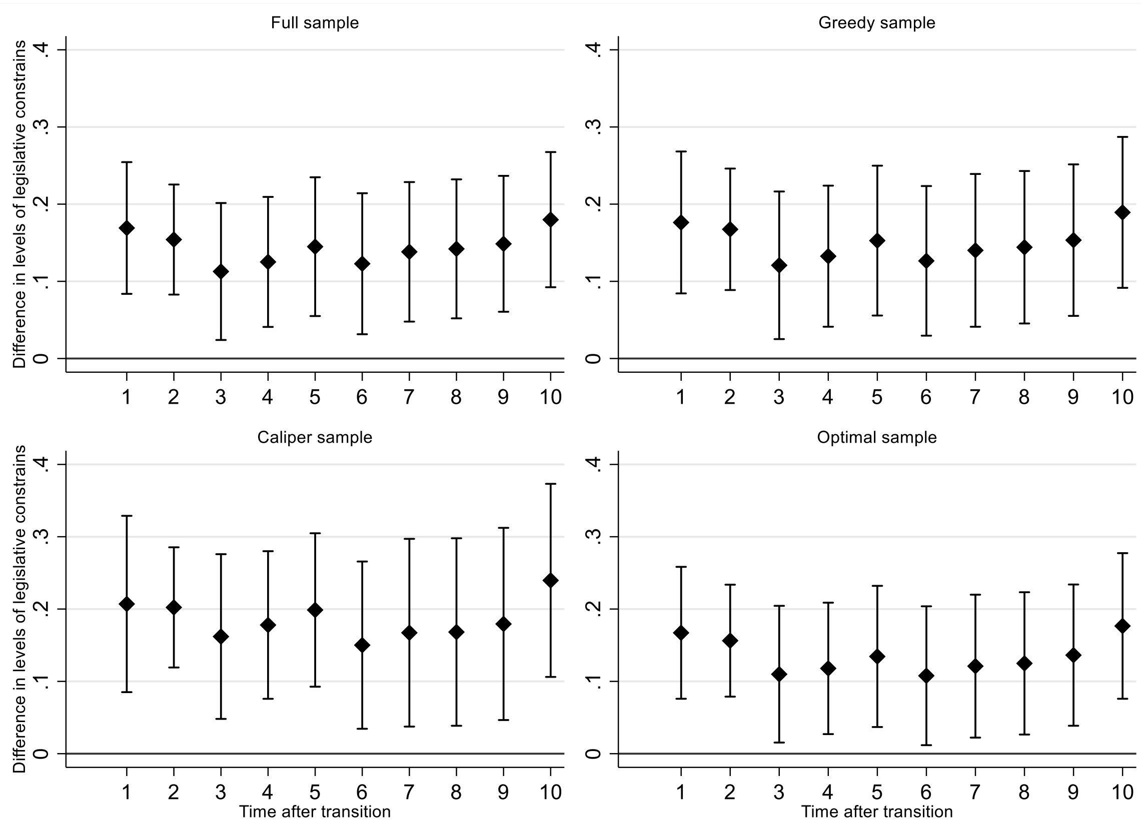
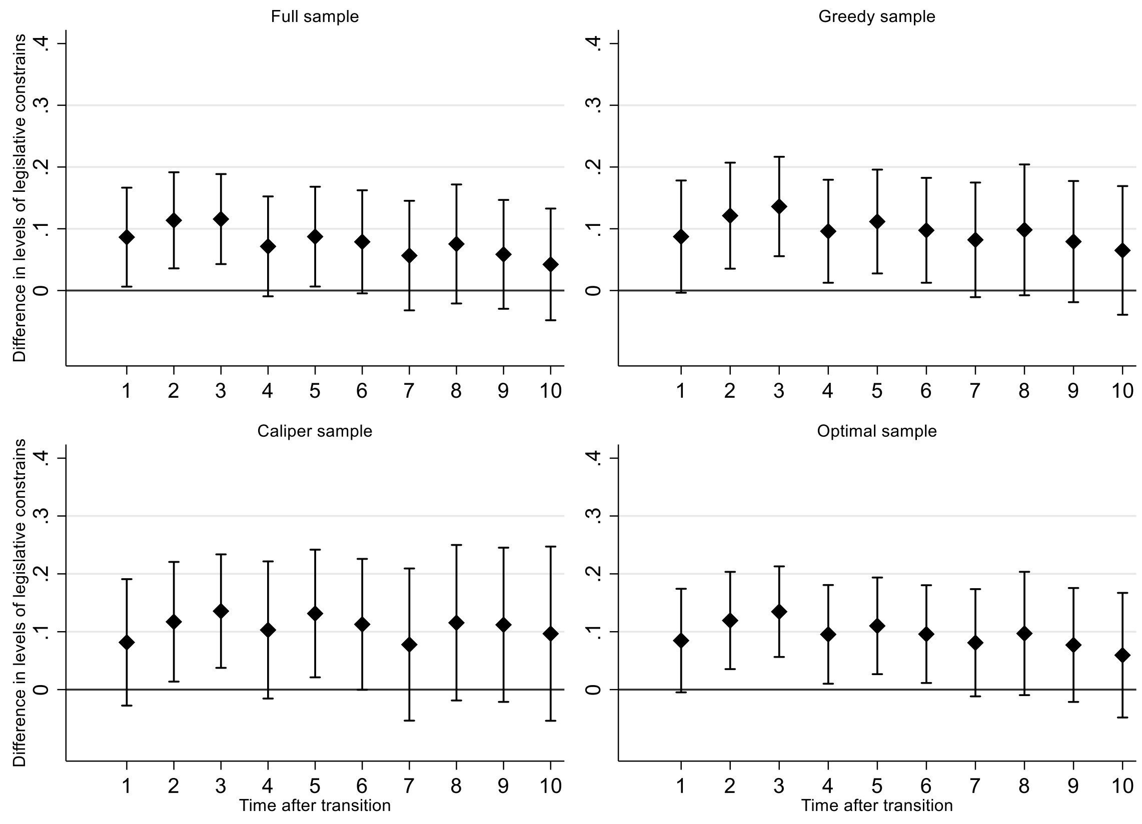


Figure A7.3. Effect of NVR on post-transition legislative constraints (Boix data)



Freedom of expression

Next, we describe the results for the effect of NVR on freedom of expression in figures A7.4, A7.5, and A7.6. The results show that there is a robust and substantial positive effect of NVR on post transition levels of freedom of expression. Up until ten years after transition NVR-induced regimes score higher on this aspect than regimes that democratized without NVR. The point estimates are positive and substantial and the estimated effect is statistically significant most specifications. However, figure A7.6 indicates that when using the Boix data, the positive effect of NVR maybe diminishing over time, because for later time periods after transition the effect gets substantially smaller and in some cases is not significant anymore. However, considering all specifications from all three datasets, there appears to be robust evidence for a positive effect of NVR on freedom of expression.

Figure A7.4. Effect of NVR on post-transition freedom of expression (Ulfelder data)

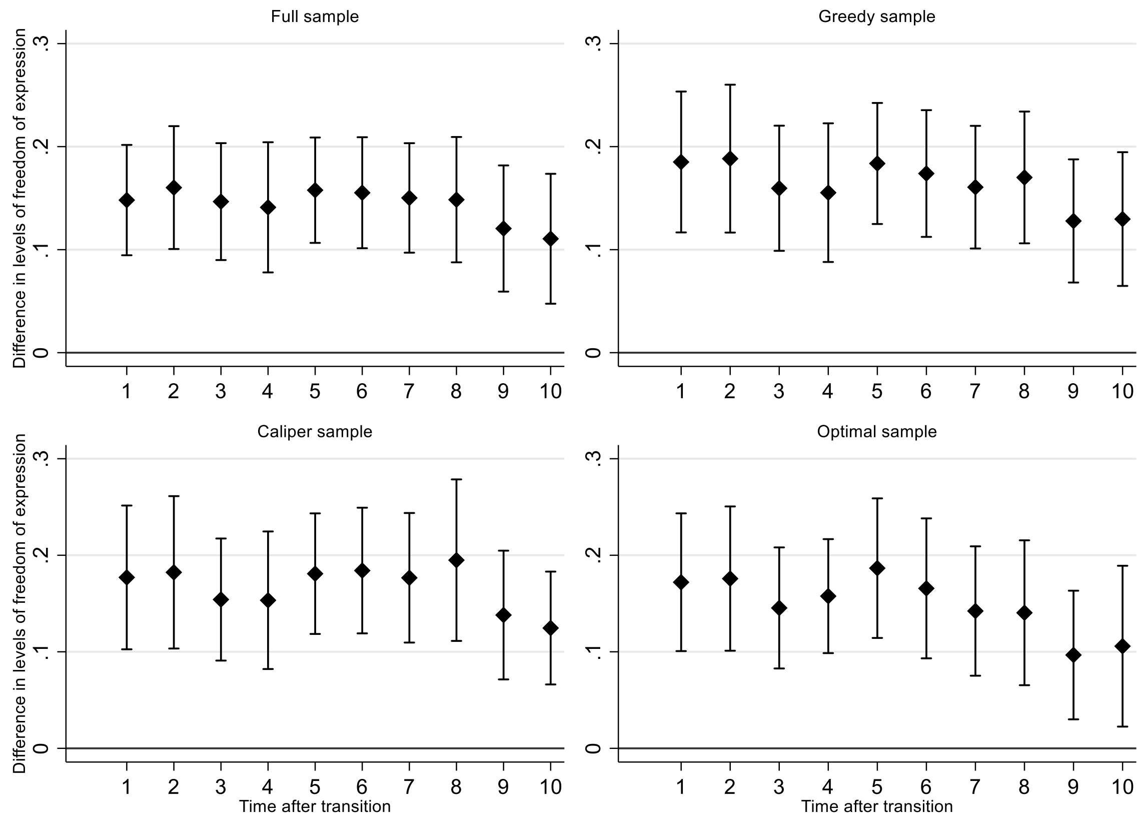


Figure A7.5. Effect of NVR on post-transition freedom of expression (Geddes data)

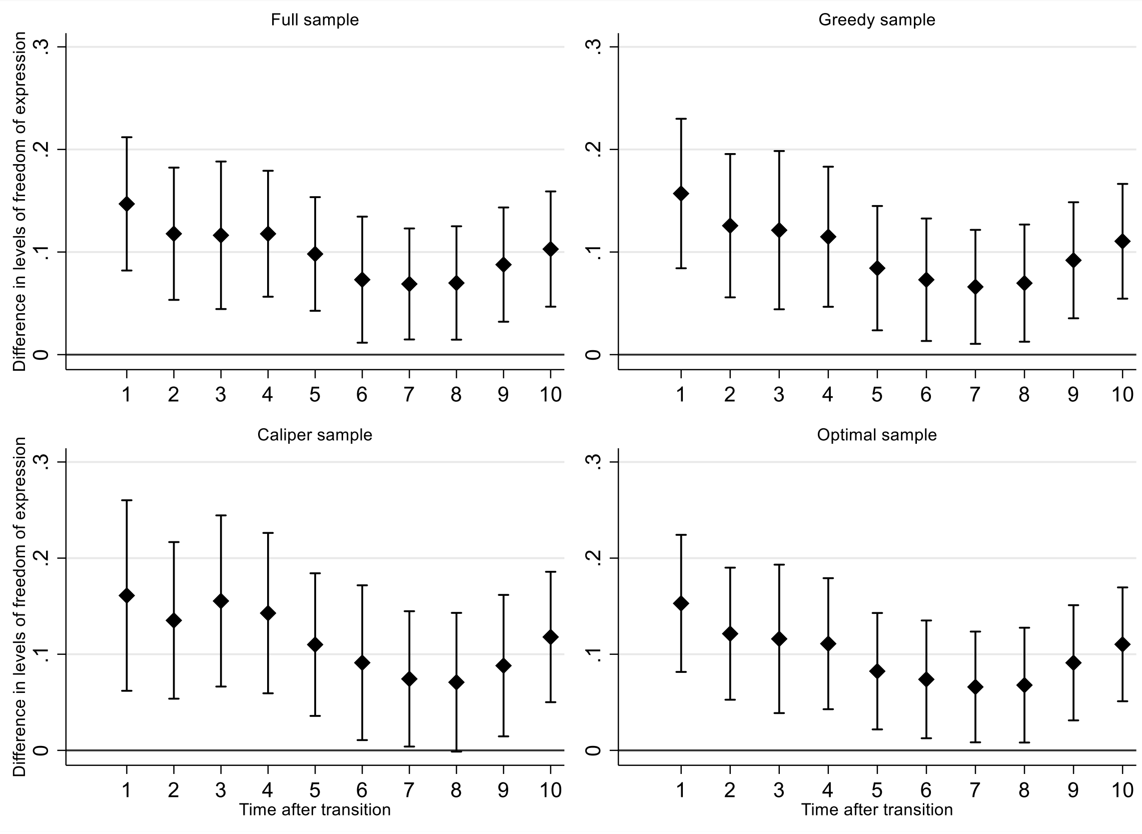
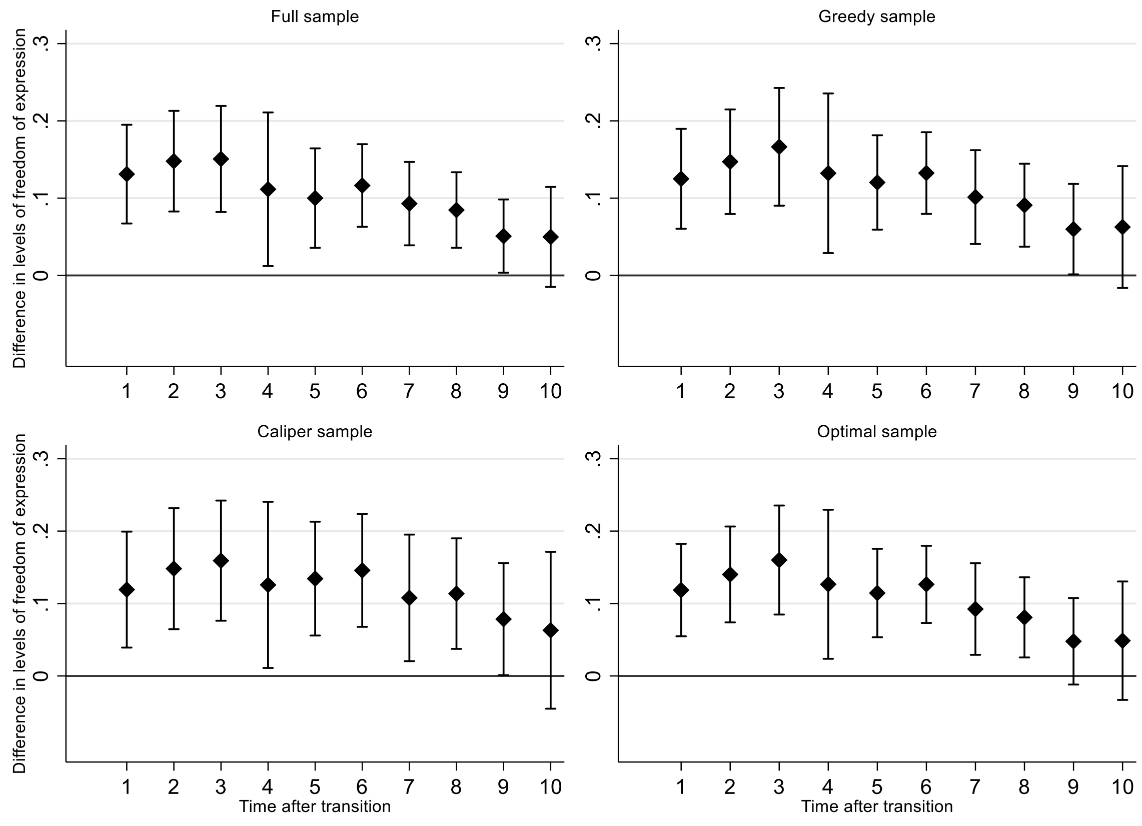


Figure A7.6. Effect of NVR on post-transition freedom of expression (Boix data)



Freedom of association

Similarly, we describe the results for freedom of association in figures A7.7, A7.8, and A7.9. As shown there, the effect reported in the main text in chapter 4 appears to be quite robust in various specifications. In all specifications there is evidence of a positive and substantial effect of NVR on post-transition levels of freedom of association. However, as shown in figures A7.8 and A7.9, there is some indication of diminishing effect sizes and especially in later time periods the effect is not statistically significant anymore. In sum, however, there appears to be robust evidence that NVR substantially advances freedom of association after transition.

Figure A7.7. Effect of NVR on post-transition freedom of association (Ulfelder data)

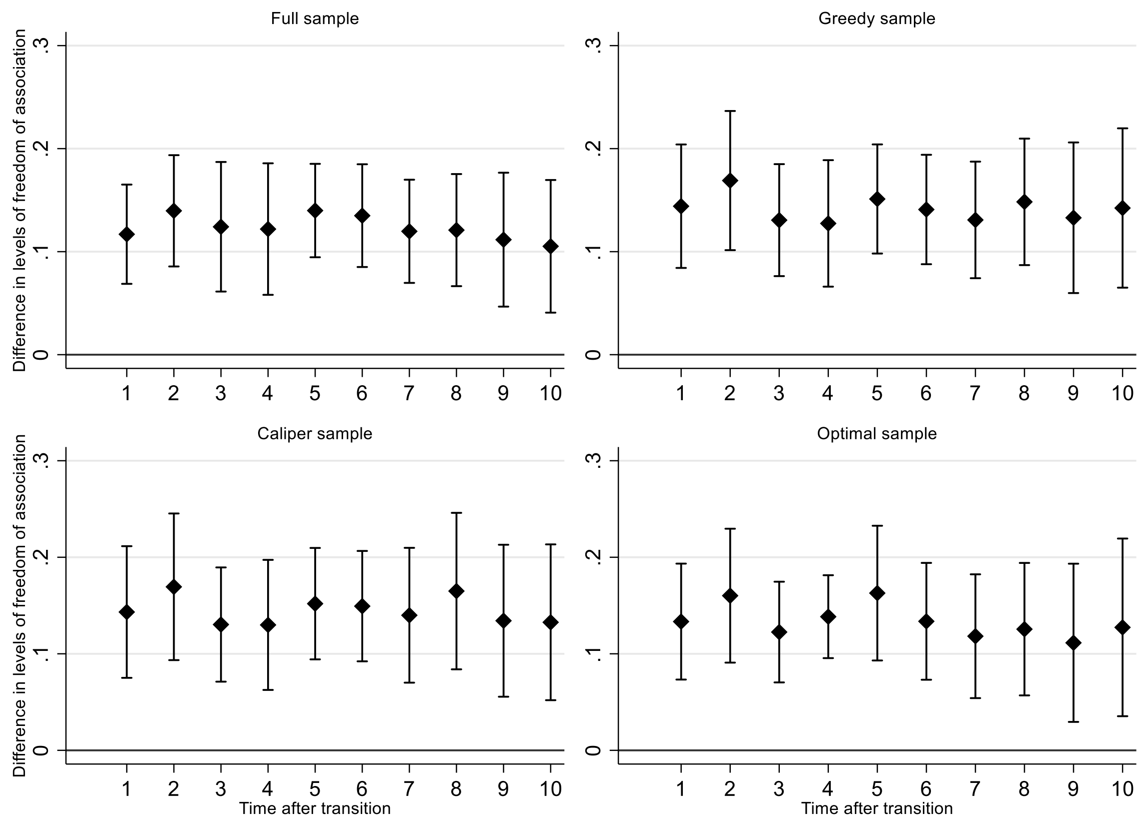


Figure A7.8. Effect of NVR on post-transition freedom of association (Geddes data)

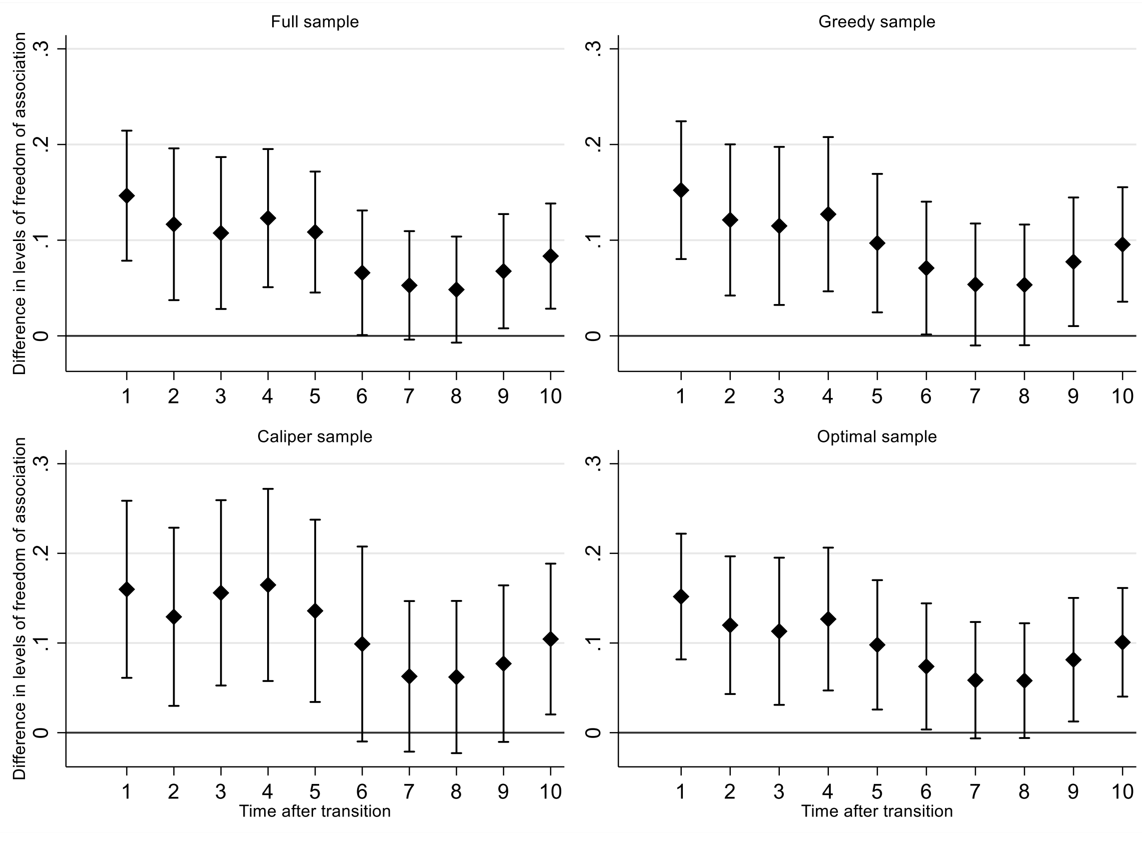
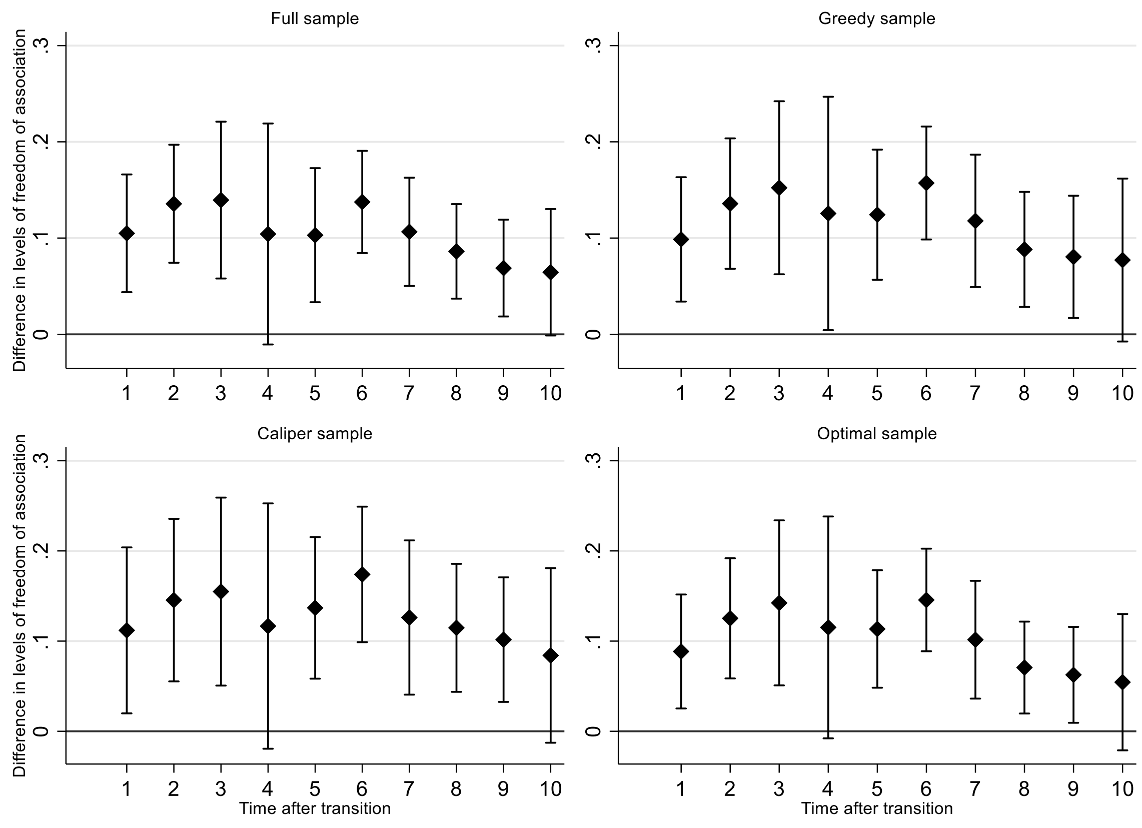


Figure A7.9. Effect of NVR on post-transition freedom of association (Boix data)



CSO independence

Next, in figures A7.10, A7.11 and A7.12, we evaluate the effect of NVR on post-transition independence for CSOs. Figure A7.10 reports the results for the Ulfelder data. As shown there, NVR has a positive and substantial impact on the independence of CSO in the post-transition political environment. This effect is robust across all different specifications. We obtain similar results when using the Geddes and Boix data, which are shown in figures A7.11 and A7.12 respectively. Only when using the Geddes data and the caliper matching procedure three specifications do not provide an effect that is statistically significant. In sum, however, the results indicate robust support for the positive effect of NVR on CSO independence.

Figure A7.10. Effect of NVR on post-transition CSO independence (Ulfelder data)

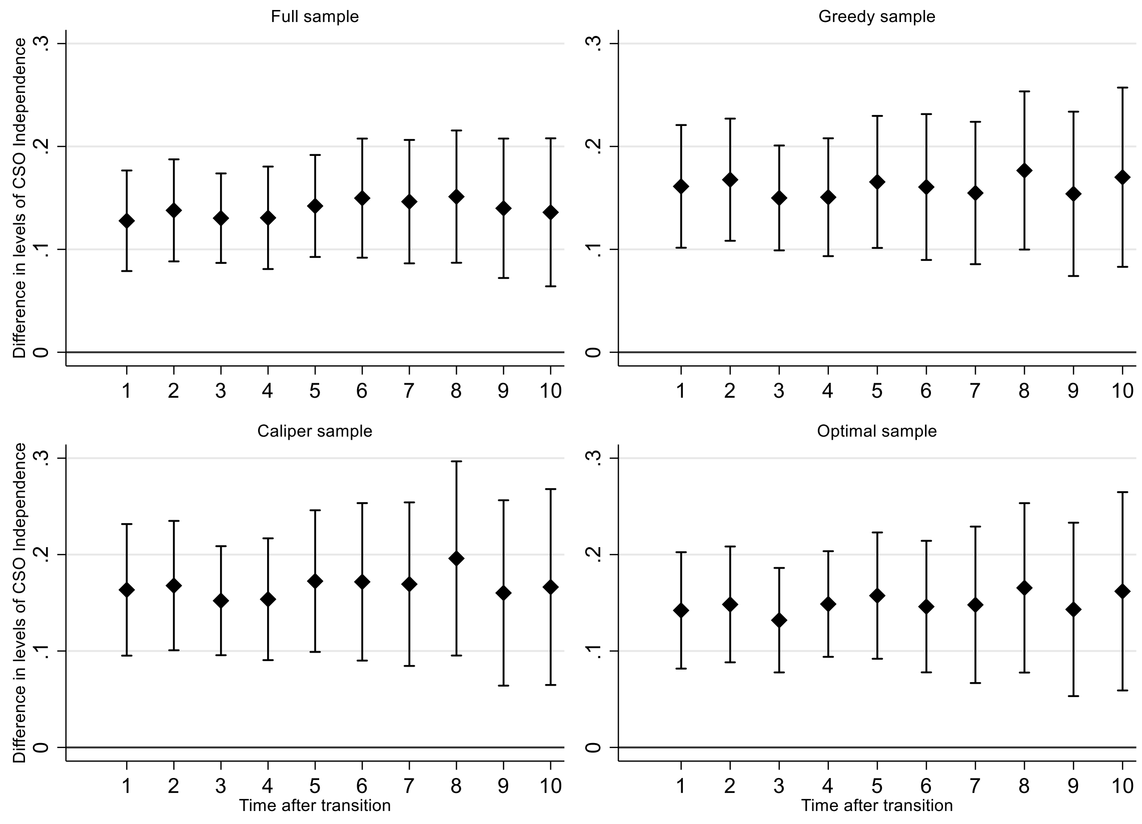


Figure A7.11. Effect of NVR on post-transition CSO independence (Geddes data)

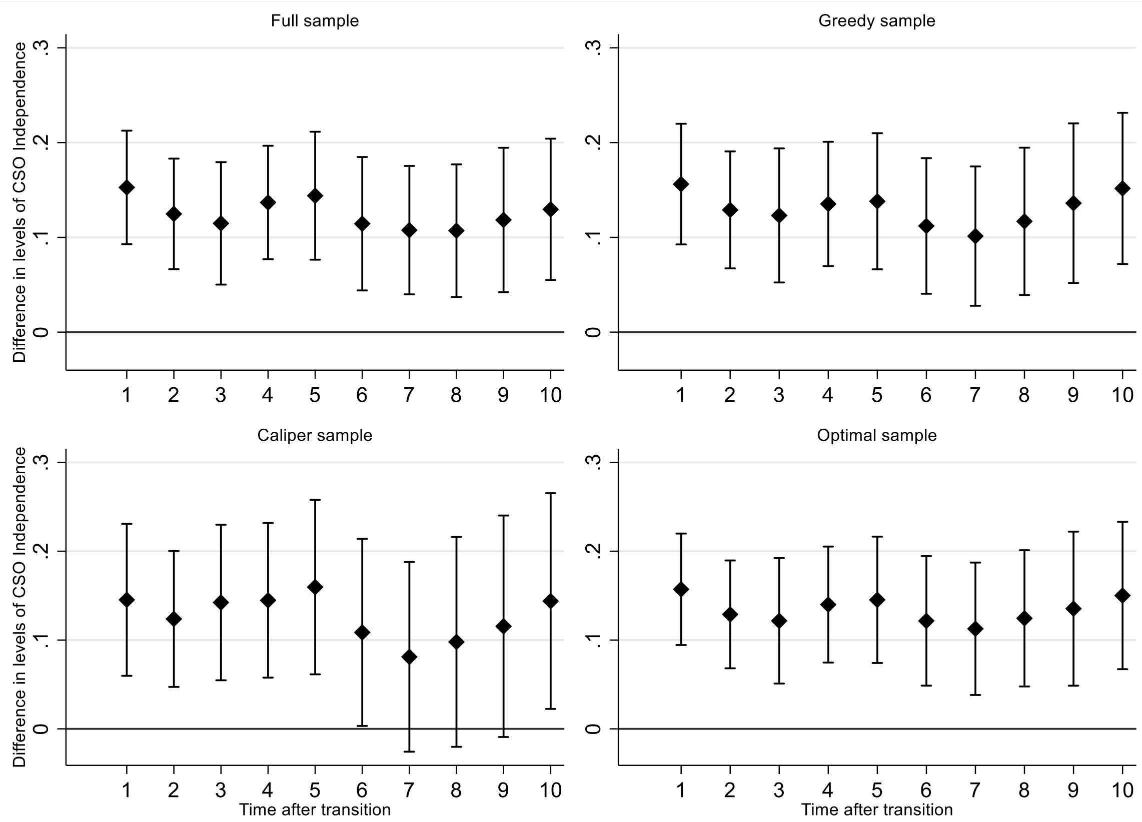
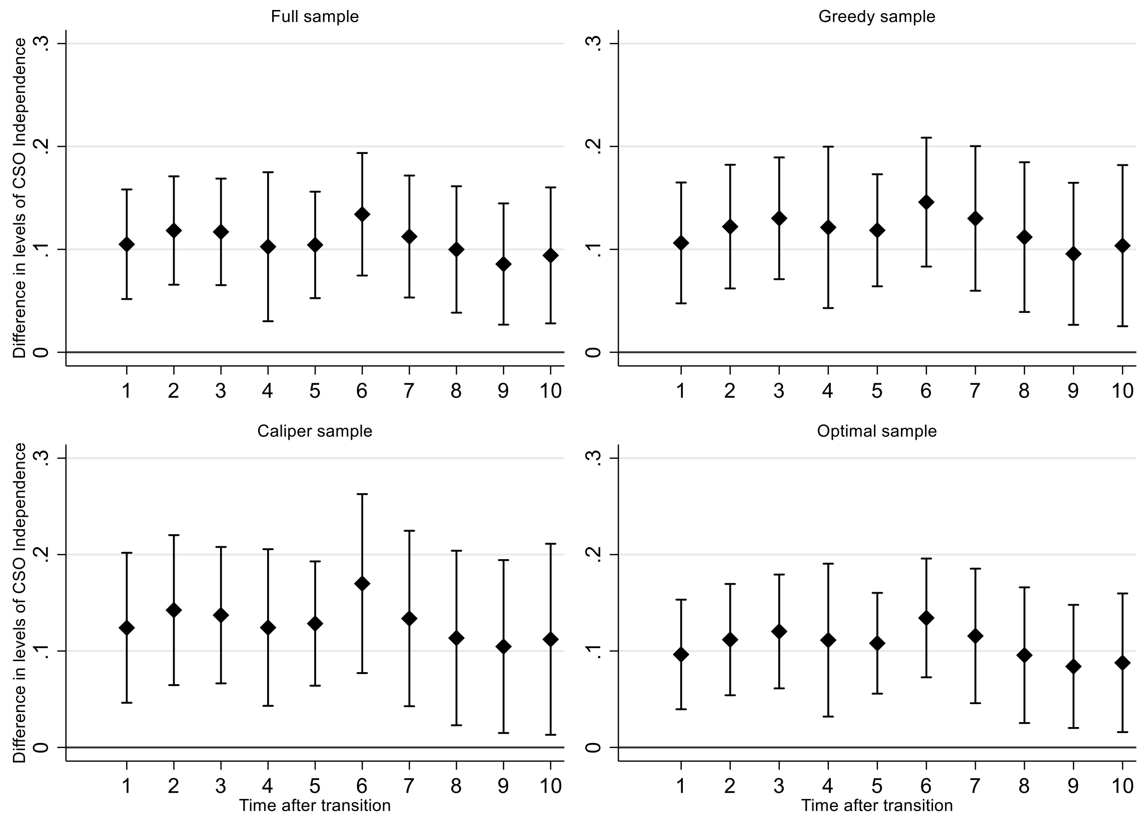


Figure A7.12. Effect of NVR on post-transition CSO independence (Boix data)



CSO freedom of repression

In figures A7.13, A7.14, and A7.15 we describe the results of the analysis of the effect of NVR on post-transition freedom of repression for CSOs. Figure A7.13 provides the results for the Ulfelder data, which indicate a substantial positive impact, which is also quite robust. Only in one specification the effect is not statistically significant. With regard to direction and size, we obtain similar estimates when using alternative datasets. However, as shown in figures A7.14, the effect estimates are not statistically significant in some specifications when using the Geddes data, especially when considering time periods between six and eight years after transition.

Figure A7.13. Effect of NVR on post-transition CSO freedom of repression (Ulfelder)

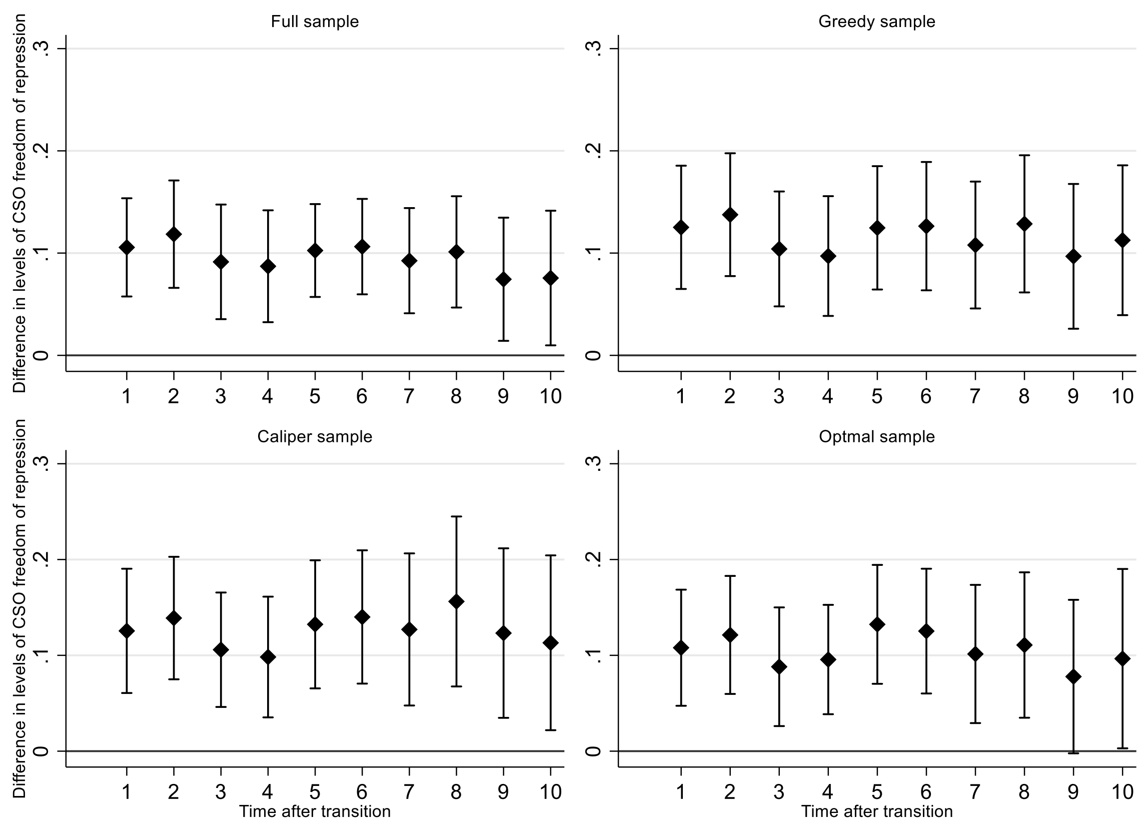


Figure A7.14. Effect of NVR on post-transition CSO freedom of repression (Geddes data)

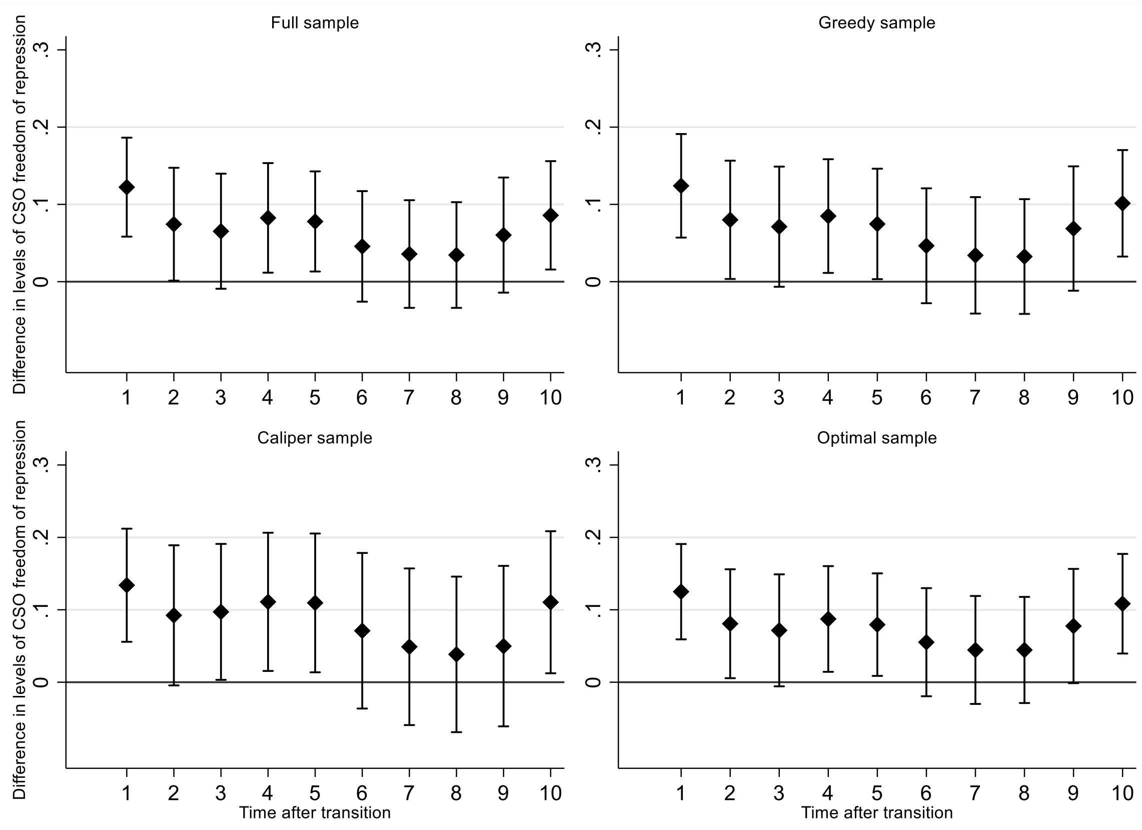
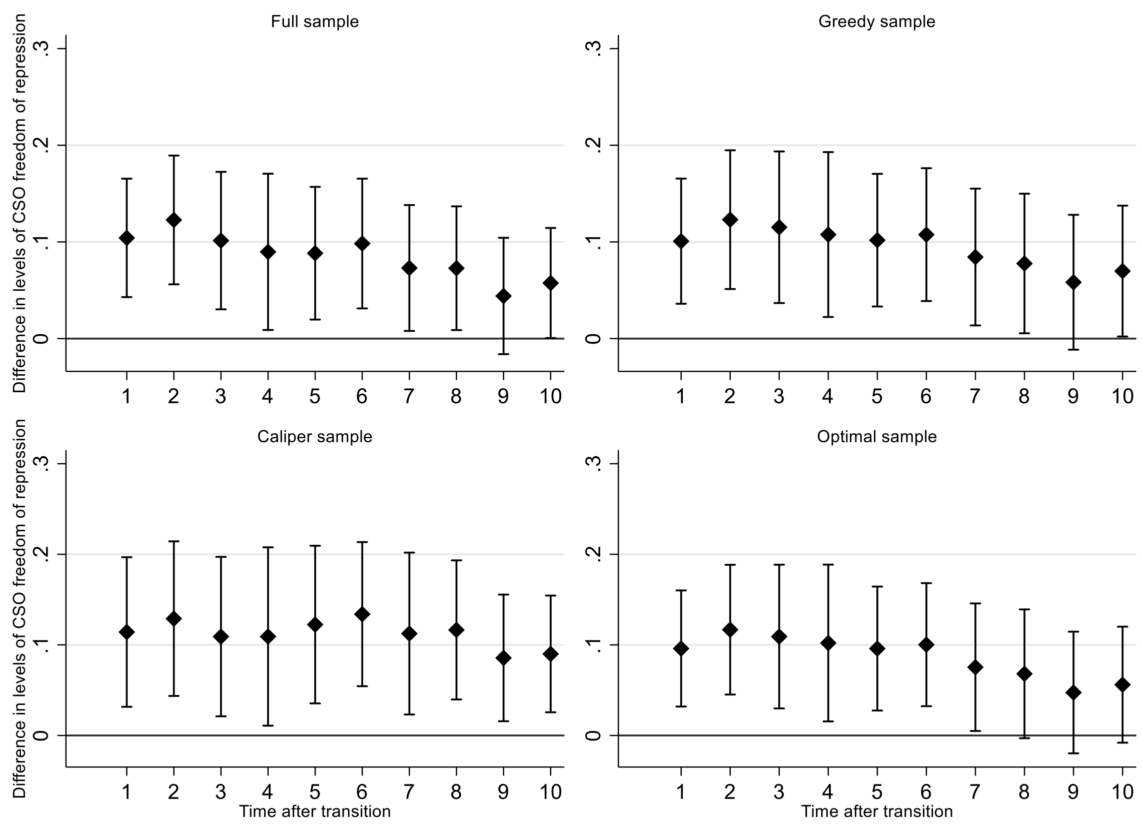


Figure A7.15. Effect of NVR on post-transition CSO freedom of repression (Boix data)



Also, when using the Boix data the effect is not entirely robust as shown in figure A7. 15. Here the estimates for the time periods nine and ten years after transition are not statistically significant in all specifications. In sum, there appears to be a substantial positive effect of NVR on post-transition freedom of repression for CSOS, which is, however, not entirely robust.

CSO consultation

In figures A7.16, A7.17, and A7.18, we describe the results of the regression analysis for the effect of NVR on post-transition consultation of CSOs. As shown there, the estimates indicate a substantial positive effect. This effect is also quite robust to different matching specifications and datasets used in the analysis. In only very few specifications it is not statistically significant. Therefore, we conclude that there appears to be robust evidence for a positive impact of NVR induced transitions on this aspect of the political environment for CSOs.

Figure A7.16. Effect of NVR on post-transition CSO consultation (Ulfelder data)

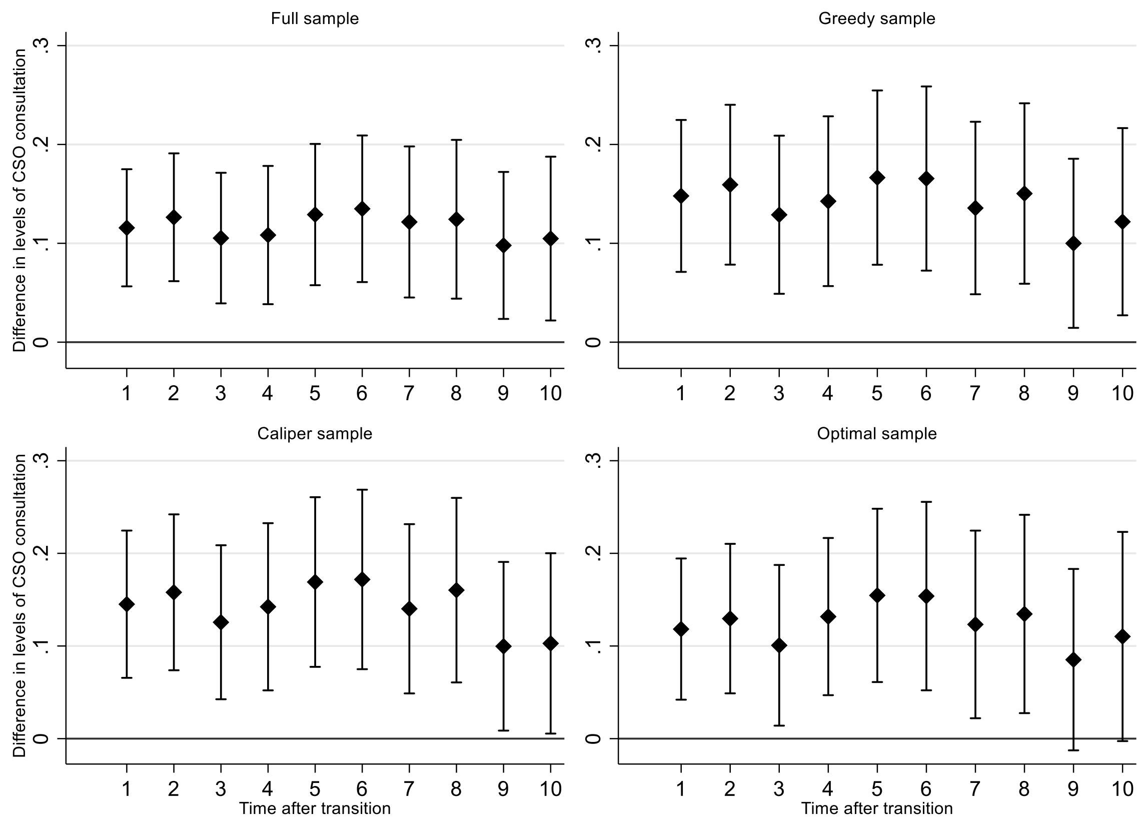


Figure A7.17. Effect of NVR on post-transition CSO consultation (Geddes data)

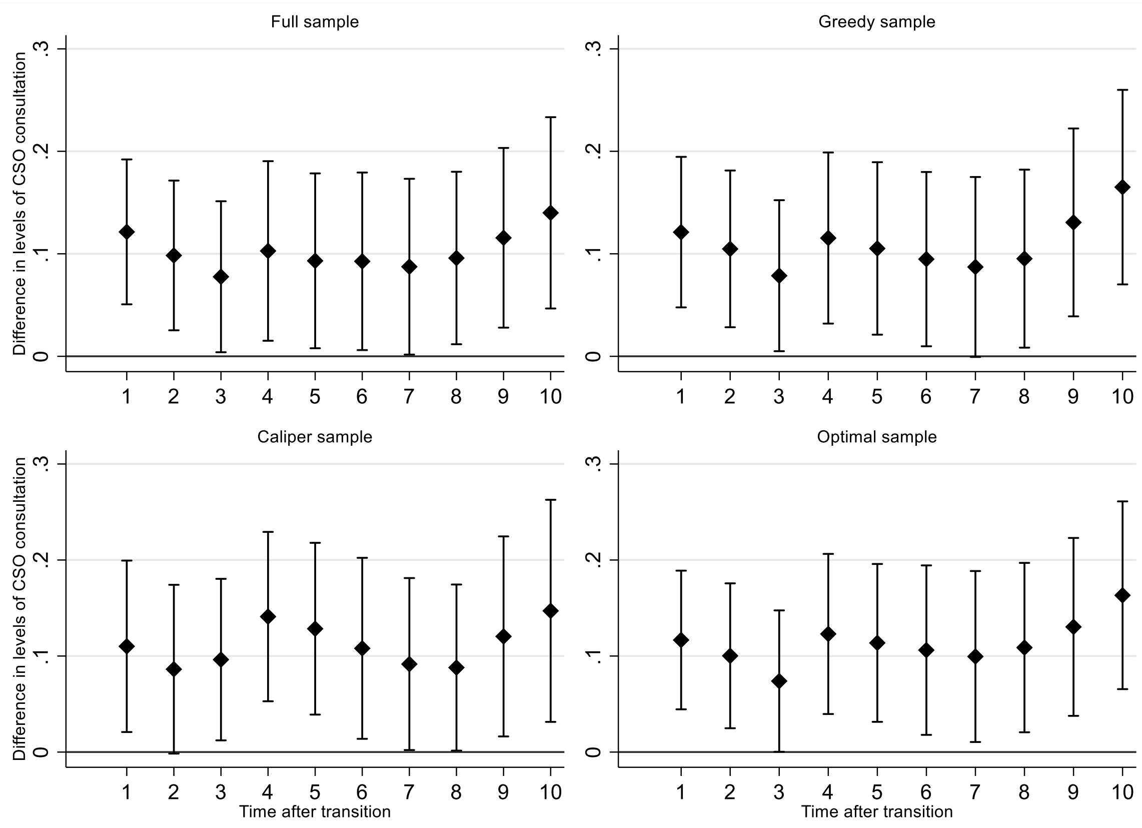
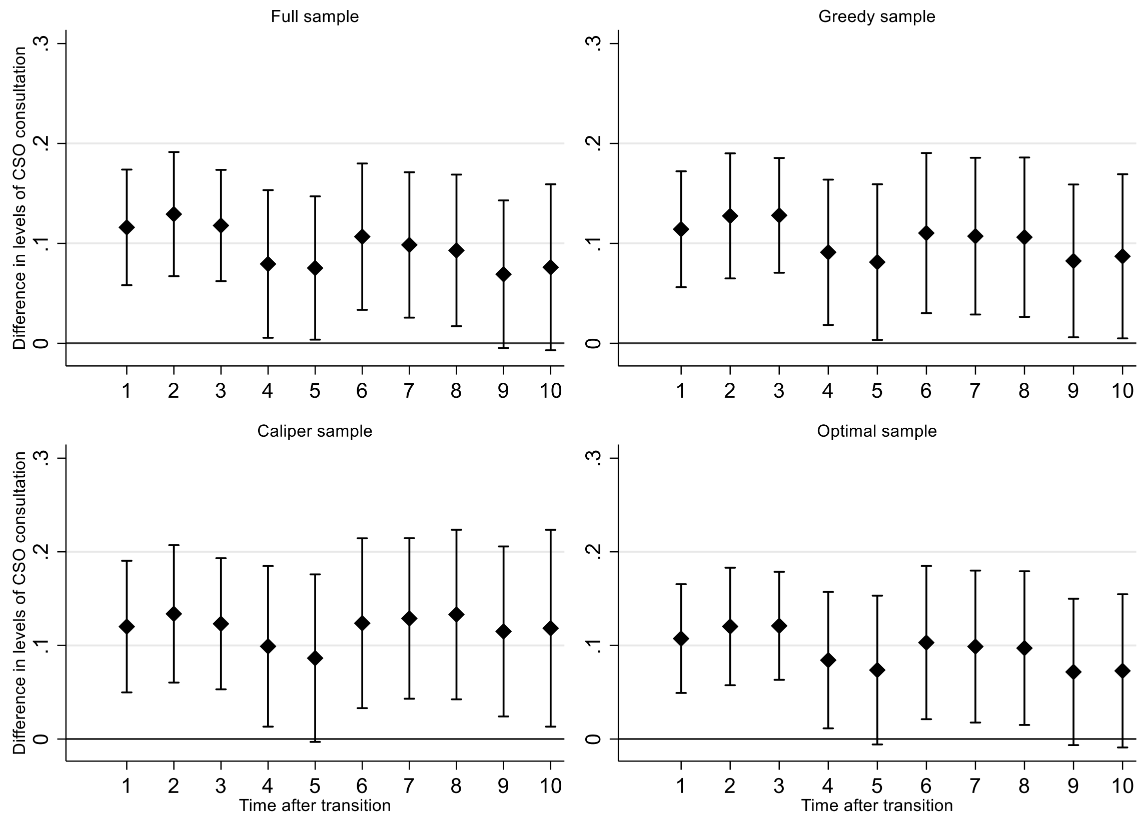


Figure A7.18. Effect of NVR on post-transition CSO consultation (Boix data)



CSO participation

Finally, figures A7.19, A7.20, and A7.21 describe the results of the analysis of the effect of NVR on post-transition participation in CSOs. Here, the estimates are very similar across different datasets and matching specifications. There appears to be a small positive effect, which is however, not statistically significant. Thus, we conclude that our results are inconclusive about the effect of NVR on post-transition CSO participation.

Figure A7.19. Effect of NVR on post-transition CSO participation (Ulfelder data)

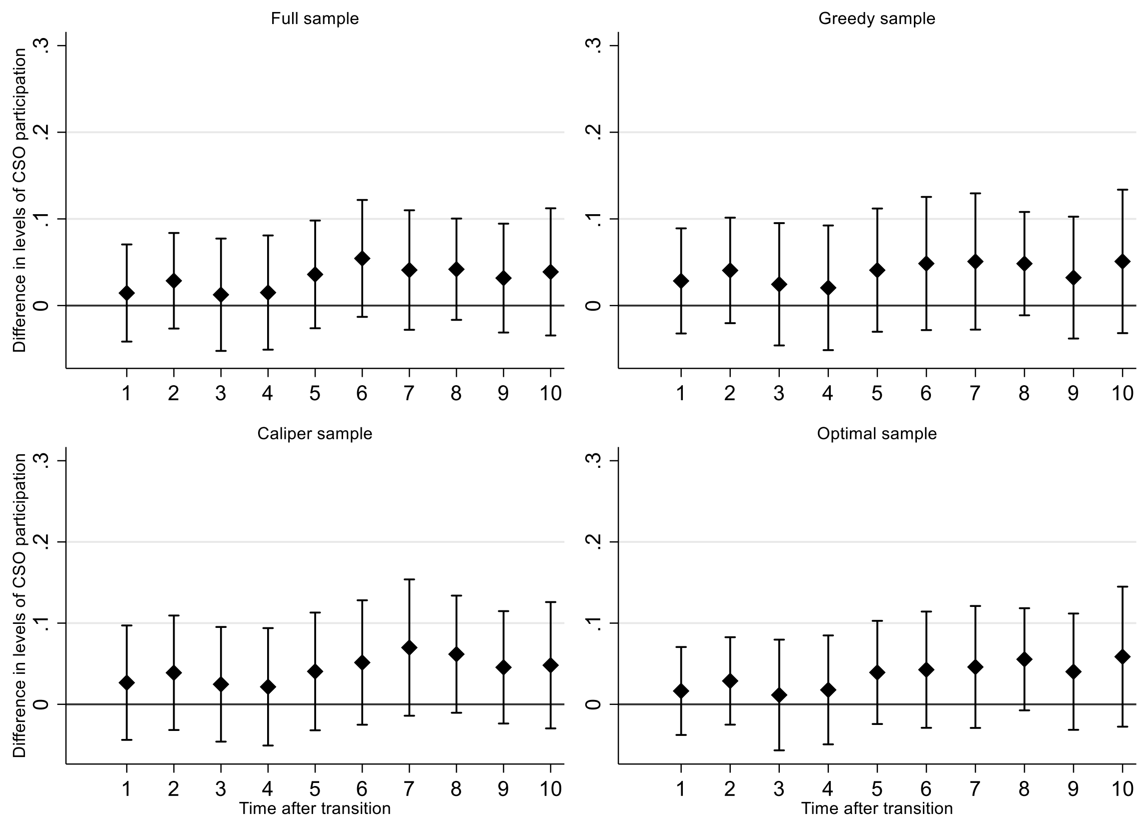


Figure A7.20. Effect of NVR on post-transition CSO participation (Geddes data)

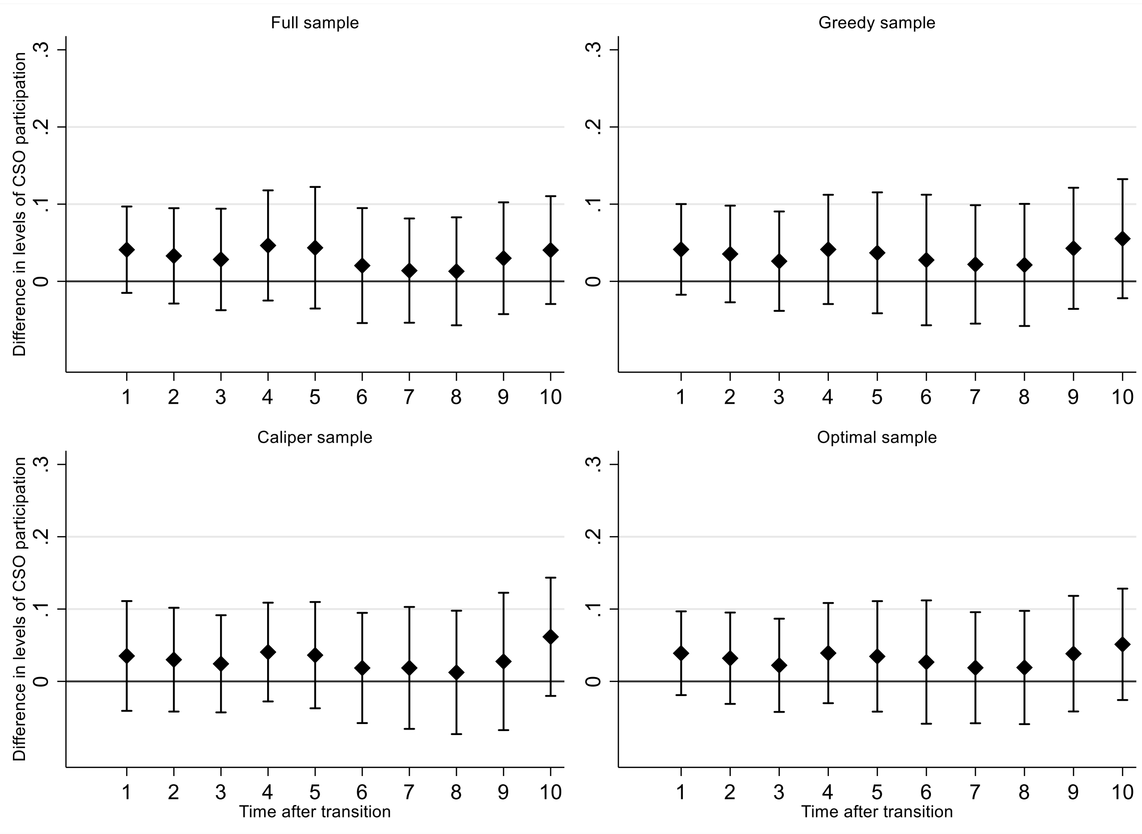
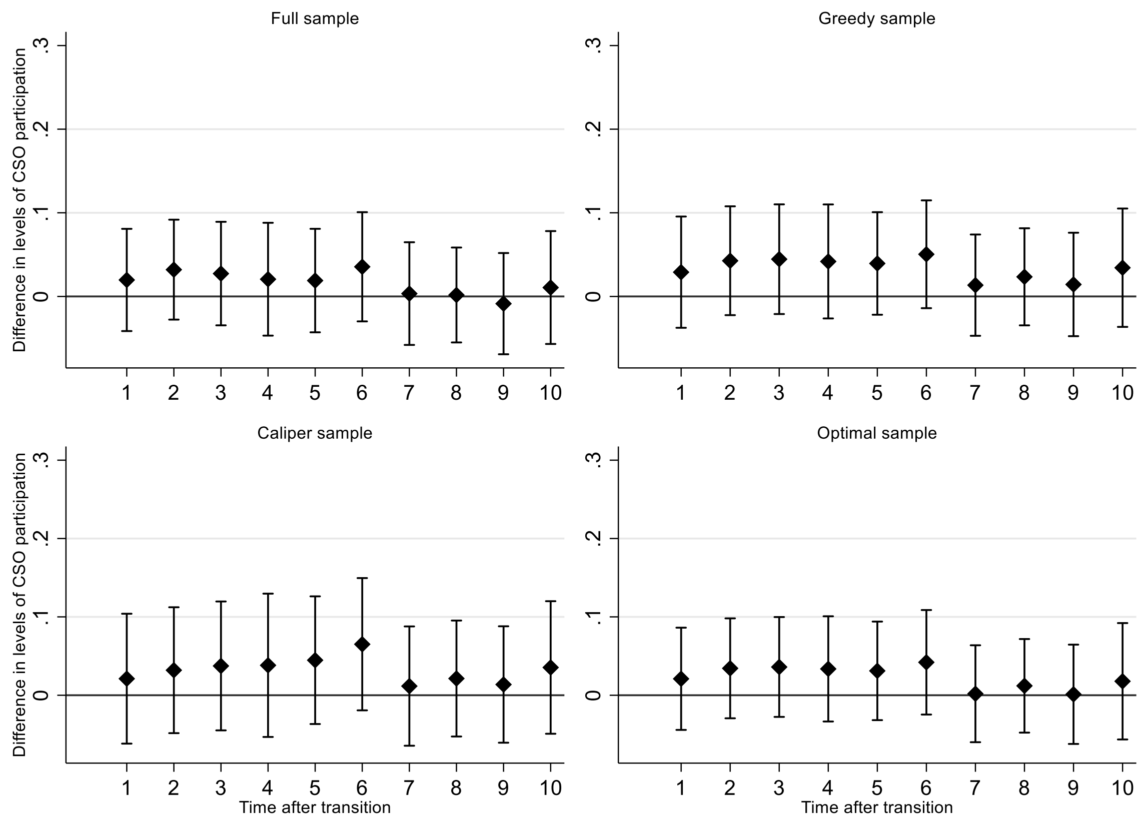


Figure A7.21. Effect of NVR on post-transition CSO participation (Boix data)



Coups

Next, we provide detailed results for the effect of NVR-induced transition on post-transition coups. Because, we only coded the timing and success of post-transition coups for the Ulfelder data, the results are limited to this one dataset. Table A7.1, describes the results of a logistic regression model estimating the effect of NVR on the probability of a coup attempt after transition and its success.

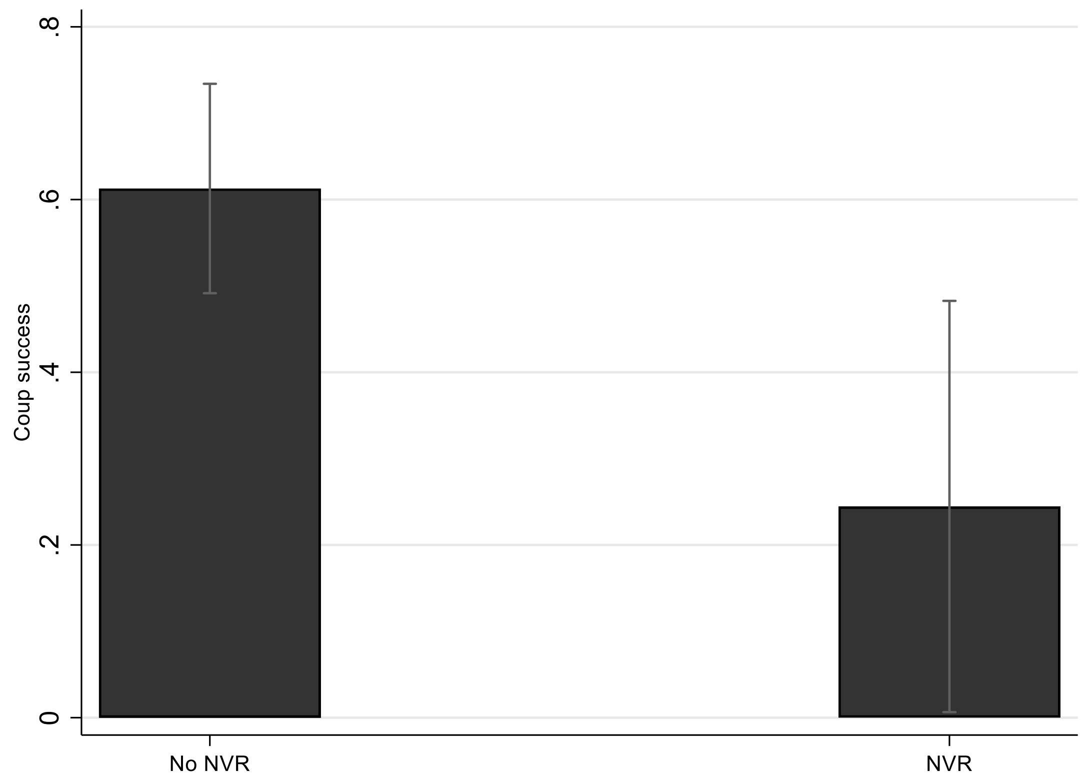
Table A7.1. Logit models with time-varying covariates (Ulfelder data)

	(1) Coup attempt	(1) Coup success
Neighboring Democracies$_{(t-1)}$	0.99**	0.98*
	(0.01)	(0.01)
GDP p.c.$_{(log,\ t-1)}$	0.85	0.47*
	(0.17)	(0.20)
Urbanization$_{(t-1)}$	0.04**	3.42
	(0.05)	(13.64)
Total Population$_{(log,\ t-1)}$	0.84	1.09
	(0.09)	(0.24)
Previous Instability	1.08	0.92
	(0.09)	(0.14)
Military Legacy	2.86**	2.46
	(1.01)	(1.36)
NVR	0.64	0.20**
	(0.24)	(0.15)
Observations	1448	77
AIC	558.60	104.38
BIC	600.82	123.13
Log lik.	-271.30	-44.19
Chi-squared	45.59	14.41

*Odds ratios; Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$*

As shown in model 1 in table A7.1, there appears to be no significant impact of NVR on coup attempts. However, the estimated coefficient for NVR in model 2 indicates a negative effect on the success of coups. Figure A7.22 illustrates the substantial size of the effect based on the estimates provided by model 2.

Table A7.22. Probability of coup success for regimes induced by NVR and regimes without NVR



The findings confirm the descriptive results provided in the main text. Coups in democratic regimes induced by NVR have a success rate of about 24%. Coups in regimes that democratized without the help of NVR succeed at a rate of approximately 61%. To conclude, the probability of a coup attempt after transition is not much different for democratic regimes that came about by means of NVR and regimes without this feature. However, the success rate of post-transition coups is substantially lower in NVR-induced regimes.

References

Aleman, Jose & David D. Yang (2011) A Duration Analysis of Democratic Transitions and Authoritarian Backslides. *Comparative Political Studies* 44(9): 1123-1151.

Austin, Peter C. (2011) An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research* 46(3): 399-424.

Austin, Peter C. (2014) The use of propensity score methods with survival or time-to-event outcomes: reporting measures of effect similar to those used in randomized experiments. *Statistics in Medicine* 33: 1242-1258.

Boix, Carles (2011) Democracy, development, and the international system. *American Political Science Review* 105(4): 809-828.

Boix, Carles & Susan Carol Stokes (2003) Endogenous democratization. *World Politics* 55(4): 517-549.

Celestino, Mauricio Rivera & Kristian Skrede Gleditsch (2013) Fresh carnations or all thorn, no rose? Nonviolent campaigns and transitions in autocracies. *Journal of Peace Research* 50(3): 385-400.

Cheibub, Jose A. (2007) *Presidentialism, parliamentarism, and democracy.* Cambridge: Cambridge University Press.

Chenoweth, Erica & Orion A. Lewis (2013) Unpacking nonviolent campaigns: Introducing the NAVCO 2.0 Dataset. *Journal of Peace Research* 50(3): 415-423.

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

Gassebner, Martin; Michael J. Lamla & James R. Vreeland (2012) Extreme Bounds of Democracy. *Journal of Conflict Resolution* 57(2): 171-197.

Geddes, Barbara; Joseph Wright & Erica Frantz (2014) Autocratic Breakdown and Regime Transitions: A New Data Set. *Perspectives on Politics* 12(2): 313-331.

Gelman, Andrew & Jennifer Hill (2006) *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.

Gleditsch, Kristian Skrede & Michael D. Ward (2006) Diffusion and the international context of democratization. *International Organization* 60(4): 911-933.

Guo, Shenyang & Mark W. Fraser (2010) *Propensity Score Analysis: Statistical Methods and Applications*. Thousand Oaks, CA: SAGE Publications.

Hansen, Ben B. & Stephanie Olsen Klopfer (2006) Optimal full matching and related designs via network flows. *Journal of Computational and Graphical Statistics* 15(3): 609-627.

Ho, Daniel E.; Kosuke Imai, Gary King & Elizabeth A. Stuart (2007) Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference.

Political Analysis 15(3): 199-236.

Holland, Paul W. (1986) Statistics and causal inference. *Journal of the American Statistical Association* 81(396): 945-960.

Lunt, Mark (2014) Selecting an Appropriate Caliper Can Be Essential for Achieving Good Balance with Propensity Score Matching. *American Journal of Epidemiology* 179(2): 226-235.

Maeda, Ko (2010) Two modes of democratic breakdown: A competing risks analysis of democratic durability. *The Journal of Politics* 72(4): 1129-1143.

Maoz, Zeev (1996) *Domestic sources of global change.* Cambridge: Cambridge University Press.

Park, Sunhee & David J. Hendry (2015) Reassessing Schoenfeld Residual Tests of Proportional Hazards in Political Science Event History Analyses. *American Journal of Political Science* 59(4): 1072-1087.

Przeworski, Adam (2000) *Democracy and development: political institutions and well-being in the world, 1950-1990.* Cambridge: Cambridge University Press.

Rosenbaum, Paul R. & Donald B. Rubin (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1): 41-55.

Rubin, Donald B. (2005) Causal inference using potential outcomes. *Journal of the American Statistical Association* 100(469): 322-331.

Rueschemeyer, Dietrich; Evelyne Huber Stephens & John D. Stephens (1992) *Capitalist development and democracy*. Cambridge, Chicago: Polity Press.

Sekhon, Jasjeet S. (2009) Opiates for the matches: Matching methods for causal inference. *Annual Review of Political Science* 12: 487-508.

Sing, Ming (2010) Explaining democratic survival globally (1946-2002). *The Journal of Politics* 72(2): 438-455.

Singer, J. David (1987) Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816-1985. *International Interactions* 14: 115-132.

Splawa-Neyman, Jerzy (1990 [1923]) On the application of probability theory to agricultural experiments. Essay on principles. Section 9. *Statistical Science* 5(4): 465-472, translated by D. M. Dabrowska and T. P. Speed.

Stuart, Elizabeth A. (2010) Matching methods for causal inference: A review and a look forward. *Statistical Science: A Review Journal of the Institute of Mathematical Statistics* 25(1): 1-21.

Stuart, Elizabeth A. & Donald B. Rubin (2008) Best practices in quasi-experimental designs. In: Jason W. Osborne (ed.) *Best practices in quantitative methods*. Thousand Oaks, CA: Sage:155-176.

Teorell, Jan; Stefan Dahlberg, Sören Holmberg, Bo Rothstein, Felix Hartmann & Richard Svensson (2015) *The Quality of Government Standard Dataset, version Jan15*. University of Gothenburg: The Quality of Government Institute.

Ulfelder, Jay (2010) *Dilemmas of Democratic Consolidations. A Game-theory Approach*. Boulder, CO: First Forum Press.

Ulfelder, Jay (2012) *Democracy/Autocracy Data Set (DAD) Handbook,* Harvard Dataverse, V1 (http://hdl.handle.net/1902.1/18836).