

Regional borders and the importance of ethnicity in national politics

Roland Hodler (with Richard Bluhm and Paul Schaudt)

CEPR Workshop on Ethnicity, Identity and Conflict, June 2025

Zooming in

Political economics of development: From countries to regions

- ▶ Strong focus on country-level outcomes in the early 2000s.
- ▶ Spatial data led to a focus on regional outcomes around 10 years ago (e.g., Hodler and Raschky 2014).
- ▶ Scholars are now looking at ever smaller spatial units.

Objective for my current research in this field:

- ▶ Keep the regional focus, but bring back country-wide economic and political interactions.

Regional borders:

- ▶ Regional border designs may affect economic, political, and social outcomes at the regional and the national level; and they can be changed relatively easily.
- ▶ However, little is known about these effects in economics and other social sciences.
- ▶ Today, I focus on political effects of border designs.

Motivation

- ▶ Ethnic identities shape political organization and mobilization in many countries.
- ▶ Ethnic politics, e.g., ethnic favoritism/patronage and ethnic voting, often leads to political instability and conflict.

Figure: Ethnic violence after the 2007 Kenyan elections



Source: The Guardian (2010)

- ▶ How can regional borders affect the importance of ethnicity in national politics?
- ▶ More specifically, should regional borders follow a country's ethnic geography?

One-slide summary

What we do:

1. Introduce a new framework for measuring and classifying regional border designs and reforms at the micro level.
2. Exploit the quasi-experimental setting created by a constitutional reform in Kenya to estimate the effect of border changes on ethnic voting in national elections.
3. Identify the border design that would have minimized ethnic voting (and simulate the effects of alternative reform proposals).

What we find:

- ▶ Kenya's border reform increased ethnic voting.
- ▶ Ethnofederal border reforms, which increase the alignment between political borders and ethnic geography, would have reduced ethnic voting.

What we conclude:

- ▶ Border designs should be seen as a nation-building policy for reducing the importance of ethnicity in national politics.

Contributions to the literature

Nation-building, e.g., via large-scale resettlement programs (e.g., Rohner and Zhuravskaya 2024; Bazzi et al. 2019, Carlitz et al. 2024).

⇒ Political border reforms are feasible, fast, and not too invasive.

Political borders and ethnic diversity (e.g., Lipset 1960, Lijphart 1977, Horowitz 1985, Alesina et al. 2011, Michalopolus & Papaioannou 2016, Bazzi & Gudgeon 2021).

⇒ New measures to characterize border designs at the subnational level.

Ethnic politics in Kenya and beyond (e.g., Padró i Miquel 2007, Franck & Rainer 2012, Burgess et al. 2015, Kramon & Posner 2016, De Luca et al. 2018, Marx et al. 2021, Amodio et al. 2024).

⇒ Evidence that ethnofederal border reforms help to reduce ethnic politics.

Decentralization, federalism, and political accountability (e.g., Tiebout 1956, Alesina & Spolaore 1997/2005, Boffa et al. 2015, Narasimhan & Weaver 2024).

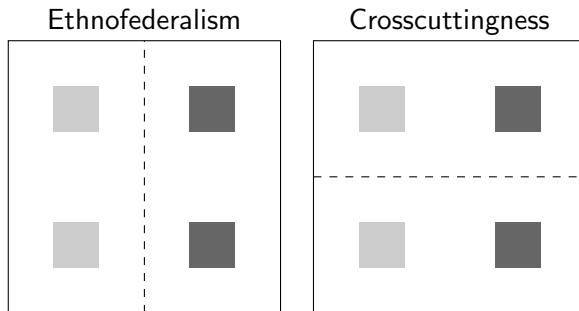
⇒ The effects of devolution and decentralization depend on border designs.

Road map

- ▶ Framework and new measures
- ▶ Background: Kenya and its 2010 constitutional reform
- ▶ Data and empirical strategy
- ▶ Main results, robustness, and mechanism
- ▶ Optimal border design
- ▶ Conclusions

Prominent border designs

Border designs for a stylized country with 4 locations and 2 groups:



Reasons why these border designs may reduce ethnic tensions at the national level:

- ▶ Ethnofederal designs (paired with political devolution and fiscal decentralization): Tranquilizing effect by moving sensitive issues like education or language policy to the regional level (Horowitz 1985).
- ▶ Crosscutting designs: Alternative cleavages.

Measurement

- ▶ **Regional fractionalization:** Probability that two randomly selected individuals from a given political region r identify with different ethnic groups e :

$$RF_r = 1 - \sum_{e=1}^n (s_e^r)^2,$$

where s_e^r is the share of individuals living in region r who identify with group e .

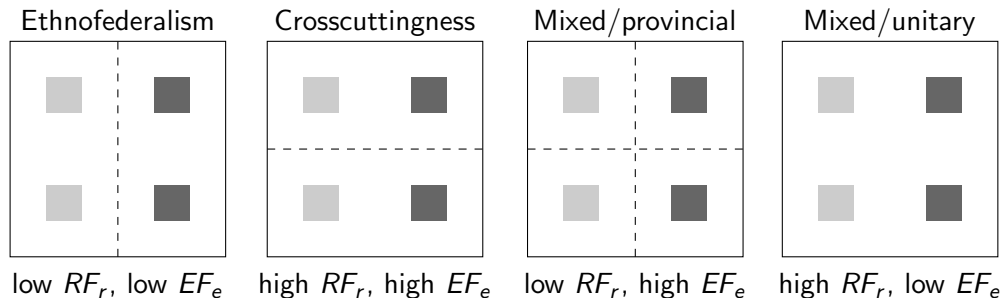
- ▶ **Ethnic fragmentation:** Probability that two randomly selected individuals of a given ethnic group e live in different political regions r :

$$EF_e = 1 - \sum_{r=1}^m (s_r^e)^2,$$

where s_r^e is the share of individuals identifying with group e who live in region r .

Measurement

Our stylized country:



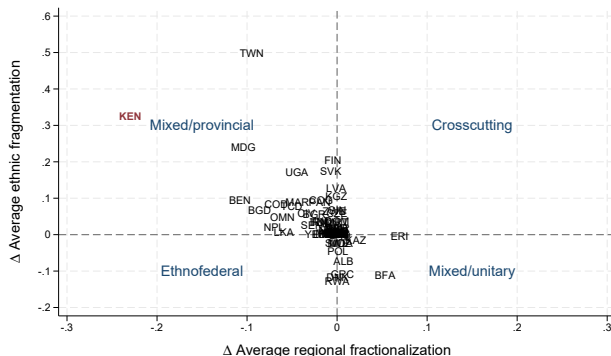
Real-world countries:

- ▶ Regional fractionalization RF_r differs across regions.
- ▶ Ethnic fragmentation EF_e differs across ethnic groups.

Applications of our measures

1. Country-level classification of border designs or reforms based on average regional fractionalization and average ethnic fragmentation or changes thereof:

Figure: Cross-country comparisons of border reforms (1995–2015)



2. Micro-level analyses on the effects of border designs or reforms by exploiting variation in RF_r and EF_e across individuals depending on their location and their ethnic identity.

Background: Kenya

Demographics and ethnic politics:

- ▶ 55m inhabitants, 40⁺ ethnic groups/tribes.
- ▶ Four large groups (Kikuyu, Kalenjin, Luo, Kamba) compete for power with (vice) presidential candidates in most elections.
- ▶ Ethnicity is salient in national politics: ethnic voting, ethnic favoritism/patronage, inter-group violence.
- ▶ Before reform: strong executive and high centralization; provinces as main subnational units.

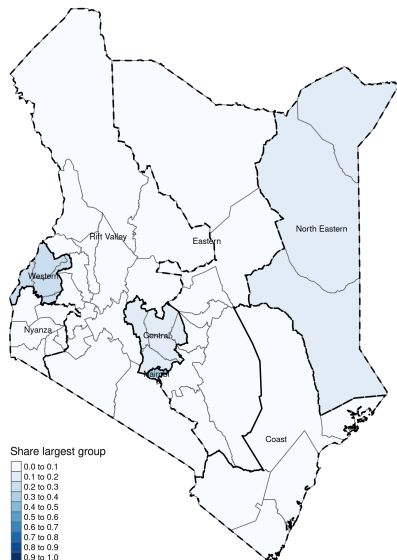
Ethnic violence and constitutional change:

- ▶ Ethnic violence after the 2007 election: 1,000 died and hundreds of thousands were temporarily internally displaced.
- ▶ African Union mediated a power-sharing deal between the main rivals – Kibaki (Kikuyu) and Odinga (Luo) – and pushed for a constitutional reform.

Background: 2010 constitutional reform

Territorial reform from 8 provinces to 47 counties:

- ▶ Many different (old and new) border reform proposals. $\Rightarrow$ Final border design hard to anticipate before 2010.
- ▶ County borders similar to last borders drawn by the British. $\Rightarrow$ Borders not gerrymandered based on contemporary political motivations.
- ▶ Counties are nested within former provinces.



▶ More maps

Background: 2010 constitutional reform (cont.)

Political devolution and fiscal decentralization:

- ▶ Elections of county governors and county assemblies.
- ▶ 15% of central government revenues to counties (formula-based).
- ▶ Devolution of responsibilities: health care, education, and roads.

Timing:

- ▶ Proposed in 2010 and supported by 68% in a national referendum in 2010.
- ▶ Implemented with the 2013 elections.

Data: Regional fractionalization and ethnic fragmentation

1989 Kenya Population and Housing Census:

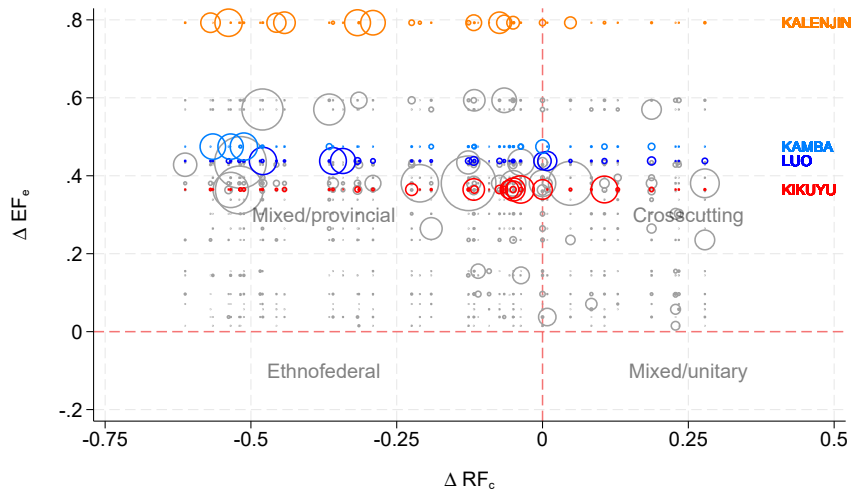
- ▶ Geocoded 5% sample, representative at the level of 3,600 sublocations.
- ▶ Last census that discloses ethnicity in microdata.

Treatment variables:

- ▶ ΔRF_c : Difference between post-reform county-level RF_c and pre-reform province-level $RF_{p(c)}$.
- ▶ ΔEF_e : Difference between an ethnic group's EF_e given post-reform county borders and its EF_e given pre-reform province borders.

Data: Regional fractionalization and ethnic fragmentation

Figure: Identifying variation



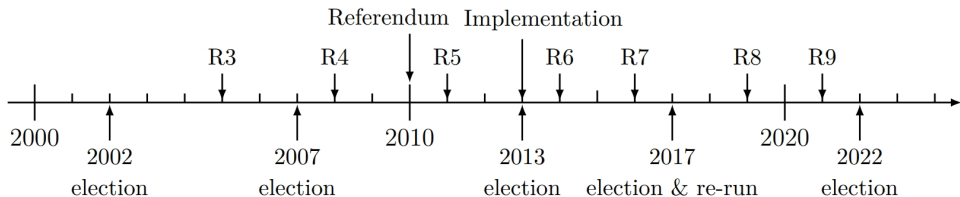
Notes: The figure plots ΔRF_c and ΔEF_e for each county-ethnic group combination. The size of the circles is proportional to the population of this combination.

Data: Ethnic voting

Afrobarometer survey rounds 3–9:

- ▶ Key question: “If presidential elections were held tomorrow, which party’s candidate would you vote for?”
- ▶ Similar group-level voting patterns as in the exit poll data.
- ▶ Timeline:

▶ Figure



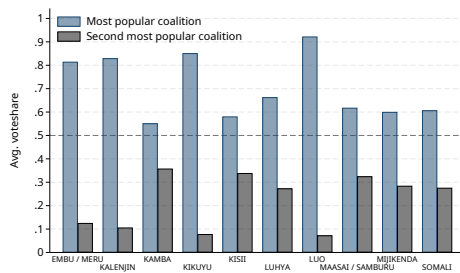
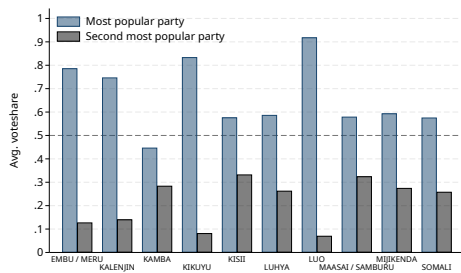
▶ Spatial coverage

Data: Ethnic voting

Main dependent variable: Plurality indicator for ethnic voting

- ▶ $EV_{icet} = 1$ if respondent i would vote for the party (coalition) for which the highest number of other respondents of ethnic group e would vote in t ; and 0 otherwise.
- ▶ Suitable given the voting cohesion within ethnic groups, with most groups having a strongly preferred party (coalition) in most survey rounds.

Figure: Most popular parties and coalitions per group (time-averaged)



Empirical strategy

Generalized diff-in-diff specification:

$$EV_{icet} = \beta(\Delta RF_c \times D_t) + \gamma(\Delta EF_e \times D_t) + \mathbf{Z}'_{it}\boldsymbol{\xi} + \mu_{ce} + \lambda_t + \epsilon_{ijt}$$

- ▶ EV_{icet} : plurality indicator for ethnic voting
- ▶ D_t : indicator for post-reform Afrobarometer survey rounds
- ▶ ΔRF_c : change in regional fractionalization in county c
- ▶ ΔEF_e : change in ethnic fragmentation for group e
- ▶ $\mathbf{Z}_{it}$: proxy variables for other cleavages (urban/rural, age, gender, economic status)
- ▶ μ_{ce} : county-by-ethnicity-fixed effects, which absorb all time-invariant differences across groups and counties and combinations thereof (including $\overline{EV_{ce}}$)
- ▶ λ_t : survey round-fixed effects, which absorb trends in ethnic voting and uniform effects of devolution and decentralization
- ▶ ϵ_{ijt} : standard errors two-way clustered on the province-by-ethnicity and the county level

Empirical strategy

Coefficients of interest:

- ▶ β and γ capture the effects of the reform-induced changes in RF_c and EF_e on ethnic voting (given political devolution and fiscal decentralization).

▶ More

Tests:

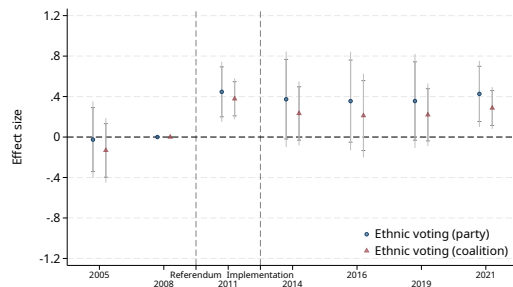
- ▶ No pre-trends (see next slide).
- ▶ No systematic relation between treatments and respondent characteristics.
- ▶ No effects of treatments on intention to vote and willingness to reveal party preferences.

▶ Figure

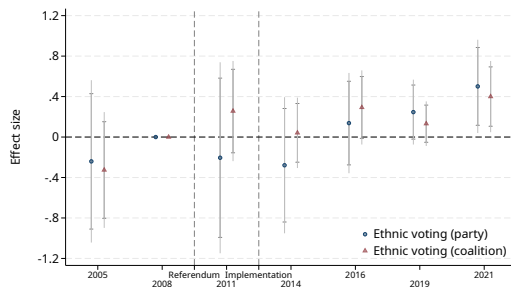
▶ Figure

Results: Event study

Figure: Event study estimates



(a) Time-varying effects of ΔRF_c



(b) Time-varying effects of ΔEF_e

Notes: The figure shows the coefficients from estimating $EV_{it} = \sum_t \beta_t \Delta RF_c + \sum_t \gamma_t \Delta EF_e + \mathbf{Z}'_{it} \boldsymbol{\xi}_t + \mu_{ce} + \lambda_t + \epsilon_{it}$, where t are survey years and $t = 2008$ the omitted reference category. Panel A reports the β_t 's, and panel B the γ_t 's. Standard errors are two-way clustered at the province-by-ethnicity and the county level. 95% and 90% confidence intervals are plotted as gray bars.

Results: Diff-in-diff

Table: Main results: Difference-in-differences

	<i>Dependent variables: Ethnic voting</i>					
	<i>Party-based plurality</i>			<i>Coalition-based plurality</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta RF_c \times D_t$	0.377** (0.172)	0.379** (0.175)	0.353* (0.196)	0.307*** (0.087)	0.307*** (0.089)	0.278*** (0.100)
$\Delta EF_e \times D_t$	0.341** (0.169)	0.335* (0.168)	0.301* (0.160)	0.433*** (0.100)	0.428*** (0.098)	0.397*** (0.093)
County-by-ethnicity FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Controls	—	✓	✓	—	✓	✓
Controls $\times D_t$	—	—	✓	—	—	✓
Observations	6357	6357	6357	6357	6357	6357

Notes: The table reports the results of regressing our plurality indicators of ethnic voting (party-based plurality in columns 1–3 and coalition-based plurality in columns 4–6) on interaction terms between the post-treatment indicator D_t and the changes in regional fractionalization (ΔRF_c) and ethnic fragmentation (ΔEF_e). All specifications include county-by-ethnicity and time-fixed effects and control for the age of a respondent, indicators if the respondent lives in an urban survey cluster, is female, owns a TV, radio, or motorized vehicle, has completed primary, secondary, or post-secondary education. Standard errors are two-way clustered at the province-by-ethnicity and county levels. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Results: Assessing the border reform

Table: Aggregate effects of border reform on ethnic voting

Ethnic group	Predicted change in ethnic voting due to			Population share
	ΔEF	ΔRF	both	
Kikuyu	0.122	-0.015	0.107	0.208
Luhya	0.128	-0.024	0.104	0.148
Luo	0.147	-0.103	0.043	0.124
Kamba	0.159	-0.171	-0.012	0.114
Kalenjin	0.266	-0.122	0.144	0.113
Embu/Meru	0.145	-0.135	0.010	0.062
Kisii	0.122	-0.179	-0.057	0.062
Mijikenda	0.191	-0.134	0.057	0.046
Maasai/Samburu	0.199	-0.048	0.151	0.023
Turkana	0.144	-0.164	-0.020	0.013
Predicted change across all groups:				0.066

Notes: The table reports the predicted aggregate changes in ethnic voting for the ten largest groups and the country as a whole. These changes are computed by multiplying the diff-in-diff coefficients in column (2) of the main table by the group-county-specific populations.

Results: Robustness tests

Context-specific robustness tests:

- ▶ Controlling for other administrative-territorial changes. [▶ Table](#)
- ▶ Accounting for possibly heterogenous effects of political devolution and fiscal decentralization. [▶ Table](#)

Standard robustness tests:

- ▶ Logistic regressions and marginal effects. [▶ Table](#)
- ▶ Alternative ethnic voting measures, e.g., majority indicator or parties/coalitions with co-ethnic (vice) presidential candidates. [▶ Figure](#)
- ▶ Sample perturbations, e.g., Afrobarometer or census weights, or exclusion of Rift Valley. [▶ Figure](#)
- ▶ Alternative clustering of the standard errors. [▶ Figure](#)

Mechanism: Why can ethnofederal reforms reduce ethnic voting?

- ▶ Ethnofederal reforms create relatively homogenous and sizable counties.
- ▶ The largest group in a county controls the county executive. [▶ Table](#)
- ▶ Given political devolution and fiscal decentralization, regional control allows this group to meet its group-specific preferences with regional public goods.
- ▶ This “tranquilizes” ethnic tensions in national politics: The effects of ΔRF_c and ΔEF_e are mainly driven by respondents of majority groups with regional control.

[▶ Table](#)

Optimal border design

We know:

- ▶ Low regional fractionalization RF_r and low ethnic fragmentation EF_e reduce ethnic voting.
- ▶ Trade-off: Having more regions tends to decrease RF_r but to increase EF_e .

We want: Border design that minimizes ethnic voting

- ▶ Our coefficient estimates suggest balancing the countervailing effects with weights $w = 0.531$ for RF_r and $(1 - w)$ for EF_e .
- ▶ Constraints on the border design:
 - ▶ New regions must be contiguous.
 - ▶ New regions must be nested within former provinces.
 - ▶ New regions must not subdivide 70 districts.
- ▶ 30.4 million possible maps with $K \in \{8, 9, 10, \dots, 70\}$ contiguous regions.

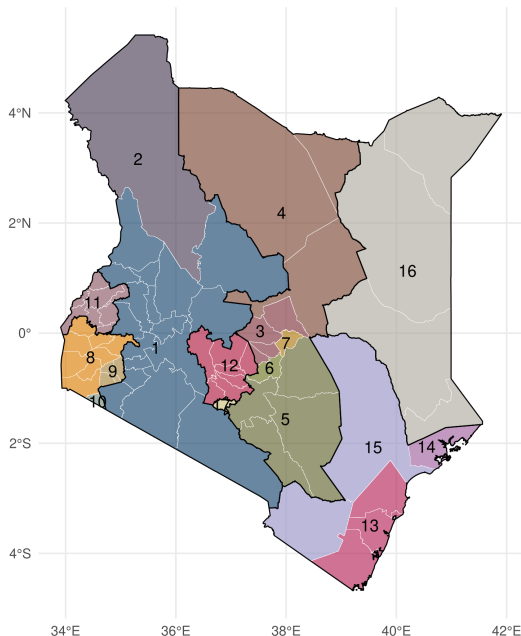
Optimal border design

Optimal map with 17 regions:

- ▶ Predicted change in ethnic voting:
–5.6 p.p. compared to provinces,
–12.2 p.p. compared to counties.
- ▶ Recipe for success: Reduce regional fractionalization without fragmenting groups.

	Reform	Our map
Average ΔRF	– 0.21	– 0.17
Average ΔEF	+ 0.44	+ 0.02

▶ Evaluation of alternative reform proposals



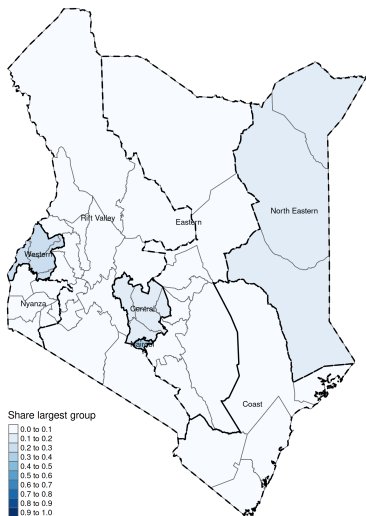
Conclusions

- ▶ We introduce two indices – regional fractionalization and ethnic fragmentation – to measure and classify border designs and reforms at the subnational level.
- ▶ We show that Kenya's reform increased ethnic voting, but propose a more ethnofederal reform that would have reduced ethnic voting.
- ▶ We argue that border designs are an important nation-building policy that can reduce the importance of ethnicity in national politics.

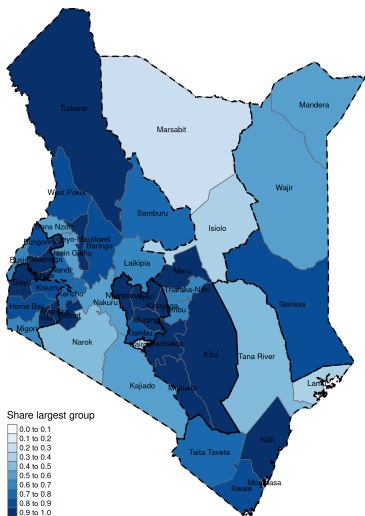
Thank you for the invitation and your attention.

Feedback most welcome.

Appendix: Regions before and after the 2010 constitutional reform



(a) Pre-reform provinces



(b) Post-reform counties

Appendix: Presidential elections in Kenya

Presidential elections: Two-round system

- (1) Win with 50% of the popular vote and 25% in at least half the regions.
- (2) Runoff among top two candidates (if no winner in 1st round).

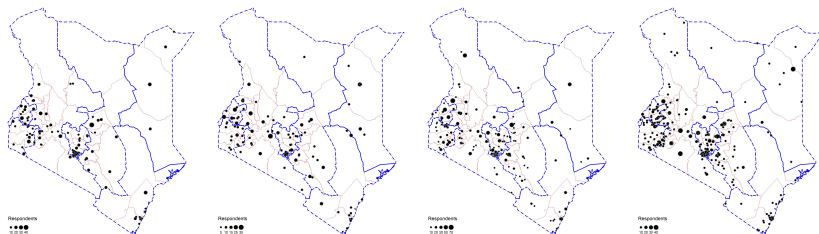
Presidential elections and presidential/vice-presidential candidates:¹

- ▶ 2002 election: **Moi** term-limited. **Kibaki** defeats **Kenyatta**.
- ▶ 2007 election: **Kibaki** narrowly defeats **Odinga** (and **Musyoka**) in a flawed election, leading to inter-group violence.
- ▶ 2010 unity government: **Kibaki**/**Odinga**.
- ▶ 2013 election: **Kenyatta**/**Ruto** defeats **Odinga**/**Musyoka**.
- ▶ 2017 election: **Kenyatta**/**Ruto** defeats **Odinga**/**Musyoka**, but Supreme Court nullifies the result. **Odinga** boycotts rerun.
- ▶ 2022 election: **Ruto**/**Gachagua** narrowly defeats **Odinga**/**Karua**.

[◀ Go back](#)

¹Big 4 groups: Kalenjin, Kamba, Kikuyu, and Luo.

Appendix: Afrobarometer coverage

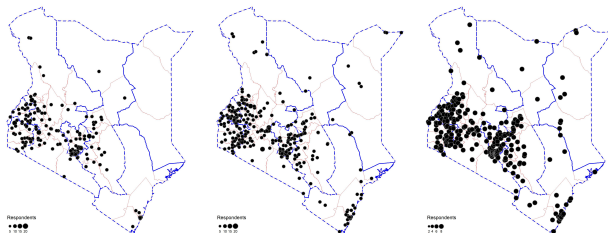


(a) R3 (2005)

(b) R4 (2008)

(c) R5 (2011)

(d) R6 (2014)

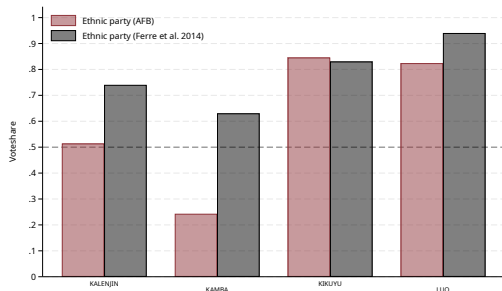


(e) R7 (2016)

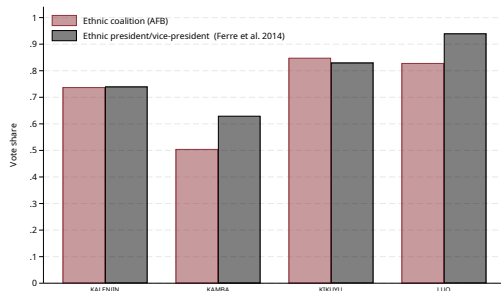
(f) R8 (2019)

(g) R9 (2021)

Appendix: Ethnic voting in Afrobarometer and exit polls



(a) Parties



(b) Coalitions

Notes: Panel A plots the share of respondents from the big four ethnic groups who intend to vote for the party with a coethnic presidential or vice-presidential candidate according to the 2014 Afrobarometer survey data (in cyan-green) and the share of voters from these groups who voted for the presidential ticket with coethnic presidential or vice-presidential candidate according to the exit polls by Ferree et al. (2014, in light brown). Panel B replicates panel A but aggregates the political parties at the level of coalitions.

[Go back](#)

Appendix: Summary statistics

Table: Summary statistics

Variable	Mean	SD	Min	Max	N
<i>Panel A: Variables at respondent level</i>					
Ethnic voting: party-based plurality	0.71	0.46	0.00	1.00	7,262
Ethnic voting: coalition-based plurality	0.75	0.43	0.00	1.00	7,262
ΔRF_c	-0.20	0.24	-0.61	0.28	7,262
ΔEF_e	0.45	0.16	0.02	0.79	7,262
Age	36.00	13.79	18.00	96.00	7,249
Female	0.47	0.50	0.00	1.00	7,262
Urban cluster	0.37	0.48	0.00	1.00	7,262
Respondent owns a radio	0.87	0.34	0.00	1.00	7,262
Δ Regional population	-0.72	0.28	-0.98	0.00	7,262
Δ Capital proximity	-0.46	0.34	-0.94	0.00	7,262
Support referendum	0.58	0.49	0.00	1.00	7,262
Largest group in county	0.72	0.45	0.00	1.00	7,262
Unitary government (county)	0.38	0.48	0.00	1.00	7,262
<i>Panel B: Variables at county-by-ethnicity level</i>					
Largest group (county)	0.03	0.17	0.00	1.00	1,666
Coethnic governors (initial)	0.03	0.17	0.00	1.00	1,666
County executive board (CEC) share	0.03	0.14	0.00	1.00	1,666
County public service board (CPSB) share	0.03	0.14	0.00	1.00	1,666
CEC & CPSB share	0.03	0.14	0.00	1.00	1,666

Appendix: Identification assumptions

Identification assumptions: Strong parallel trends

- ▶ **Standard parallel trends:** Unobserved time-varying confounds behave similarly in groups or counties that are treated differently by the border reform in terms of ΔRF_c or ΔEF_e .
- ▶ **No selection bias:** Respondents with higher potential treatment effects do not select into higher intensities of ΔRF_c or ΔEF_e .

If these assumptions are satisfied, TWFE estimation of β and γ identifies weighted averages of the causal responses at different treatment intensities of ΔRF_c or ΔEF_e .

◀ Go back

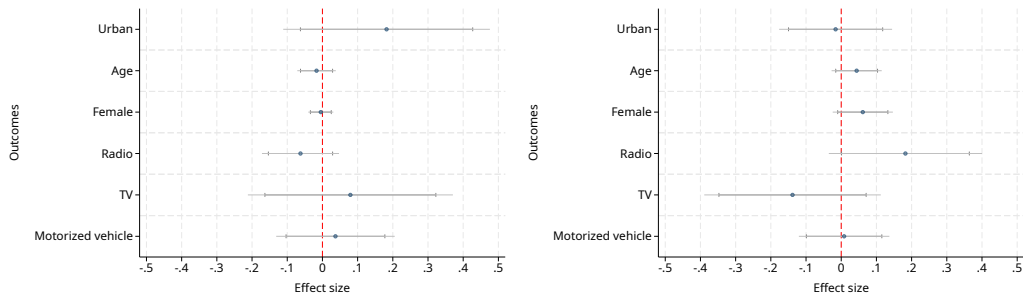
Appendix: Identification assumptions (cont.)

- ▶ TWFE estimation does not always recover treatment effects in diff-in-diff and event study designs with staggered treatment adoption (e.g., Callaway and Sant'Anna 2021, Sun and Abraham 2021).
- ▶ In our case, treatments occur simultaneously, but their intensity varies across units.
- ▶ Callaway et al. (2024) show that the weights on the underlying causal responses are always positive in this case if the data are a proper panel. We interpret our repeated cross-section estimates along these lines.

◀ Go back

Appendix: Balancing checks

Figure: Balancing results for individual characteristics



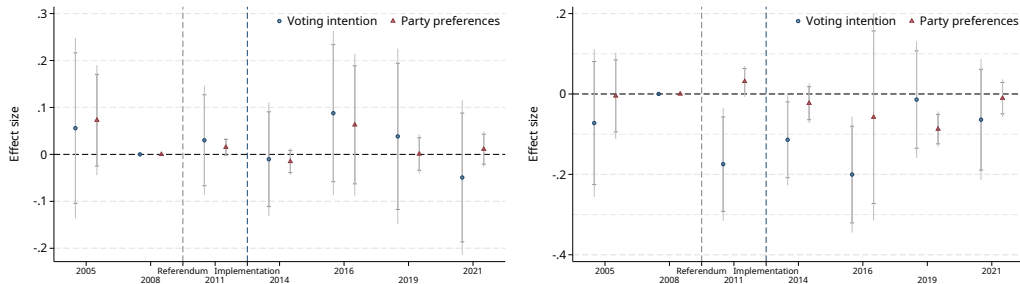
(a) Changes in regional fractionalization (ΔRF_c) (b) Changes in ethnic fragmentation (ΔEF_e)

Notes: The figure summarizes the balancing tests on individual characteristics for our treatments. Panels A and B report the point estimates and 95% confidence intervals of a change in regional fractionalization and ethnic fragmentation over different individual controls. The 95% confidence intervals based on two-way clustered standard errors at the county and province-by-ethnicity level are plotted as gray error bars.

[Go back](#)

Appendix: Intention to vote and willingness to reveal party preferences

Figure: Intention to vote and willingness to reveal party preferences



(a) Effects of changes in regional fractionalization (b) Effects of changes in ethnic fragmentation

Notes: The figure shows event study coefficients of reform-induced changes in regional fractionalization and ethnic fragmentation on whether a respondent indicates an intention to vote (circles) and whether they are willing to reveal the party's candidate they intend to vote for (triangles). The event study specifications include county-by-ethnicity and time-fixed effects as well as the following individual-level control variables: age and indicator variables for residence in an urban cluster, being female, owning a radio, TV, and motorized vehicle. 95% confidence intervals based on standard errors clustered at the province-by-ethnicity and county level are plotted as gray error bars.

[Go back](#)

Appendix: Other boundary changes and devolution

Other boundary changes:

- ▶ Constitutional reform increased the number of electoral constituencies from 210 to 290.
- ▶ Constitutional reform abolished 175 elected local authorities and subsumed them into county governments.
- ▶ High Court declared 210 districts created by Moi and Kibaki illegal in 2009.

Heterogeneous effects of political devolution and fiscal decentralization:

- ▶ Controlling for changes in size (population) of regions, changes in proximity to regional capital, and support for the referendum.

◀ Go back

Appendix: Other boundary changes and devolution (cont.)

Table: Controlling for other constitutional changes

	Dependent variable: Ethnic voting (party-based plurality)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta RF_c \times D_t$	0.384** (0.179)	0.380** (0.185)	0.356* (0.188)	0.386* (0.199)	0.362* (0.188)	0.409*** (0.141)
$\Delta EF_e \times D_t$	0.374** (0.167)	0.334* (0.170)	0.325* (0.182)	0.332** (0.164)	0.346** (0.170)	0.334*** (0.123)
$\Delta RF_c(Constituencies) \times D_t$	0.006 (0.079)					
$\Delta EF_e(Constituencies) \times D_t$	0.678** (0.279)					
$\Delta RF_c(Local\ authority) \times D_t$		-0.019 (0.037)				
$\Delta EF_e(Local\ authority) \times D_t$		-0.005 (0.388)				
$\Delta RF_c(District) \times D_t$			-0.009 (0.050)			
$\Delta EF_e(District) \times D_t$			0.073 (0.381)			
$\Delta\ A dm.\ unitsize(pop) \times D_t$				-0.013 (0.079)		
$\Delta\ Capital\ proximity \times D_t$					0.038 (0.069)	
$Support\ referendum \times D_t$						0.130* (0.065)
County-by-ethnicity FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Observations	6174	6332	6072	6357	6357	6357

Appendix: Logistic regressions and marginal effects

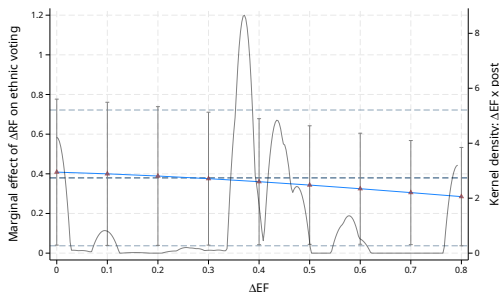
Table: Logit regressions and OLS with triple-interaction

	<i>Dependent variables: Ethnic voting (party-based plurality)</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta RF_c \times D_t$	2.062** (0.925)	2.066** (0.945)	5.064*** (1.935)	5.121*** (1.951)	0.373** (0.163)	0.371** (0.166)
$\Delta EF_e \times D_t$	1.584* (0.931)	1.533* (0.931)	-0.202 (1.242)	-0.294 (1.230)	0.245 (0.150)	0.230 (0.150)
$\Delta RF_c \times \Delta EF_e \times D_t$			-6.617* (3.388)	-6.740** (3.351)	-1.065 (0.674)	-1.099 (0.663)
Est-M.	Logit	Logit	Logit	Logit	LPM	LPM
County-by-ethnicity FE	✓	✓	✓	✓	✓	✓
Time-FE	✓	✓	✓	✓	✓	✓
Controls	—	✓	—	✓	—	✓
Observations	6143	6121	6143	6121	6379	6357

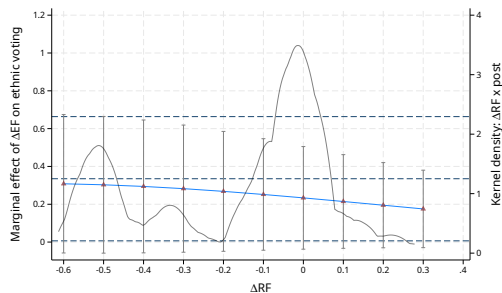
Notes: The table reports the regression results of regressing our party-based plurality indicator of ethnic voting on the post-treatment indicator times the change in regional fractionalization ($\Delta RF_C \times D_t$), and the change ethnic fragmentation ($\Delta EF_e \times D_t$). All specifications include county-by-ethnicity and time-fixed effects and control for the age of a respondent, indicators if the respondent lives in an urban survey cluster, is female, owns a TV, radio, or motorized vehicle, has completed primary, secondary, or post-secondary education. Standard errors are two-way clustered at the province-by-ethnicity and county levels. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix: Logistic regressions and marginal effects (cont.)

Figure: Marginal effects



(a) Effect of ΔRF_c



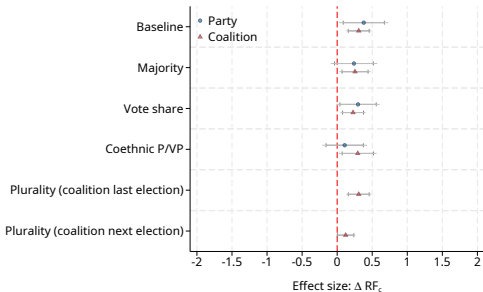
(b) Effect of ΔEF_e

Notes: The figure plots the marginal effects on $\Delta RF \times D_t$ (panel A) and $\Delta EF \times D_t$ (panel B), over the empirical distribution of the respective other measure on party-based ethnic voting. The plots are based on the logit specification depicted in column (2) on the previous slide.

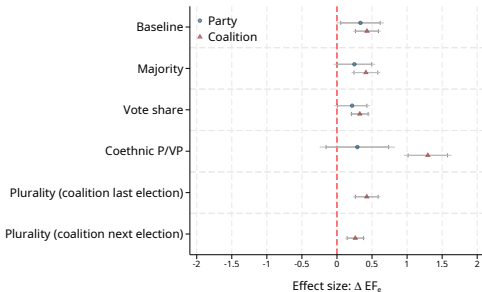
[Go back](#)

Appendix: Alternative measures of ethnic voting

Figure: Robustness tests with alternative voting outcomes



(a) Effect of changes in regional fractionalization

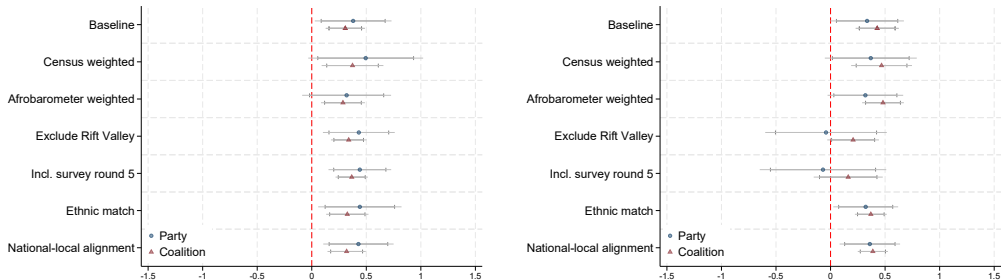


(b) Effect of changes in ethnic fragmentation

Notes: The figure checks the robustness of our main results (see main table, columns 2 and 5) to alternative measures of ethnic voting. Panel A shows estimates for regional fractionalization, and panel B for ethnic fragmentation. Blue circles indicate party-based measures of ethnic voting, and red triangles coalition-based measures. Baseline indicates the use of our plurality indicator of ethnic voting. Majority indicates the use of a majority indicator of ethnic voting, which is equal to one only for parties/coalitions that receive 50% of more of the vote share of an ethnic group in a survey round. Group share indicates the use of the vote share of the plurality party. Co-ethnic P/VP indicates the use of an indicator of voting for a party headed by a co-ethnic presidential or vice-presidential candidate (only defined for the Kalenjin, Kamba, Kikuyu, and Luo). Plurality (coalition past election) and plurality (coalition next election) indicate use of a coalition-based plurality indicator based on coalitions in the past or the next election (rather than the coalition known at the survey date). 95% confidence intervals based on two-way clustered standard errors at the province-by-ethnicity and county level are plotted as gray error bars.

Appendix: Sample perturbations

Figure: Robustness tests with perturbed samples

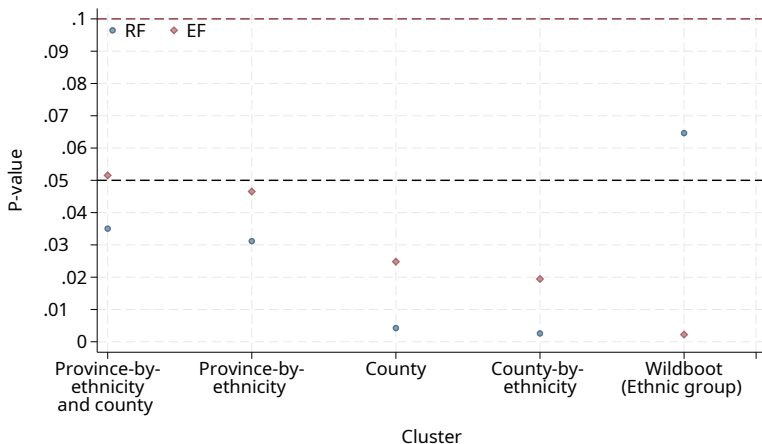


(a) Effect of changes in regional fractionalization (b) Effect of changes in ethnic fragmentation

Notes: The figure checks the robustness of our main results (see main table, columns 2 and 5) to samples. Panel A shows estimates for regional fractionalization, and panel B for ethnic fragmentation. Census weighted weights each Afrobarometer respondent in proportion to the census population. Afrobarometer weights indicate the use of the Afrobarometer survey weights in the regression. Exclude Rift Valley indicates that we exclude all counties of the former Rift Valley province. Incl. Survey round 5 indicates that we add survey round 5 as the first post-treatment period to the regression. Ethnic match indicates that we exclude all respondents whose language spoken at home does not match the language spoken at home by the Afrobarometer interviewer. National-local alignment indicates that we control for the alignment of governors in 2013 with the current co-ethnic presidential or vice-presidential candidate for an ethnic group within a county. 95% confidence intervals based on two-way clustered standard errors at the province-by-ethnicity and county level are plotted as gray error bars.

Appendix: Alternative clustering of standard errors

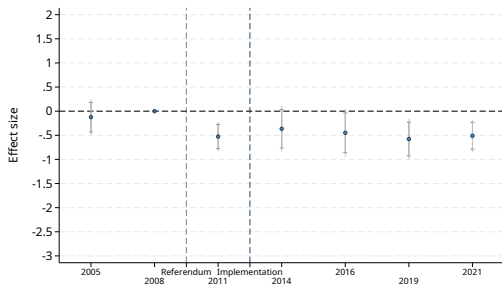
Figure: Alternative clustering of standard errors



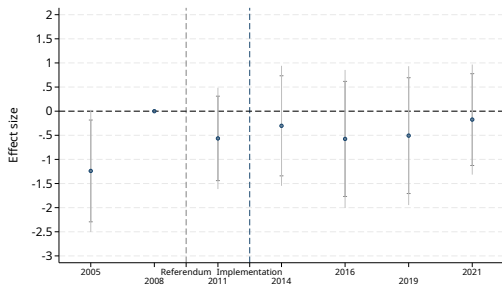
Notes: The figure reports estimated p-values of regional fractionalization (blue dots) and ethnic fragmentation (red diamonds) of our baseline specification (column 2 of main table) for alternative forms of standard error clustering or their calculation.

Appendix: Symmetric effect for changes in regional fractionalization

Figure: Effect symmetry



(a) Decreasing regional fractionalization



(b) Increasing regional fractionalization

Notes: The figure shows event study coefficients of the effect of reform-induced changes in regional fractionalization on ethnic voting split by whether regional fractionalization increases or decreases. Panel A shows estimates for reductions in regional fractionalization, and panel B shows estimates for increases in regional fractionalization. Note that **the negative ΔRF_c values in panel A are multiplied by -1 to report the effect of reductions in regional fractionalization**. All specifications include changes in ethnic fragmentation (ΔEF_e), county-by-ethnicity and time-fixed effects as well as the following individual level control variables: age and indicator variables for residence in an urban cluster, being female, owning a radio, TV, and motorized vehicle. Note that ethnic fragmentation increases for all groups, so we do not report the effect of ΔEF_e in this figure. 95% confidence intervals based on two-way clustered standard errors at the province-by-ethnicity and county level are plotted as gray error bars.

Appendix: Single component effects

Table: Single component effects

	<i>Dependent variables: Ethnic voting</i>					
	<i>Party-based plurality</i>			<i>Coalition-based plurality</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta RF_c \times D_t$	0.379** (0.174)	0.205* (0.106)		0.309*** (0.089)	0.219** (0.084)	
$\Delta EF_e \times D_t$	0.334* (0.168)		0.185 (0.269)	0.428*** (0.098)		0.394* (0.230)
County-by-ethnicity FE	✓	✓	✓	✓	✓	✓
Time FE	✓	—	—	✓	—	—
Ethnicity-by-time FE	—	✓	—	—	✓	—
County-by-time FE	—	—	✓	—	—	✓
Controls	✓	✓	✓	✓	✓	✓
Observations	6370	6370	6370	6370	6370	6370

Notes: In columns 1 and 4, we replicate our main specifications. In columns 2 and 5, we include ethnicity-by-time fixed effects to isolate the partial impact of a change in regional fractionalization. In columns 3 and 6, we include county-by-time fixed effects to isolate the partial impact of a change in ethnic fragmentation. Standard errors are two-way clustered at the province-by-ethnicity and county levels. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix: Largest groups and regional control

Table: Determinants of regional control (at the county-group level)

	<i>Dependent variables: Regional control</i>			
	<i>Coethnic governor</i> (1)	<i>CEC share</i> (2)	<i>CPSB share</i> (3)	<i>CEC & CPSB share</i> (4)
<i>Largest group_{ce}</i>	0.844*** (0.054)	0.333*** (0.082)	0.352*** (0.093)	0.341*** (0.086)
<i>Coethnic governor_{ce}</i>		0.417*** (0.086)	0.410*** (0.096)	0.414*** (0.090)
County FE	✓	✓	✓	✓
Ethnicity FE	✓	✓	✓	✓
Observations	1637	1637	1637	1637

Notes: The table reports the results from regressing four different proxies for political control at the county level on an indicator for the largest ethnic group in the county (all columns) and an indicator for having a coethnic governor (columns 2–4). The proxies for regional control are an indicator for having a coethnic governor (column 1), the share of coethnics in the county executive committee (CEC; column 2), the share of coethnics in the county public service board (CPSB; column 3), and the share across both the CEC and the CPSB (column 4). The regressions include county- and ethnicity-fixed effects, and standard errors are clustered at the county level. The dataset for these regressions are a county-by-ethnicity cross-section based on the census data and data from the 2016 diversity audit of the county administration. * $p < 0.1$. ** $p < 0.05$. *** $p < 0.01$.

Appendix: Regional control and ethnic voting

Table: Regional control and ethnic voting

Dependent variables: <i>Ethnic voting</i>	<i>party-based</i>	<i>coalition-based</i>	<i>party-based</i>	<i>coalition-based</i>
Interaction term: <i>Regional control</i>	<i>Largest group</i>		<i>Unitary government</i>	
	(1)	(2)	(3)	(4)
$\Delta RF_c \times D_t$	0.270*	0.091	0.419**	0.302***
	(0.156)	(0.172)	(0.169)	(0.087)
$\Delta EF_e \times D_t$	0.012	0.324**	0.130	0.271***
	(0.212)	(0.134)	(0.176)	(0.093)
$\Delta RF_c \times D_t \times \text{Regional control}_{ce}$	0.273	0.319	0.178	0.221**
	(0.243)	(0.204)	(0.168)	(0.083)
$\Delta EF_c \times D_t \times \text{Regional control}_{ce}$	0.549**	0.238	0.525**	0.409***
	(0.254)	(0.143)	(0.196)	(0.127)
County-by-ethnicity FE	✓	✓	✓	✓
Time FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Observations	6370	6370	6370	6370

Notes: The table reports the results from regressing the plurality indicator of ethnic voting (based on parties and coalitions) on the same explanatory variables as in 1 plus interactions with measures of regional government control. These measures are an indicator for the largest group in the county (columns 1–2) and an indicator for unitary regional government control, which is equal to one if an ethnic group is the largest group in the county, has a coethnic governor, and has a share across the CEC and the CPSB of at least 94 percent (which is the median share for groups with a coethnic governor). All specifications in panel B include county-by-ethnicity and time-fixed effects and the same controls as in 1. Standard errors are two-way clustered at the province-ethnicity and county levels. * $p < 0.1$. ** $p < 0.05$. *** $p < 0.01$.

Appendix: Self-identification

Table: Self-identification: Nation vs. ethnic group

	<i>Dependent variable: National self-identification</i>					
	<i>Continuous</i>			<i>Indicator</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta RF_c \times D_t$	-0.013 (0.029)	-0.013 (0.030)	-0.029 (0.033)	-0.039 (0.071)	-0.040 (0.075)	-0.065 (0.073)
$\Delta EF_e \times D_t$	-0.040 (0.032)	-0.046 (0.036)	-0.044 (0.037)	-0.086 (0.067)	-0.091 (0.074)	-0.077 (0.077)
County-by-ethnicity FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓
Controls	—	✓	✓	—	✓	✓
Controls $\times D_t$	—	—	✓	—	—	✓
Observations	6283	6274	6274	6283	6274	6274

Notes: The table reports the results of regressing two proxies for self-identification (continuous and an indicator for more national) on the same explanatory variables as in the main table. All specifications include county-by-ethnicity and time-fixed effects and control for the respondent's age, indicators if the respondent lives in an urban survey cluster, is female, owns a TV, radio, or motorized vehicle, has completed primary, secondary, or post-secondary education. Standard errors are two-way clustered at the province-ethnicity and county levels. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix: Ethnic discrimination

Table: Ethnic discrimination

	<i>Dependent variable: Ethnic discrimination</i>			
	<i>Continuous</i>		<i>Indicator</i>	
	(1)	(2)	(3)	(4)
<i>Largest group</i>	0.003 (0.007)		-0.003 (0.010)	
<i>Unitary government</i>		0.014 (0.020)		0.006 (0.029)
County FE	✓	✓	✓	✓
Ethnicity-by-time FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Observations	5240	5240	5240	5240

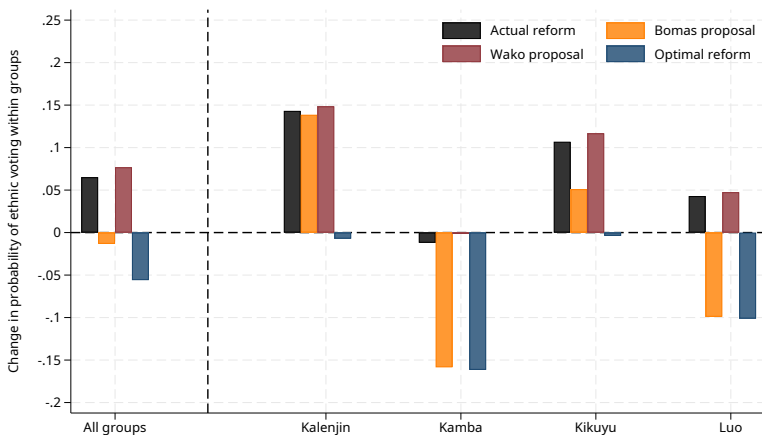
Notes: The table reports the results of regressing two proxies for ethnic discrimination (continuous and an indicator for feeling often or regularly discriminated by the government based on one's ethnicity) on measures of regional control in the post-treatment period (i.e., from Afrobarometer survey round 6 onwards). These measures are an indicator for the largest group in the county (columns 1 and 3) and an indicator for unitary local government control in columns 2 and 4. All specifications include county and ethnicity-by-time-fixed effects and control for the respondent's age, indicators if the respondent lives in an urban survey cluster, is female, owns a TV, radio, or motorized vehicle, has completed primary, secondary, or post-secondary education. Standard errors are two-way clustered at the province-ethnicity and county levels. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix: Counterfactual border designs

Prominent alternative border proposals:

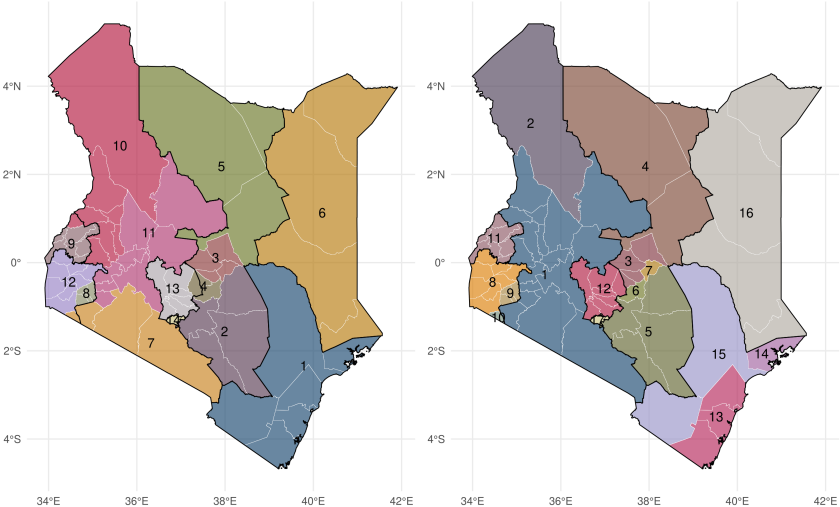
- ▶ Bomas: Broad national constitutional conference, strong devolution, 14 regions.
- ▶ Wako: Counterproposal by the Kibaki government, limited devolution, 70 districts.

Figure: Predicted effects of counterfactual border reforms on ethnic voting



Appendix: Counterfactual border designs

Figure: Comparing the Bomas draft and the optimal design



(a) Bomas draft (14 regions)

(b) Optimal design (17 regions)