Kurdish Studies
November 2018
Volume: 06, No: 01, pp 195-214
ISSN: 2051-4883 (Print) | ISSN 2051-4891 (Online)
www.KurdishStudies.net

**DOI:** 10.67145/ks.v6i1.4160

# Reproducible Markov Transition Estimation for Ethnic Power Positions Applied to Kurdish Groups

**Luis Eduardo Muñoz Guerrero**

Universidad Tecnológica de Pereira, Facultad de Ingenierías, Pereira, Risaralda, Colombia. lemunozg@utp.edu.co. ORCID 0000-0002-9414-6187

**Abstract**

**Ethnic groups hold different, changing positions of state power. This article models transitions among three power states as Markov chains, using Ethnic Power Relations data for 1946–2013. A likelihood-ratio test rejects equal transition probabilities before and after 1990 ($G^2 = 38.50$, $p < 0.001$), and yearly differences stay within ±1.16 points. Exploratory analyses show that Sub-Saharan Africa carries most absolute change, the best cut depends on region and outside state, many groups never move, and state-clustered inference is borderline. Kurdish groups, with few changes, permit description only. The reproducible pipeline traces every number to its source.**

***Keywords:*** Markov chains · cluster bootstrap · ethnic power relations · reproducible pipeline , Kurdish groups.

## 1. Introduction

Ethnic groups differ in their share of state power. Some groups monopolise state power, others share it as junior or senior partners, and others are excluded altogether. These positions are not fixed: groups gain and lose power as states change. Quantitative work links exclusion from state power to armed challenges (Wimmer et al. 2009) and to conflict with governments (Cederman et al. 2010).

This paper asks a narrower question, which the data can answer directly. How stable are the power positions of ethnic groups from one year to the next, and did the probabilities of moving between positions differ between two periods? The Kurdish groups are the substantive focus.

The approach models the positions as a three-state Markov chain, estimates each transition probability by maximum likelihood, and compares the two periods with a likelihood-ratio test of homogeneity of the kind that Anderson and Goodman (1957) develop for first-order chains. Uncertainty is estimated with a cluster bootstrap that resamples whole groups (Efron 1979). The confirmatory hypothesis, the validity checks and the relevance threshold were fixed in an analysis plan before the results were computed, and exploratory analyses, kept apart from the verdict, then ask when, where and for whom the probabilities moved. The paper also treats the analysis as an engineering object: every number in the text is generated from a stored output and traced to its source.

## 2. Background

### 2.1 Measuring ethnic power

The Ethnic Power Relations (EPR) data set identifies the politically relevant ethnic groups of the world and their access to central state power (Cederman et al. 2010), and its 2014 version extends the temporal coverage from 2009 to 2013 (Vogt et al. 2015). Measurement is contested. Chandra and Wilkinson (2008) argue that many tests of ethnic hypotheses use data or measures that conflate ethnic structure and ethnic practice, and the EPR status variable, which records the outcome of political practice at the centre, is closer to the second. Birnir et al. (2017) present a sample of socially recognised groups meant to be free of the selection issues known in this field, a warning that applies to any list of politically relevant groups, since a group enters the data only once coded as relevant.

### 2.2 Exclusion and inclusion

Exclusion is a choice that rulers can revisit. Roessler (2011) finds that in Africa ethnic exclusion substitutes civil war risk for coup risk, and Cederman et al. (2017) offer an account of declining ethnic civil war that stresses group rights, regional autonomy and inclusion in power-sharing. A transition from exclusion to access registers such accommodative moves, which gives the transition probabilities a substantive interpretation. Fearon and Laitin (2003) attribute the prevalence of internal war to a steady accumulation of protracted conflicts since the 1950s and 1960s and not to a sudden change at the end of the Cold War, which poses the question of whether a sudden post-Cold War change is visible in power positions at all. For Africa, popular challenges to neopatrimonial rule have been analysed comparatively (Bratton and van de Walle 1994), and Carbone (2007) describes the last decade of the twentieth century as a resurgence of multipartism.

### 2.3 The Kurdish setting and the scope of the review

The Kurdish case suits a study of power positions because its political setting differs from state to state. Harris (1977) notes that the division of the Kurdish core area among Turkey, Iran and Iraq after the First World War gave Kurdish nationalism major opponents. For Turkey, Updegraff (2012) describes a state policy that evolved from denial and mandatory assimilation

to cultural recognition and then to acknowledgement of a political problem. For Iraq, Rogg and Rimscha (2007) write that after decades of fighting the Kurds achieved far-reaching self-rule, and Soguk (2015) argues that the Kurdistan Regional Government behaves empirically as a sovereign entity while remaining constitutionally part of Iraq. For Syria, Federici (2015) describes a de facto autonomous Kurdish region established during the conflict there, and for Iran, Gresh (2009) reports a rise in ethnonationalist activity among the Kurdish population since 2003. These studies concern autonomy, mobilisation, identity and policy, and few of those dimensions are visible in a yearly code for access to central power (Section 5.7).
The review is bounded by design. It covers works published through December 2017, and later publications were neither reviewed for this article nor used in it. The sources reviewed for the novelty check were read through their abstracts, none of which report estimates of transition probabilities of power positions for Kurdish groups, and one record that lists neither a DOI nor an abstract could not be read.

**3. Data**

The data are the Ethnic Power Relations Core 2014 release (Vogt et al. 2015), whose 3,802 rows record the status of 853 groups over ranges of consecutive years from 1946 to 2013. Nine EPR statuses are collapsed into three states, access, excluded and outside, and a move between statuses inside one state, such as from junior to senior partner, counts as a stay (Table 1). The codebook used here does not define IRRELEVANT and STATE COLLAPSE, so the outside state rests on the labels alone, and a transition into or out of it is partly a transition in the attention of the coders. The three-state chain sees a different part of the movement according to where a group stands: most exits from JUNIOR PARTNER and POWERLESS are also changes of state, whereas fewer than a third of those from DISCRIMINATED are.
A pair is two consecutive years of the same group, with sgt the state of group g in year t. It belongs to the first period when its origin year lies between 1946 and 1989 and to the second when it lies between 1990 and 2012, and π(t) denotes the period of origin year t, so that the period of a pair does not depend on its destination. The data give 42,229 pairs, 24,883 in the first period and 17,346 in the second, and only 526 of them change state. Only

**Table 1. The nine EPR statuses, grouped by the state of the chain**

| Status | Meaning | Pairs | Groups | Exit rate | To other state | Kurdish pairs |
|---|---|---|---|---|---|---|
| Access | | | | | | |
| Monopoly | Sole holder of executive power | 1,919 | 49 | 1.51 % | 6.9 % | 0 |
| Dominant | Dominates; others only tokens | 2,253 | 86 | 2.40 % | 11.1 % | 0 |
| Senior partner | Senior member of power-sharing | 3,425 | 167 | 3.24 % | 23.4 % | 6 |
| Junior partner | Junior member of power-sharing | 6,816 | 265 | 2.58 % | 67.6 % | 15 |
| Excluded | | | | | | |
| Powerless | No national influence, not targeted | 15,282 | 476 | 1.60 % | 66.1 % | 72 |
| Discriminated | Targeted exclusion by the state | 5,743 | 239 | 3.06 % | 30.1 % | 163 |
| Self-exclusion | Territory declared independent | 583 | 31 | 3.95 % | 17.4 % | 34 |
| Outside | | | | | | |
| Irrelevant | Not defined in the codebook | 6,121 | 202 | 2.35 % | 97.2 % | 0 |
| State collapse | Not defined in the codebook | 87 | 17 | 20.69% | 77.8 % | 0 |

Pairs by status at origin. Exit rate: yearly probability of leaving the status; to other state: share of those exits that change state. The meaning paraphrases the codebook of the data.

305 of the 853 groups ever record a change of state, and the 85 most mobile groups, a tenth of the total, hold 51.5%

of all changes, so that the number of informative groups, not of pairs, limits precision.
Figure 1 shows the whole panel. Long unbroken rows dominate, and only 43 distinct sequences of the three states occur (panel b). The set of groups is not fixed: the first year holds 336 groups, and the largest later entries, in 1960 (72 groups) and 1991 (79 groups), coincide with African decolonisation and with the dissolution of the Soviet Union. For that reason the analysis works with transitions and not with stocks: the number of groups in access rises by 243, from 74 to 317, but moves between states add only 47 of that gain, while groups entering the list in access add 232 and those leaving it take away 36.
Each EPR row also reports the share of the state population that the group represents and an indicator of regional autonomy, and each state is assigned to one of five regions by the leading digits of its Gleditsch and Ward code. That rule places Sudan and South Sudan in the Middle East and North Africa and the Caucasus in Europe, so the Kurdish group of Armenia falls in Europe and the other four, those of Iran, Turkey, Iraq and Syria, in the Middle East and North Africa. The five Kurdish groups contribute 290 pairs, 21 from access, 269 from excluded and none from outside, with 2 changes of state in total.

## 4. Method

### 4.1 Estimator and homogeneity test

The yearly state of each group is modelled as a first-order Markov chain (Equation (1)), with transition probabilities that are the same for all groups and constant within a period; Sections 5.3 and 5.4 examine these assumptions. With Nij the number of pairs in a period from origin state i to destination state j and Ni· the number from origin i, the maximum-likelihood estimate of each transition probability, conditional on the first state of each series, is the observed share of its row (Equation (2)).

$$\Pr\left(s_{g,t+1} = j \middle| s_{gt} = i,\ s_{g,t-1}, \ldots\right) = P_{ij}^{(\pi(t))} \qquad (1)$$

$$\widehat{P}_{ij} = \frac{N_{ij}}{N_{i\cdot}} \qquad (2)$$

The confirmatory test (H1) asks whether the transition probabilities are the same in both periods. With $\bar{P}_{ij}$ the pooled estimate from the pairs of both periods, the likelihood-ratio statistic is given in Equation (3), where cells without pairs contribute nothing: $G^2$ is twice the log-likelihood gained by giving each period its own matrix, each cell adding its pairs times the log-ratio of its period and pooled estimates. Under homogeneity, $G^2$ is asymptotically chi-squared, with the degrees of freedom of Equation (4), where ri is the number of periods with pairs from origin i and ci the number of destinations reached from it, which gives 6 for the three states.

$$G^2 = 2 \sum_{p} \sum_{i} \sum_{j} N_{ij}^{(p)} \ln \frac{\widehat{P}_{ij}^{(p)}}{\bar{P}_{ij}} \qquad (3)$$

$$\mathrm{df} = \sum_{i} (r_i - 1)\,(c_i - 1) \qquad (4)$$

### 4.2 Uncertainty and validity checks

Pairs from the same group share that group's history, so the bootstrap resamples whole groups instead of pairs, and its 2,000 replicates give percentile intervals for each transition probability and each period difference. Three checks run before any verdict. V1 simulates 500 data sets under homogeneity, keeping the origin state of each real pair and redrawing its destination from the pooled estimate, and its rejection rate at the nominal level must fall inside the band fixed in the plan, from 0.03 to 0.07. V2 checks that the period boundary follows the origin year and that permuting the destinations leaves the period and origin of every pair unchanged. V3 resamples whole groups under the pooled estimate and recomputes $G^2$ in 2,000 replicates, and the bootstrap test must agree with the chi-squared test on which side of the nominal level the result falls. The chi-squared p-value assumes independent pairs, and neither V1 nor V3 corrects for dependence within groups, because both redraw destinations from one pooled chain.

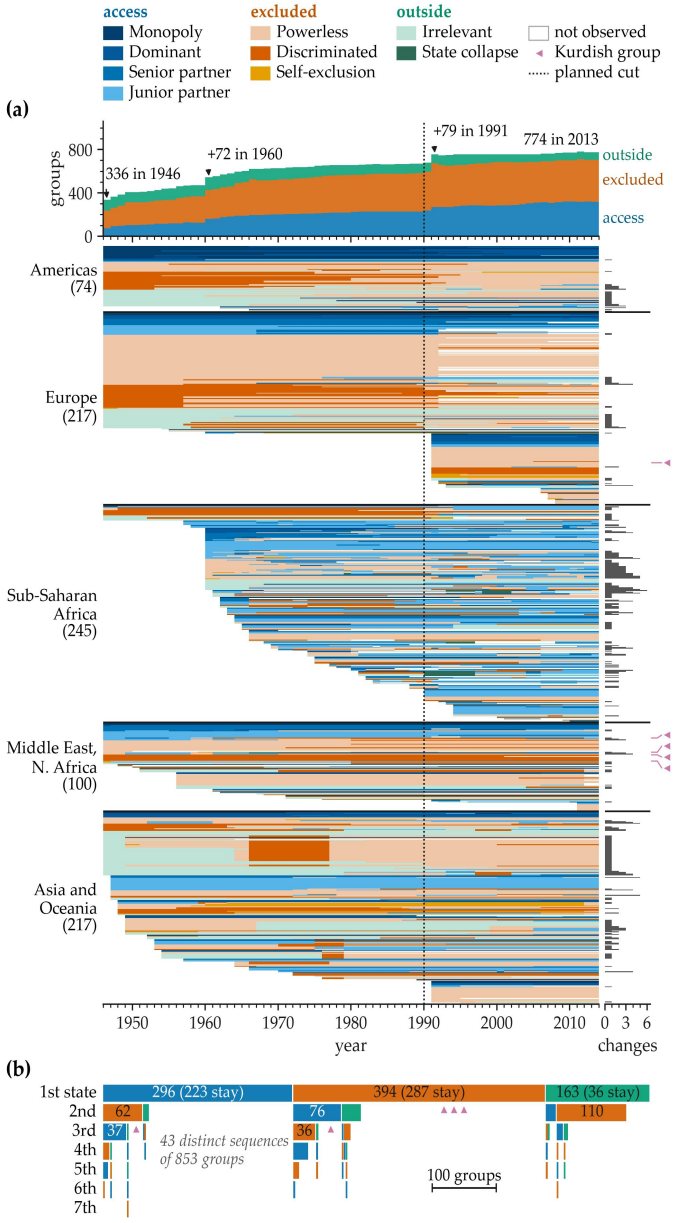


**Figure 1. Every group-year of the data, 1946–2013. (a) Rows are groups, by region, and columns years, coloured by EPR status (white, not observed); above, the groups observed each year by state; at right, the changes of state of each group; the dotted line marks the planned cut. (b) Prefix tree of the 43 sequences of the three states: row d holds the d-th state and block width the groups (in parentheses, those that never change state); blank space marks sequences that end there. Triangles mark Kurdish groups. Source: EPR Core 2014.**

### 4.3 Decision rule and secondary hypotheses

The confirmatory decision uses $\alpha = 0.05$ for H1, with a Holm step-down correction over H1 and H2; the result is positive when H1 survives the correction, and it passes its checks only when V1, V2 and V3 all pass. H2 compares the Kurdish groups with all other groups, with both periods pooled, using the same statistic; its outside row has no Kurdish pairs, so it has 4 effective degrees of freedom rather than the 6 of the plan, and both values are reported. H3 is descriptive: it reports the Kurdish transition probabilities with their 95% cluster-bootstrap intervals. The relevance threshold $\delta$, 5 percentage points in a yearly transition probability, was fixed in advance as a declared judgment of scale.

4.4 Exploratory analyses

The analyses of this section were added after the plan was fixed and the results were known, and none of them enters the verdict. With $\hat{D}$ the six free differences between the periods and $\Sigma$ their covariance over 20,000 cluster-bootstrap replicates, the Wald statistic of Equation (5) allows any dependence within the clusters, which are groups or states; its null distribution is obtained by recentring each replicate difference on $\hat{D}$, and an analytic sandwich covariance gives a second version. A test on triples of consecutive years checks the Markov property (Anderson and Goodman 1957).

**Table 2. Primary test, secondary test and validity checks**

| Test | Statistic (df) | p-value | Outcome |
|---|---|---|---|
| H1, confirmatory: equal periods | $G^2$ = 38.50 (6) | < 0.001 Holm < 0.001 | Rejected |
| H2, secondary: Kurdish groups against others | $G^2$ = 3.695 (4; 6 in the plan) | 0.449 (0.718 with 6) | Not rejected |
| V1, null simulations | Rejection rate 0.058 (0.041 to 0.082) | — | Pass: inside 0.03 to 0.07 |
| V2, period labels | — | — | Pass |
| V3, cluster bootstrap | 0 of 2,000 replicates reach the observed $G^2$ | < 0.0005 | Pass: same side as H1 |

Holm: step-down correction over H1 and H2. V1: rejection rate with its 95% Wilson interval.

$$W = \widehat{D}^{\top}\widehat{\Sigma}^{-1}\widehat{D} \qquad 5$$

In the mover-stayer model of Goodman (1961), some units remain in one category with probability one while the others move according to a Markov chain: a group whose first observed state is i is a stayer with probability φi and otherwise moves according to a common matrix M, so that the likelihood contribution of a group with cg changes and transition counts ng,jl is that of Equation (6). The model is fitted by expectation-maximisation (Fuchs and Greenhouse 1988), and a parametric bootstrap gives the test against the single chain, because the null value of φ lies on the boundary of the parameter space. Over a horizon of h years, the single chain predicts the share of pairs in each destination by powers of the pooled matrix, and the mixture by weighting stayers and movers, with S the diagonal matrix of the shares of stayers among the pairs that start in each state (Equation (7)).

$$L_g = \varphi_i \, \mathbf{1}\{c_g = 0\} + (1-\varphi_i) \prod_{} M_{jl}^{n_{g,jl}} \qquad 6$$

$$\widehat{P}^{(h)}_{\text{chain}} = \bar{P}^{\,h}, \qquad \widehat{P}^{(h)}_{\text{ms}} = S + (I-S)M^h \qquad 7$$

To treat the break date as unknown, W(c) is computed for every candidate cut c between 1955 and 2005, and its supremum (Equation (8)) is compared with a recentred group bootstrap (Andrews 1993). The change in the excluded-to-access probability is divided among regions by the decomposition of Equation (9), in which ξ(p)r is the share of the excluded pairs of period p that lie in region r and a bar denotes the mean of the two periods.

$$W_{\sup} = \max_{c} W(c) \qquad 8$$

$$\widehat{P}^{(2)}_{EA} - \widehat{P}^{(1)}_{EA} = \sum_{r} \bar{\xi}_r \left( \widehat{P}^{(2)}_{r,EA} - \widehat{P}^{(1)}_{r,EA} \right) \qquad 9$$

A discrete-time hazard model (Equation (10)), a logistic regression on the pairs at risk with a covariance clustered by group, relates the move ygt from exclusion into access to the period and to covariates xgt: group size, the logarithm of years in state, regional autonomy, an indicator of states first coded from 1990, the share of other excluded groups and region. Alternative codings repeat the test under five other mappings of the statuses, and coding error is examined by simulation. A Monte Carlo study with the structure of the real panel measures the level of the tests under a first-order chain, a mover-stayer population, a chain whose exit rates follow a gamma distribution across groups (frailty), and a shock common to the groups of a state in a year. For the Kurdish groups, the number of changes K is compared with its exact distribution under the exit probabilities of all other groups, a sum of independent binomials over the nKi Kurdish pairs from each origin i that treats pairs as independent (Equation (11)).

$$\operatorname{logit} \Pr(y_{gt} = 1)$$

$$K \sim \sum \operatorname{Bin}\left(n_i^K,\ 1 - \bar{P}_{ii}^{\text{oth}}\right) \qquad 1$$

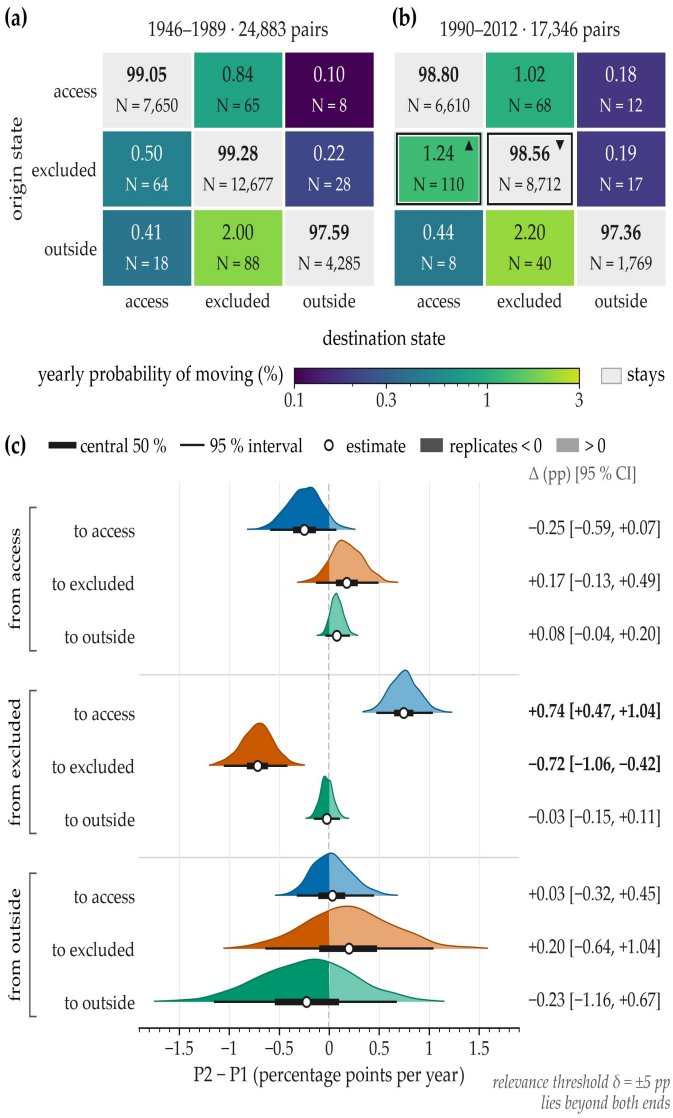


**Figure 2. Transition probabilities and period differences for the primary test (H1). (a, b) Estimated transition matrices for 1946–1989 and 1990–2012 (origin years), origin states in rows: each cell gives the yearly probability in percent and its pairs N; moves are coloured on a logarithmic scale and stays are grey. Outlined cells in (b) differ from (a) with a 95% interval that excludes zero, and triangles give the direction. (c) Differences P2 − P1 in percentage points: ridges are the 2,000 cluster-bootstrap replicates, darker below zero, bars their central half and 95% interval, and circles the estimates; bold values exclude zero.**

## 5. Results

### 5.1 Primary test

Table 2 summarises the tests: the result is positive and passes its checks. In every row and in both periods the most likely destination is the origin state itself (Figure 2, panels a and b), and the homogeneity test rejects equal transition probabilities between the periods, with $G^2 = 38.50$ on 6 degrees of freedom and $p < 0.001$. The Wilson interval of V1 extends above the upper limit of its band because 500 simulations estimate the rate with limited precision, so the check detects gross miscalibration and not a rate a few points too high.

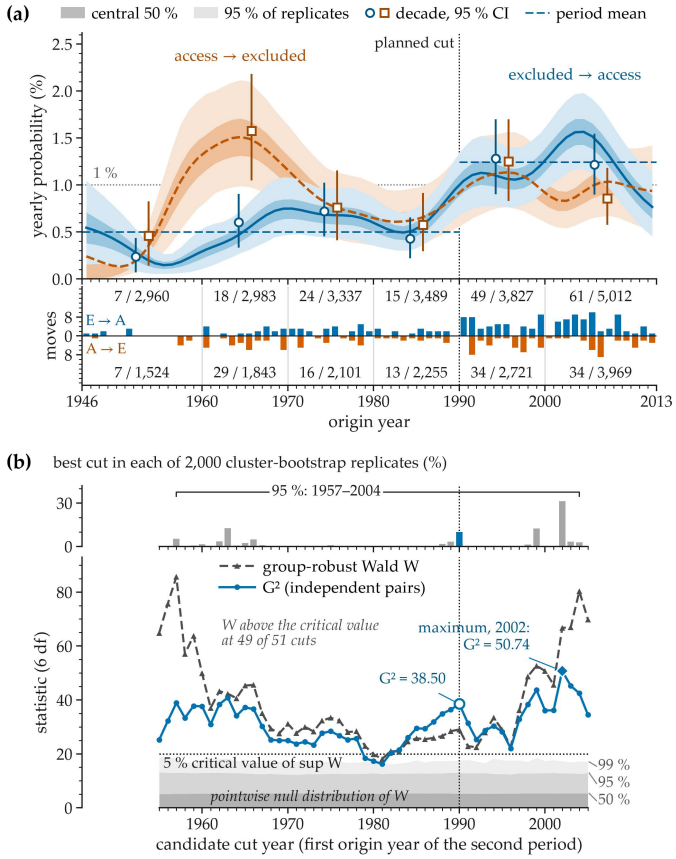


**Figure 3. Timing of the change (exploratory). (a) Yearly probabilities of moving from excluded to access and back, smoothed with a Gaussian kernel of three years, with fans for the central half and 95% of 2,000 cluster-bootstrap replicates, decade estimates with 95% intervals and the period means; bars count the moves of each year, and numbers give the moves over the pairs at risk by decade. (b) G² and the group-robust W by candidate cut, over the pointwise null distribution of W and the 5% critical value of its supremum; above, the share of replicates in which each year maximises G².**

The largest absolute change is from excluded to access, which rises from 0.50% to 1.24%, a difference of 0.74 percentage points (Figure 2, panel c). One cell carries the statistic (Table 3): the move from excluded to access contributes 35.08 of the 38.50 of G², or 91.1%, and it is the only transition between different states whose ratio between the periods has an interval that excludes one. Every 95% interval for a difference between the periods lies within ±1.16 percentage points, far from the relevance threshold of 5 points, so the difference is clear in the statistical sense and small on the yearly scale. It accumulates, however: among pairs whose ten-year window lies before the cut, 3.50% of those that start in exclusion are in access ten years later, and among those whose window lies after it, 8.83%, a difference of 5.33 points with an interval of 3.34 to 7.38.

### 5.2 Dependence and composition

Table 4 gathers exploratory diagnostics of the assumptions that the validity checks leave open. The rejection stands when dependence within groups is allowed and when only the 668 groups present in both periods are kept. Dependence between the groups of a state matters more: 74.7% of the 526 changes occur in a state-year with another change, and with states as clusters the robust test sits at the conventional boundary, below or above it according to the method, while the interval for the change from excluded to access widens to 0.33 to 1.27 points and still excludes zero.

### 5.3 When and where mobility changed

The excluded-to-access probability lies below one percent in each decade before 1990 and above it after (Figure 3, panel a). With the three states and all groups together, however, the scan of candidate cuts shows that 1990 is not a special year (panel b): 49 of the 51 cuts exceed the critical value of the supremum, and the G² curve peaks at 2002. Which year fits best depends on which moves are counted

| Table 3. Ratio of each transition probability between the periods and its part of $G^2$ | | |
|---|---|---|
| **Transition** | **Ratio P2/P1 (95% interval)** | **Part of $G^2$** |
| Access to access | 1.00 (0.99 to 1.00) | 0.02 |
| Access to excluded | 1.21 (0.86 to 1.72) | 1.18 |
| Access to outside | 1.73 (0.71 to 5.04) | 1.48 |
| Excluded to access | 2.48 (1.82 to 3.39) | 35.08 |
| Excluded to excluded | 0.99 (0.99 to 1.00) | 0.27 |
| Excluded to outside | 0.88 (0.42 to 1.62) | 0.18 |
| Outside to access | 1.07 (0.32 to 2.47) | 0.03 |
| Outside to excluded | 1.10 (0.72 to 1.59) | 0.24 |
| Outside to outside | 1.00 (0.99 to 1.01) | 0.01 |
| Percentile intervals from the 2,000 replicates that resample groups. The part of $G^2$ gathers the terms of Equation (3) for that cell over both periods; the column sums to $G^2 = 38.50$. | | |

| Table 4. Supplementary diagnostics (exploratory) | |
|---|---|
| **Diagnostic** | **Result** |
| Group-robust Wald test of H1 | $W = 28.96$ (6 df); recentred bootstrap $p < 0.001$ |
| Wald test, states as clusters | $W = 12.59$; recentred bootstrap $p = 0.058$; analytic sandwich $p = 0.035$ |
| Groups present in both periods | 668 groups; $G^2 = 54.06$, $p < 0.001$ |
| Second-order against first-order | $G^2 = 29.21$ (12 df), $p = 0.004$, triples treated as independent |
| Ten-year prediction of the chain | 5 of 9 cells exclude zero |
| The Wald statistics are recentred over 20,000 replicates. | |

| Table 5. Best-fitting single cut by chain and sample (exploratory) | | | | |
|---|---|---|---|---|
| **Chain** | **Best cut** | **$G^2$ at best cut** | **$G^2$ at 1990** | **Replicates with best cut in 1990** |
| All groups | | | | |
| Three states | 2002 | 50.74 | 38.50 | 10.1% |
| Access and exclusion only | 1990 | 36.56 | 36.56 | 62.7% |
| Sub-Saharan Africa | | | | |
| Three states | 1990 | 31.06 | 31.06 | 25.1% |
| Access and exclusion only | 1990 | 29.31 | 29.31 | 74.8% |
| Rest of the world | | | | |
| Three states | 2002 | 37.01 | 12.03 | 0.0% |
| Access and exclusion only | 2001 | 14.86 | 10.83 | 0.9% |
| Replicates resample groups. Single cuts run from 1955 to 2005. | | | | |

When the chain is limited to access and exclusion, the best cut for all groups together is 1990, chosen by 62.7% of the replicates, and within Sub-Saharan Africa it is also 1990; in the rest of the world the The change is regionally concentrated on the absolute scale (Table 6): by Equation (9), Sub-Saharan Africa accounts for 0.51 of the 0.74 points of the change for all groups, or 68.1%. In Figure 4, the Sub-Saharan probability reaches 5.93% in the 1990s and 4.31% after 2000, against 1.43% in the 1980s, while in the Middle East and North Africa it is zero from 1970 to 1999, so that the region of four of the five Kurdish groups shows no increase. Without Sub-Saharan Africa, the pooled $G^2$ falls to 12.03 on 6 degrees of freedom, with $p = 0.061$, and the group-robust test gives $p = 0.177$. The result does not hinge on a single state, since leaving out each of the

165 states in turn leaves $G^2$ between 28.07 and 45.13, nor on the changing set of groups, since among groups present in both periods the change in Sub-Saharan Africa is 4.14 points. Weighting each pair by the population share of its group gives 2.33 points for all groups and 5.86 in Sub-Saharan Africa, above the relevance threshold in its point estimate.
Group covariates do not explain the contrast. In the hazard model of moving from exclusion to access, the crude period odds ratio of 2.50 becomes 4.00 when the covariates are added, with a group-robust interval from 2.80 to 5.70; larger groups are more likely to enter access (odds ratio 4.45 per tenfold increase in size), and the odds fall as time in exclusion accumulates (0.74 per unit of the logarithm of years). On the odds scale the contrast is broader than on the absolute scale: in a variant of that model with one period coefficient for each region, the odds ratio is above one, with an interval that excludes it, in every region except the Middle East and North Africa. These are associations in observational data, and the covariates may themselves be consequences of earlier political change.

**5.4 Movers and stayers**

The first-order chain fails as a description of the data (Table 4), and the groups differ: 548 groups, or 64.2%, never change state, while the common chain would produce 472.4. The yearly risk of leaving exclusion is not constant either, as a first-order chain would imply: in spells (unbroken stretches in one state) with an observed start it is 1.53% in the first year, 4.88% between years four and six and 1.05% after two decades. The mover-stayer mixture fits far better than a single chain (Table 7), with a likelihood-ratio statistic of 153.4 against a largest value of 7.1 among 199 parametric bootstrap replicates ($p = 0.005$, the smallest p-value this number allows). For transitions between access and exclusion it closes much of the gap at long horizons (Figure 5, panels a, b, d and e): after 40 years, 87.5% of the pairs that start in access are still there, against 69.6% under the chain and 79.2% under the mixture.
The mixture does not reproduce the data either: it predicts 151 groups with one change where 185 occur, and a gamma distribution of the rates across groups predicts nearly the same number of groups without change, 551, so that count does not separate a class of stayers from continuous heterogeneity. The period contrast survives among movers and it is larger: their excluded-to-access probability rises by 1.33 points, with an interval from 0.81 to 1.86.

**5.5 Sensitivity to coding**

The rejection does not depend on the mapping of statuses into states (Table 8): $G^2$ rejects in all six specifications, and the change from any excluded status to any access status varies only between 0.74 and 0.77 points. The finding that homogeneity is not rejected without Sub-Saharan Africa is less robust: it holds in three of the six mappings, with p-values from 0.053 to 0.061, and fails with the four-state mappings and the access-exclusion chain ($p = 0.004$). Random coding errors do not account for the result: every one of the 200 replicates rejects at every error rate studied, up to 2.0% of group-years miscoded. A systematic error can only be bounded: the rise comes largely from 69 African entries into access after 1990, and $G^2$ would fall below its critical value if 36 of them, 52.2%, were artefacts returned to exclusion.
5.6 How the tests behave
The simulation in Table 9 qualifies the group-robust tests. Under a single first-order chain all three tests hold their nominal level, but under the heterogeneity found in the data they do not: with gamma frailty the recentred bootstrap rejects 10.4% of null panels, possibly because the transitions are concentrated in few groups, the setting of few clusters in which standard asymptotic tests over-reject (Cameron et al. 2008). This changes little for the main result, whose bootstrap p-value is far below the level at which an excess of a few points would matter, although that p-value allows only for dependence within groups. A shock common to the groups of a state in a year matters more: at a strength that yields the simultaneity found in the data, the chi-squared test rejects 39.7% of null panels at the nominal 5% and the Wald test with states as clusters 10.7%, so the observed statistic lies in the tail of the null distribution much less firmly than the chi-squared reference suggests.

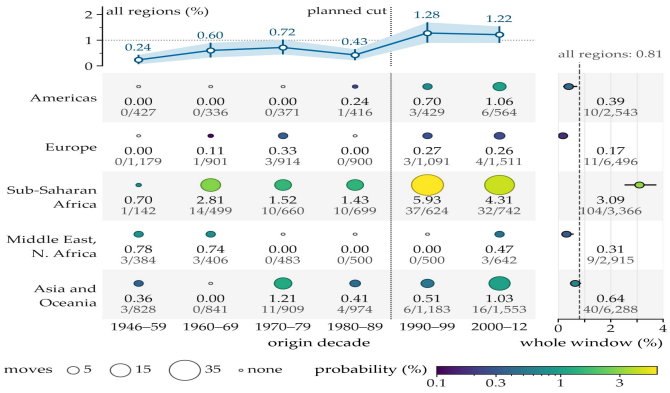


**Figure 4. Yearly probability of moving from excluded to access by region and origin decade (exploratory). Circle area shows the moves and colour the probability on a logarithmic scale; below each circle are the probability and the moves over the pairs at risk, and small open circles have no moves. Above, all regions by decade with 95% cluster-bootstrap intervals; at right, each region over the whole window with exact Poisson 95% intervals that ignore clustering.**

**Table 7. Single chain against the mover-stayer mixture (exploratory)**

| Quantity | Single chain | Mover-stayer mixture |
|---|---|---|
| Log-likelihood | −3,033.0 | −2,956.2 |
| AIC | 6,077.9 | 5,930.5 |
| Groups expected never to change | 472.4 | 552 |
| Excluded to access, P1 | 0.50% | 0.92% (movers) |
| Excluded to access, P2 | 1.24% | 2.25% (movers) |
| Stayers, first observed in access | — | 64.8% (57.6% to 71.8%) |
| Stayers, first observed in exclusion | — | 51.5% (41.7% to 60.0%) |

548 groups never change. P1 and P2 are origin years before and from 1990; for the mixture, the excluded-to-access rows come from the fit with period-specific mover matrices. Stayer shares with 95% intervals.

**5.7 Kurdish groups**

The secondary test (H2) gives no evidence against equality between the Kurdish groups and all others (Table 2), but with so few changes the chi-squared approximation is weak and the test cannot tell equality from lack of data. Only 2 Kurdish pairs change state, one from access to excluded and one from excluded to access, and only 21 start in access, so the intervals of Figure 6, panel a, are wide, and the stay probability in access is 95.2%, from 91.7% to 100.0%. The exact comparison confirms how little the data can say (panels b and c): the rates of all other groups imply 2.97 changes and 2 are observed, a ratio of 0.67 with an exact 95% interval from 0.08 to 2.41, and a Kurdish rate would need to be about 3.45 times that of other groups to be detected with 80% power. Detecting a doubling of the rate would take about 910 pairs, or 182 years for five groups.

Figure 4. Yearly probability of moving from excluded to access by region and origin decade (exploratory). Circle area shows the moves and colour the probability on a logarithmic scale; below each circle are the probability and the moves over the pairs at risk, and small open circles have no moves. Above, all regions by decade with 95% cluster-bootstrap intervals; at right, each region over the whole window with exact Poisson 95% intervals that ignore clustering.

Counting all nine statuses adds information (Table 10). The five groups have 11 changes of status but only 2 of state, 18.2%, against 54.3% for all non-Kurdish groups, so most Kurdish moves lie inside the excluded state, unseen by the chain. Their yearly rate of status change, 3.79%, is indistinguishable from that of the other groups (rate ratio 1.65, exact interval from 0.82 to 2.96). 548 groups never change. P1 and P2 are origin years before and from 1990; for the mixture, the excluded-to-access rows come from the fit with period-specific mover matrices. Stayer shares with 95% intervals.

The Kurdish codes read differently beside those of the other groups of the same states (Figure 7). In Iraq, the move of the Kurdish group into access in 2004 coincides with an inversion at the centre, in which the Sunni Arab group moves from MONOPOLY to JUNIOR PARTNER and the Shi'a Arab group from DISCRIMINATED to SENIOR PARTNER. In Iran, the Kurdish move from POWERLESS to DISCRIMINATED in 1980 is shared in that year with the Arab, Baloch and Turkmen groups. In Syria, the Kurdish, Alawi, Christian and Druze groups all lose access in 1958, and in 1961 the other three regain it while the Kurdish group does not. In Turkey, where the Kurdish group holds 18% of the population (panel b), no group changes status at all. These coincidences belong to the codes and do not establish a common origin.

Table 11 sets the coded series beside what the case literature describes. The coded status registers fully only the self-rule of the Kurds of Iraq, and in part the change of status in Syria in 2012, without any autonomy code; the other episodes fall on a status that does not change or on groups that the data do not code.

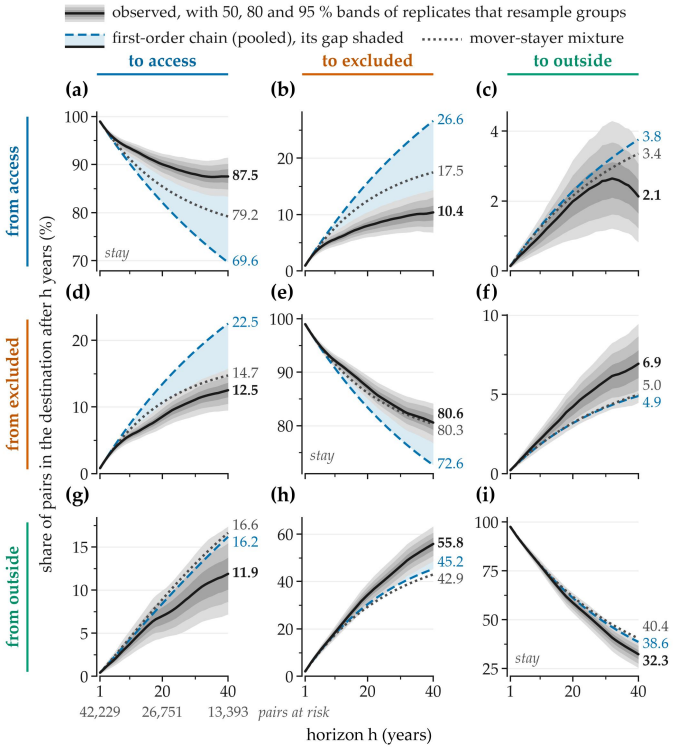


**Figure 5. Transition shares by horizon against two models (exploratory). Rows give the origin and columns the destination state; each panel shows the share of pairs there h years later, with fans of the replicates that resample groups and the gap of the pooled first-order chain shaded. End labels give the values at the last horizon; pairs at risk appear below.**

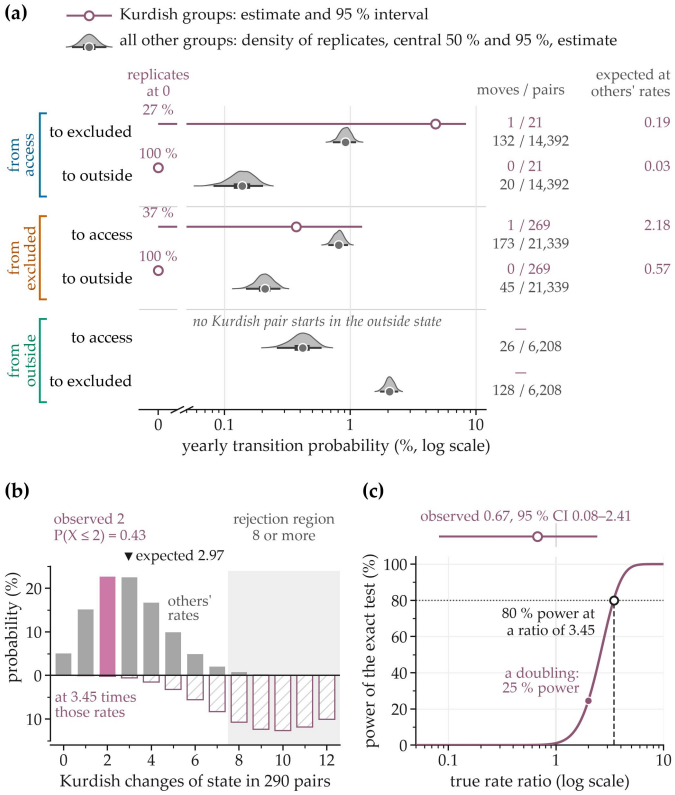


**Figure 6. Kurdish groups against all other groups (H2). (a) Transitions between states on a logarithmic scale with a strip for zero: Kurdish estimates with 95% intervals (purple) and the replicate densities of the other groups (grey); the columns give moves over origin pairs and the Kurdish moves expected at the rates of the other groups. (b) Exact distribution of the Kurdish changes at those rates (top) and at 3.45 times them (bottom, hatched); the rejection region is shaded. (c) Power of the exact test by rate ratio, with the observed ratio and its exact 95% interval.**

**Table 8. Sensitivity to the mapping of statuses into states (exploratory)**

| Mapping | States | Changes | $G^2$ (df) | Excluded to access, points |
|---|---|---|---|---|
| Three states (primary mapping) | 3 | 526 | 38.50 (6) | 0.74 |
| Self-exclusion moved to outside | 3 | 562 | 40.35 (6) | 0.77 |
| State collapse moved to excluded | 3 | 521 | 40.34 (6) | 0.76 |
| Four states: access split by rank | 4 | 625 | 59.41 (12) | 0.74 |
| Four states: discrimination apart | 4 | 723 | 67.08 (12) | 0.74 |
| Two states: outside pairs dropped | 2 | 307 | 36.56 (2) | 0.74 |

Each row applies the mapping to all pairs and repeats the test; the last column joins every excluded status and every access status.

**Table 9. Rejection rates at the nominal level in simulated panels (exploratory)**

| Process with no period change | $G^2$ against chi-squared | Sandwich Wald | Bootstrap Wald |
|---|---|---|---|
| One first-order chain | 5.0% (3.8% to 6.5%) | 5.3% (4.1% to 6.9%) | 5.7% (4.4% to 7.3%) |
| Movers and stayers | 5.9% (4.6% to 7.5%) | 7.2% (5.8% to 9.0%) | 8.1% (6.6% to 10.0%) |
| Gamma frailty in exit rates | 6.7% (5.3% to 8.4%) | 9.1% (7.5% to 11.0%) | 10.4% (8.7% to 12.4%) |

Panels with the structure of the real one; 1,000 simulations per row, Wilson 95% intervals in parentheses. Nominal level 5%; the bootstrap Wald test is recentred and resamples groups.

## 6. Reproducibility and engineering

### 6.1 Pipeline and invariants

Replication needs access to the data behind a study (King 1995), and reproducibility is treated here as an engineering requirement (Peng 2011; Sandve et al. 2013; Wilson et al. 2014). The analysis is a pipeline of five stages (Figure 9), and the manuscript is built last from the stored outputs: a number in the prose is a formatted reference to a value in a results file and not one typed by an author, so that when an analysis is rerun every affected number and table changes together. Random seeds are fixed. The confirmatory analysis has a frozen plan and a hash of its code, and exploratory scripts live in a different directory, so adding one does not alter that hash.

Four arithmetic invariants are checked in every run and stored with the outputs (Equation (12)): the counts add up to the n pairs and each row total matches its counts (I1); each estimated row sums to one or is empty (I2); $G^2$ is non-negative up to a floating-point tolerance ε (I3); and each cluster-bootstrap replicate has as many pairs as the groups it draws, where Rbg counts the draws of group g in replicate b and ng is its number of pairs (I4, which does not cover the replicates of V3). The closed-form estimates agree with those of a numerical constrained optimiser (SLSQP) to the ninth decimal place, and an independent implementation built on

SciPy's contingency-table test reproduces the primary statistics, their degrees of freedom and p-values to the same precision.

$$N_{\cdot\cdot}^{(\cdot)} = n, \qquad \widehat{P}_{i\cdot} \in \{0,\ 1\},$$

$$G^2 \geq - \varepsilon, \qquad N_{\cdot\cdot}^{*(b)} = \sum_{g} R_{bg}\ n_g \qquad (12)$$

**6.2 Vectorised resampling**

Panels with the structure of the real one; 1,000 simulations per row, Wilson 95% intervals in parentheses. Nominal level 5%; the bootstrap Wald test is recentred and resamples groups.

The bootstrap is computed from sufficient statistics. The counts of each group are stored once as a four-way array of group, period, origin and destination, and a replicate is a vector of draws per group, so that the replicated tables of B replicates are the single matrix product of Equation (13), in which A is the array flattened by group, R holds the multiplicities, nP is the number of periods and k the number of states. No individual pair is touched, so the computation cannot break the unit of resampling. The cost does not depend on the n pairs (Equation (14)): 30.7 million elementary operations for 2,000 replicates, against 84.5 million for recounting the pairs. Figure 10 compares their times on one thread.

$$\mathbf{N}^{*} = \mathbf{R}\,\mathbf{A}, \qquad \mathbf{R} \in \mathbb{N}^{B \times n_G}, \quad \mathbf{A} \in \mathbb{N}^{n_G \times n_P k^2} \qquad (13)$$

$$C_{\text{recount}} = B\,n, \qquad C_{\text{product}} = B\,n_G\,n_P\,k^2 \qquad (14)$$

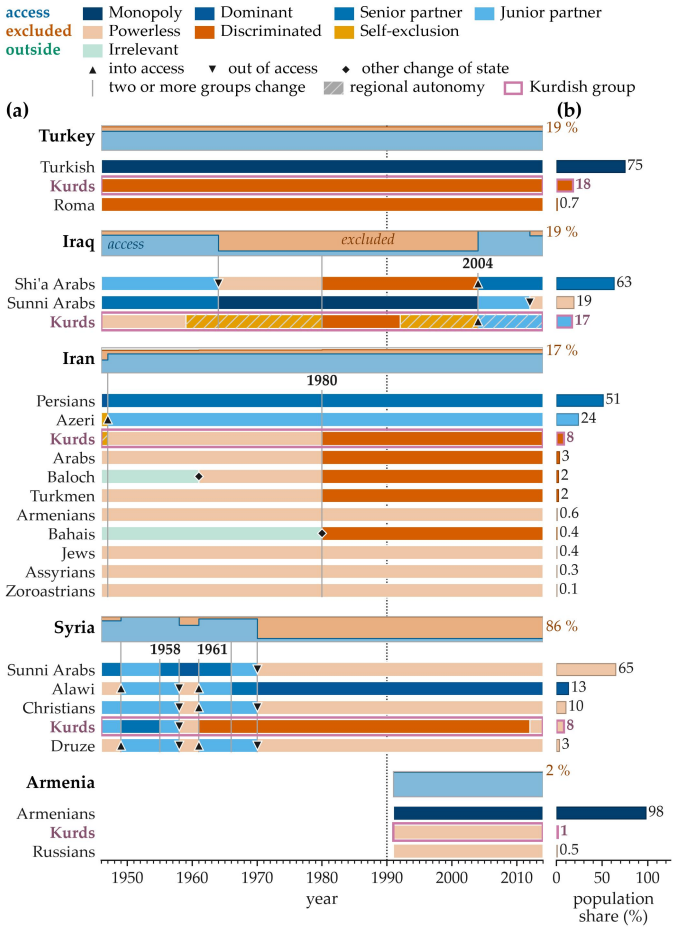


**Figure 7. Ethnic power configurations of the states with Kurdish groups. (a) EPR status of each group by year, beneath a ribbon that splits the population of the state among access, excluded and outside (blank, unassigned; the number gives the last excluded share). Triangles mark changes of state; grey lines, years in which two or more groups change, labelled where the text discusses them; hatching, Kurdish regional autonomy; dotted line, planned cut. (b) Population share of each group, coloured by its last status.**

**6.3 Release checks**

Adjoining rows with the same status are merged; the chain state follows Section 3, and no Kurdish spell falls outside. The codebook codes every self-exclusion as autonomous by rule.

A text number with no source fails a release check, a citation fails without a verified passage from the abstract of the work cited, causal wording is refused outside negated sentences, as are dates after the declared limit of the review. The checks were tested by injecting faults into a copy of the finished document: of 1,320 faults in eleven classes, 96.1% were detected, and detection was complete in every class but one: an altered digit is detected in 57.5% of the cases, because the check asks whether a number is registered and not whether it is the right one. These checks confirm that a number comes from a file, not that the analysis producing that file is right,

**Table 10. Kurdish spells in the EPR Core data**

| Host state | Years | EPR status | Chain state | Years with autonomy |
|---|---|---|---|---|
| Armenia | 1991–2013 | Powerless | Excluded | 0 of 23 |
| Iran | 1946–1946 | Self-exclusion | Excluded | 1 of 1 |
| Iran | 1947–1979 | Powerless | Excluded | 0 of 33 |
| Iran | 1980–2013 | Discriminated | Excluded | 0 of 34 |
| Iraq | 1946–1958 | Powerless | Excluded | 0 of 13 |
| Iraq | 1959–1979 | Self-exclusion | Excluded | 21 of 21 |
| Iraq | 1980–1991 | Discriminated | Excluded | 0 of 12 |
| Iraq | 1992–2003 | Self-exclusion | Excluded | 12 of 12 |
| Iraq | 2004–2013 | Junior partner | Access | 10 of 10 |
| Syria | 1946–1948 | Junior partner | Access | 0 of 3 |
| Syria | 1949–1954 | Senior partner | Access | 0 of 6 |
| Syria | 1955–1957 | Junior partner | Access | 0 of 3 |
| Syria | 1958–1960 | Powerless | Excluded | 0 of 3 |
| Syria | 1961–2011 | Discriminated | Excluded | 0 of 51 |
| Syria | 2012–2013 | Powerless | Excluded | 0 of 2 |
| Turkey | 1946–2013 | Discriminated | Excluded | 0 of 68 |

Adjoining rows with the same status are merged; the chain state follows Section 3, and no Kurdish spell falls outside. The codebook codes every self-exclusion as autonomous by rule.

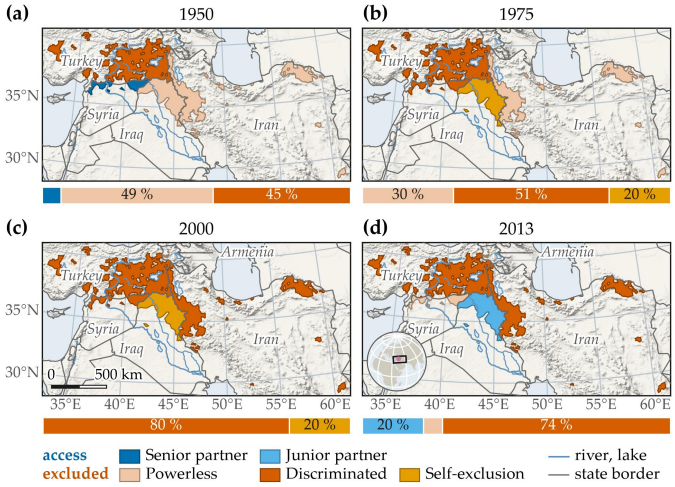


**Figure 8. Kurdish settlement areas coloured by the EPR status of their group in 1950, 1975, 2000 and 2013, in a Lambert azimuthal equal-area projection over shaded relief with rivers, lakes, borders and a graticule; a group not observed in a year is not drawn. The bar under each map splits the coded Kurdish area by status; scale bar in (c), locator globe in (d). Sources: GeoEPR 2014 and Natural Earth.**

**Table 11. Case literature against the coded series (exploratory)**

| Case | What the sources describe | What the code records |
|---|---|---|
| Turkey | Policy from denial to recognition (Updegraff 2012); self-determination read constitutionally (Bayir 2013); mobilisation and the PKK (Tezcür 2010; Mousseau 2012; Jongerden 2017); Europeanisation and the Kurdish initiative (Keyman 2012; Saylan 2012; Casier et al. 2011) | DISCRIMINATED, 1946–2013, no autonomy: one status for the whole window |
| Iraq | Far-reaching self-rule after decades of fighting (Rogg and Rimscha 2007); a regional government that acts as a sovereign entity (Soguk 2015) | SELF-EXCLUSION, 1992–2003 (autonomy 12 of 12 years), then JUNIOR PARTNER, 2004–2013 (10 of 10): recorded |
| Syria | De facto autonomy during the conflict (Federici 2015; Radpey 2016) | POWERLESS, 2012–2013 (autonomy 0 of 2): a change of status only |
| Iran | Ethnonationalist activity rising since 2003 (Gresh 2009) | DISCRIMINATED, 1980–2013, no autonomy: not recorded |
| Soviet Union | Fragmented, inconsistent policy towards its Kurds (Pohl 2017) | No Kurdish group before Armenia in 1991: not coded |
| Descriptions follow the abstracts of the works cited. No event is attributed to any change of code. | | |

and that a quoted passage exists, not that the surrounding sentence represents the source fairly; the primary statistics have an independent replica, but the exploratory analyses do not.

## 7. Discussion and limitations

### 7.1 What the evidence supports

The confirmatory result remains valid under its plan, but read alone it suggests a global shift after the end of the Cold War, whereas the exploratory analyses support a narrower statement: an average in which one region carries most of the weight on the absolute scale, with a planned cut that stands out only for moves between access and exclusion. The yearly differences are small against the threshold fixed in the plan, while over a decade the observed difference in reaching access from exclusion is 5.33 points. Read as laws of motion, the two chains imply different equilibrium distributions and mean sojourn times (Kalbfleisch and Lawless 1985), given by Equation (15) and in Table 12: the long-run share of groups in access rises from 34.2% under the first chain to 48.8% under the second, a difference of 14.6 points with an interval from 6.2 to 23.1, and the mean stay in exclusion falls from 139 to 70 years. These equilibria extrapolate average rates far beyond the window: the estimated matrices are best read as average yearly rates over a heterogeneous population, and not as the law of motion of any single group.

$$\pi^{(p)} = \pi^{(p)}\widehat{P}^{(p)}, \qquad \sum \pi_i^{(p)} = 1, \qquad m_i^{(p)} \qquad \frac{1}{}$$

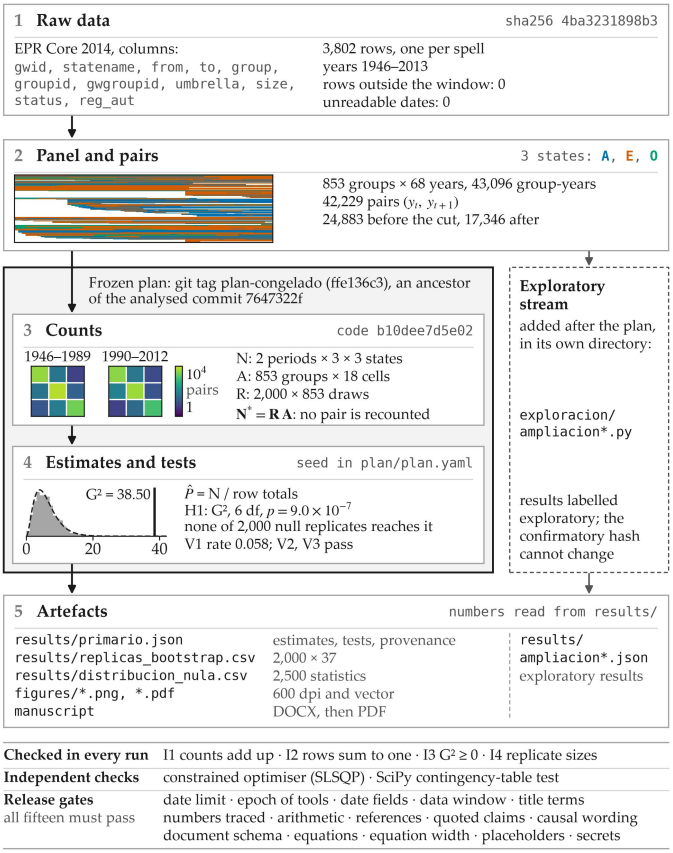


**Figure 9. The pipeline, from the raw data to the manuscript. Each stage shows its data drawn small (the columns of the data table; groups by years coloured by state, white where absent; the count tables of both periods; the bootstrap null distribution of G² against its observed value), its dimensions and its fingerprint. A dark frame marks the frozen plan around the confirmatory stages, and the dashed lane is the separate exploratory stream. Rows below list the invariants, the independent checks and the release gates.**

### 7.2 Competing readings of the African concentration

The regional concentration is compatible with the political opening described for African states in the 1990s (Bratton and van de Walle 1994; Carbone 2007), but this design cannot attribute it to that opening: within the region the best-fitting cut is 1990, with an interval from 1966 to 2003. A political shock may also register as several group-level transitions in states with many relevant groups, but the rate of state-years with at least one entry into access rises by a factor of 2.43 in Sub-Saharan Africa, against 2.89 when pairs are counted, so the multiplication of moves within one political event accounts for only a small part of the contrast. More than half of the entries after the cut, 56.4%, are returns of groups that had held access before, which suits the reopening of earlier coalitions at least as well as the admission of new groups. Finally, the information available to coders about cabinet composition may have become richer after the 1980s, which only a comparison of successive releases of the data could test.

### 7.3 Implications for the Kurdish case

For Kurdish studies, the main contribution of the quantitative analysis is a precise statement of what it cannot deliver: with 2 changes in the three-state chain, no inference about period differences or about differences from other groups is possible, and a longer series would not remedy that. Three descriptive observations remain. There is no evidence that Kurdish groups are less mobile in status than others when all nine categories are counted; the regional setting of most of them has seen no increase in entry into access; and the recorded movements are concentrated inside the category of exclusion, for example between POWERLESS and DISCRIMINATED.

A variable defined as access to central power is poorly suited to the politics that much of the Kurdish literature describes: Iraq after

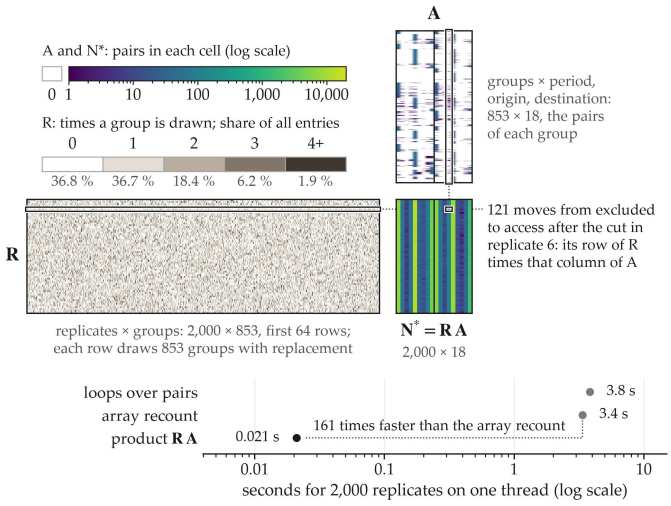


**Figure 10. The cluster bootstrap as one matrix product (Equation (13)), drawn with the real matrices: R at left, A above and N* = R A where they meet. R gives the first 64 rows of the multiplicities (warm greys: times each group is drawn), A the counts of each group by period, origin and destination, and N* the replicate tables, on a logarithmic colour scale with zero in white; outlines mark one entry. Below, the seconds each method of resampling takes on one thread.**

Table 12. Long-run implications of the chain of each period (exploratory)

| State | Long-run share, P1 | Long-run share, P2 | Difference, points (95% CI) | Mean stay, P1 (years) | Mean stay, P2 (years) |
|---|---|---|---|---|---|
| Access | 34.2 % | 48.8 % | 14.6 (6.2 to 23.1) | 1.06 | 1.84 |
| Excluded | 59.0 % | 44.6 % | −14.4 (−22.4 to −6.2) | 1.39 | 1.70 |
| Outside | 6.8% | 6.6% | −0.3 (−4.1 to 3.9) | 1.41 | 1.38 |

Equilibrium shares and mean stays of each period's pooled chain (Equation (15)). P1 and P2 are origin years before and from 1990; percentile intervals from the 2,000 replicates that resample groups.

1991 and Syria after 2012 involve territorial control and self-government, and the Turkish case involves a sequence of policy positions and party-political struggles while the coded status stays the same for the entire window. Quantitative work on the Kurds would gain more from a purpose-built, multidimensional coding of autonomy, recognition, representation and repression than from further modelling of the present series. A low probability of a change in code says nothing about the absence of political conflict, of violence or of change in the lives of the people concerned, and description of this kind can complement the work of area specialists and of Kurdish scholars but does not substitute for it.

### 7.4 Limitations

Equilibrium shares and mean stays of each period's pooled chain (Equation (15)). P1 and P2 are origin years before and from 1990; percentile intervals from the 2,000 replicates that resample groups.

The design is descriptive: it compares two periods and does not identify what moved groups between positions. The EPR labels are judgments made at a distance from the populations they summarise, the outside state rests on labels that the codebook used here does not define, and a miscoded year would enter the pairs as if it were correct. The destination years from 2010 to 2013 were added by the 2014 release, whose coding was not checked against the earlier releases, and in them the probability of moving from excluded to access is 0.52%, on 8 moves, against 1.35% for the destination years 1991 to 2005 and 1.58% for 2006 to 2009, which the analysis cannot separate from an effect of the release. The data hold one status per group and year, so moves within a year are invisible, and whether the yearly chains embed in a continuous-time process is a question especially pertinent with such fragmentary data (Singer and Spilerman 1976). The exploratory analyses were designed after the first results were known, so their p-values lack the meaning they would have in a pre-registered analysis; the regression models omit economic conditions, regime type and conflict history; and the regional classification follows numeric state codes, so that moving Sudan and South Sudan to Sub-Saharan Africa leaves the three-state test without that region at $p = 0.079$.

The panel also conditions on the list of groups. Groups enter and leave it over the window, and an exit ends a series without a transition, so the chain describes movement among listed groups and treats each exit as uninformative censoring: if groups left the list because they lost power, the probabilities of leaving access would be understated. The list is itself selected for political relevance, the selection issue raised in Section 2.1. The two periods also differ in the population at risk, since the second holds fewer pairs, so its intervals are wider and its estimates less stable in small cells, such as the transitions out of the outside state.

## 8. Conclusion

Ethnic power positions are persistent, and the transition probabilities differ between 1946–1989 and 1990–2012. The difference is statistically clear and small on the yearly scale, and over a decade it reaches 5.33 points in the observed share of excluded groups that reach access. Exploratory analyses show that it is carried largely by Sub-Saharan Africa on the absolute scale, that a single first-order chain does not describe all groups, because many never move, and that the robust test sits at the conventional boundary when states are the clusters. For the Kurdish groups, the data contain too few changes to estimate their chain with confidence, and the analysis specifies how far they fall short: detecting a doubling of their rate would take about 182 years of data. The paper contributes a validity protocol and a pipeline in which every number in the text is traced to its source, and its practical lesson is that the confirmatory test alone would have supported a broader claim than the data warrant.

Three extensions follow from the design. A hidden Markov model would separate coding error from movement and replace the simulated miscoding of Section 5.5 with an explicit measurement model; group-level random effects would replace the split between movers and stayers with a continuous distribution of propensities; and a continuous-time formulation would make the yearly matrices consistent across horizons. For the Kurdish groups, each of these first requires codes that register autonomy and representation directly. All three can run inside the same pipeline, so that their numbers stay traceable to their source and pass the same release checks.

### Declarations

**Data availability:** the input data are the Ethnic Power Relations Core 2014 release (Vogt et al. 2015); the map also uses GeoEPR 2014 and the Natural Earth borders. [Code and output availability: to be confirmed by the author.]

Ethics: the study analyses an aggregate, published data set of ethnic groups and involves no human participants and no personal data, and no individual or organisation is named as a research subject.

Funding and competing interests: [to be declared by the author.]

### References

1. Anderson, T. W., and Leo A. Goodman. 1957. "Statistical Inference about Markov Chains." The Annals of Mathematical Statistics 28 (1): 89–110. https://doi.org/10.1214/aoms/1177707039.
2. Andrews, Donald W. K. 1993. "Tests for Parameter Instability and Structural Change With Unknown Change Point." Econometrica 61 (4): 821. https://doi.org/10.2307/2951764.
3. Bayir, Derya. 2013. "Turkey, the Kurds, and the Legal Contours of the Right to Self-Determination." Kurdish Studies 1 (1): 5–27. https://doi.org/10.33182/ks.v1i1.383.
4. Birnir, Jóhanna K., David D. Laitin, Jonathan Wilkenfeld, David M. Waguespack, Agatha S. Hultquist, and Ted R. Gurr. 2017. "Introducing the AMAR (All Minorities at Risk) Data." Journal of Conflict Resolution 62 (1): 203–226. https://doi.org/10.1177/0022002717719974.
5. Bratton, Michael, and Nicolas van de Walle. 1994. "Neopatrimonial Regimes and Political Transitions in Africa." World Politics 46 (4): 453–489. https://doi.org/10.2307/2950715.

6. Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." Review of Economics and Statistics 90 (3): 414–427. https://doi.org/10.1162/rest.90.3.414.
7. Carbone, Giovanni M. 2007. "Political Parties and Party Systems in Africa: Themes and Research Perspectives." World Political Science 3 (3). https://doi.org/10.2202/1935-6226.1023.
8. Casier, Marlies, Joost Jongerden, and Nic Walker. 2011. "Fruitless Attempts? The Kurdish Initiative and Containment of the Kurdish Movement in Turkey." New Perspectives on Turkey 44: 103–127. https://doi.org/10.1017/s0896634600005951.
9. Cederman, Lars-Erik, Andreas Wimmer, and Brian Min. 2010. "Why Do Ethnic Groups Rebel? New Data and Analysis." World Politics 62 (1): 87–119. https://doi.org/10.1017/s0043887109990219.
10. Cederman, Lars-Erik, Kristian Skrede Gleditsch, and Julian Wucherpfennig. 2017. "Predicting the decline of ethnic civil war." Journal of Peace Research 54 (2): 262–274. https://doi.org/10.1177/0022343316684191.
11. Chandra, Kanchan, and Steven Wilkinson. 2008. "Measuring the Effect of "Ethnicity"." Comparative Political Studies 41 (4-5): 515–563. https://doi.org/10.1177/0010414007313240.
12. Efron, B. 1979. "Bootstrap Methods: Another Look at the Jackknife." The Annals of Statistics 7 (1): 1–26. https://doi.org/10.1214/aos/1176344552.
13. Fearon, James D., and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." American Political Science Review 97 (01): 75–90. https://doi.org/10.1017/s0003055403000534.
14. Federici, Vittoria. 2015. "The Rise of Rojava: Kurdish Autonomy in the Syrian Conflict." SAIS Review of International Affairs 35 (2): 81–90. https://doi.org/10.1353/sais.2015.0023.
15. Fuchs, Camil, and Joel B. Greenhouse. 1988. "The EM Algorithm for Maximum Likelihood Estimation in the Mover-Stayer Model." Biometrics 44 (2): 605. https://doi.org/10.2307/2531873.
16. Goodman, Leo A. 1961. "Statistical Methods for the Mover-Stayer Model." Journal of the American Statistical Association 56 (296): 841–868. https://doi.org/10.1080/01621459.1961.10482130.
17. Gresh, Geoffrey. 2009. "Iranian Kurds in an Age of Globalisation." Iran and the Caucasus 13 (1): 187–196. https://doi.org/10.1163/160984909x12476379008241.
18. Harris, George S. 1977. "Ethnic Conflict and the Kurds." The Annals of the American Academy of Political and Social Science 433 (1): 112–124. https://doi.org/10.1177/000271627743300111.
19. Jongerden, Joost. 2017. "A spatial perspective on political group formation in Turkey after the 1971 coup: The Kurdistan Workers Party of Turkey (PKK)." Kurdish Studies 5 (2): 134–156. https://doi.org/10.33182/ks.v5i2.441.
20. Kalbfleisch, J. D., and J. F. Lawless. 1985. "The Analysis of Panel Data under a Markov Assumption." Journal of the American Statistical Association 80 (392): 863–871. https://doi.org/10.1080/01621459.1985.10478195.
21. Keyman, E. Fuat. 2012. "Rethinking the 'Kurdish question' in Turkey." Philosophy & Social Criticism 38 (4-5): 467–476. https://doi.org/10.1177/0191453711435655.
22. King, Gary. 1995. "Replication, Replication." PS: Political Science & Politics 28 (3): 444–452. https://doi.org/10.2307/420301.
23. Mousseau, Demet Yalcin. 2012. "An inquiry into the linkage among nationalizing policies, democratization, and ethno-nationalist conflict: the Kurdish case in Turkey." Nationalities Papers 40 (1): 45–62. https://doi.org/10.1080/00905992.2011.638279.
24. Peng, Roger D. 2011. "Reproducible Research in Computational Science." Science 334 (6060): 1226–1227. https://doi.org/10.1126/science.1213847.
25. Pohl, Jonathan Otto. 2017. "Kurds in the USSR, 1917–1956." Kurdish Studies 5 (2): 157–171. https://doi.org/10.33182/ks.v5i2.442.
26. Radpey, Loqman. 2016. "Kurdish Regional Self-rule Administration in Syria: A new Model of Statehood and its Status in International Law Compared to the Kurdistan Regional Government (KRG) in Iraq." Japanese Journal of Political Science 17 (3): 468–488. https://doi.org/10.1017/s1468109916000190.
27. Roessler, Philip. 2011. "The Enemy Within: Personal Rule, Coups, and Civil War in Africa." World Politics 63 (2): 300–346. https://doi.org/10.1017/s0043887111000049.
28. Rogg, Inga, and Hans Rimscha. 2007. "The Kurds as Parties to and Victims of Conflicts in Iraq." International Review of the Red Cross 89 (868): 823–842. https://doi.org/10.1017/s1816383108000143.
29. Sandve, Geir Kjetil, Anton Nekrutenko, James Taylor, and Eivind Hovig. 2013. "Ten Simple Rules for Reproducible Computational Research." PLoS Computational Biology 9 (10): e1003285. https://doi.org/10.1371/journal.pcbi.1003285.
30. Saylan, İbrahim. 2012. "The Europeanization process and Kurdish nationalism in Turkey: the case of the Democratic Society Party." Nationalities Papers 40 (2): 185–202. https://doi.org/10.1080/00905992.2011.652608.
31. Singer, Burton, and Seymour Spilerman. 1976. "The Representation of Social Processes by Markov Models." American Journal of Sociology 82 (1): 1–54. https://doi.org/10.1086/226269.
32. Soguk, Nevzat. 2015. "With/Out a State, Kurds Rising: The Un/Stated Foreign Policy and the Rise of the Kurdish Regional Government in Iraq." Globalizations 12 (6): 957–968. https://doi.org/10.1080/14747731.2015.1100857.
33. Tezcür, Güneş Murat. 2010. "When democratization radicalizes: The Kurdish nationalist movement in Turkey." Journal of Peace Research 47 (6): 775–789. https://doi.org/10.1177/0022343310386156.
34. Updegraff, Ragan. 2012. "Turkey under the AKP: The Kurdish Question." Journal of Democracy 23 (1): 119–128. https://doi.org/10.1353/jod.2012.0010.

35. Vogt, Manuel, Nils-Christian Bormann, Seraina Rüegger, Lars-Erik Cederman, Philipp Hunziker, and Luc Girardin. 2015. "Integrating Data on Ethnicity, Geography, and Conflict." Journal of Conflict Resolution 59 (7): 1327–1342. https://doi.org/10.1177/0022002715591215.
36. Wilson, Greg, D. A. Aruliah, C. Titus Brown, Neil P. Chue Hong, Matt Davis, Richard T. Guy, Steven H. D. Haddock, et al. 2014. "Best Practices for Scientific Computing." PLoS Biology 12 (1): e1001745. https://doi.org/10.1371/journal.pbio.1001745.
37. Wimmer, Andreas, Lars-Erik Cederman, and Brian Min. 2009. "Ethnic Politics and Armed Conflict: A Configurational Analysis of a New Global Data Set." American Sociological Review 74 (2): 316–337. https://doi.org/10.1177/000312240907400208.
38. Wucherpfennig, Julian, Nils B. Weidmann, Luc Girardin, Lars-Erik Cederman, and Andreas Wimmer. 2011. "Politically Relevant Ethnic Groups across Space and Time: Introducing the GeoEPR Dataset." Conflict Management and Peace Science 28 (5): 423–437. https://doi.org/10.1177/0738894210393217.