Protest from Above: online appendix*

Jonathan Schoots† Johan Fourie‡

Contents

A. Non-parametric tests of wealth and household differences

B. Pre-1779 wealth trend test

C. Network analysis: logit specifications and permutation tests

D. Impact of signatories outside the common wealth sample

E. Heemraad reconstruction and linkage under alternative specifications

*This is the supplementary material for Schoots, Jonathan, and Johan Fourie. 2026. "Protest from Above: Elite Decoupling and Public Petitioning in the VOC Cape Colony." Working Paper, Department of Economics, Stellenbosch University. This paper was created with the help of Anthropic's Claude Code (Opus 5) and OpenAI's Codex (GPT-5.5). Cite the paper, not this appendix.

†Department of Sociology, University of the Witwatersrand. Email: jonathan.schoots@wits.ac.za.

‡Department of Economics, Stellenbosch University. Email: johanf@sun.ac.za.

Appendix A: Non-Parametric Tests of Wealth and Household Differences

Table A1: Kolmogorov-Smirnov tests with Benjamini–Hochberg correction

Asset	D	Raw p	BH-adjusted p
Wine	0.325	1.4×10^{-12}	1.2×10^{-11}
Total index	0.287	$< 10^{-9}$	$< 10^{-8}$
Slaves	0.265	$< 10^{-7}$	$< 10^{-6}$
Horses	0.251	$< 10^{-6}$	$< 10^{-5}$
Settlers	0.174	$< 10^{-3}$	$< 10^{-2}$
Cattle	0.145	.007	.010
Wheat	0.124	.033	.038
Sheep	0.065	.637	.637

Note: R's default `ks.test()` was used; the function emits the warning "p-value will be approximate in the presence of ties" on all of these discrete series. Wheat moves from raw-marginal (0.033) to slightly-more-marginal (0.038) under BH. Only sheep remains non-significant.

Table A2: Mann-Whitney U tests with Benjamini–Hochberg correction

Asset	Pet. median	Non-pet. median	U	BH-adjusted p	r_b
Wine	0	0	99,971	$< .001$	0.33
Total wealth	394	147	99,249	$< .001$	0.32
Slaves	4	0	98,423	$< .001$	0.30
Horses	5	1	94,450	$< .001$	0.25
Settlers	4	2	90,351	$< .001$	0.20
Cattle	20	20	87,218	.002	0.16
Wheat	0	0	84,896	$< .001$	0.13
Sheep	0	100	71,460	.262	-0.05

Note: Mann–Whitney U tests compare petitioners and non-petitioners. Reported p-values are Benjamini–Hochberg-adjusted. r_b is the rank-biserial correlation; positive values indicate that petitioners tend to rank above non-petitioners, while negative values indicate the reverse. All variables except sheep are significant.

Appendix B: Pre-1779 wealth trend test

Table 1 reports the within-household pre-trend model discussed in §6.3 of the main text. The sample is restricted to pre-petition observations (1764–1778) for households in the common 1779 wealth sample. We demean both wealth and year within household, so the interaction term captures whether future petitioners were on a different pre-1779 trajectory from other observed households. Standard errors are clustered at the household level.

Table 1: Within-household pre-1779 wealth trend test.

Term	Estimate	SE	p
Year (demeaned within household)	5.18	1.46	$<.001$***
Future petitioner × year	−1.59	7.27	0.827

Note: Outcome is total production-asset value in rixdollars. Household fixed effects are implemented through within transformation. The positive year term reflects overall growth in observed asset value in the pre-petition panel; the null interaction shows that future petitioners were not a stratum in relative decline. Two simpler tests agree. Defining growth as the change in a household's production-asset value between its first and last observed pre-1779 year, a t-test on mean growth yields $p = .43$ and a Mann–Whitney test on the same quantity yields $p = .52$. The minimum-detectable-slope calculation at 80 per cent power (Roth, 2022) gives $|\Delta\beta| = 20.4$ Rds/year. The year-by-year counterpart is plotted in Figure 1 and tabulated in Table 2.

Table 2 reports and Figure 1 plots the year-by-year counterpart: an event-study regression of household wealth on future-petitioner status interacted with a full set of year dummies, with 1778 as the reference year and the same household fixed effects as Table 1. Standard errors here are the CR2 cluster-robust estimator of `estimatr::lm_robust`; Table 1 uses the cluster-robust sandwich estimator of `sandwich::vcovCL`, so the two are clustered on the household but are not the same estimator. The year-specific point estimates fluctuate around zero through the late 1760s, with wide confidence intervals straddling zero. Five coefficients (1770, 1771, 1772, 1774, 1777) reach individual significance, but their sign pattern runs in the wrong direction for a deprivation reading: because 1778 is the reference year, a negative coefficient in year t means that petitioners' within-household wealth premium over non-petitioners was *smaller* in t than in 1778 – petitioners' relative wealth position was improving, not eroding, over the pre-petition decade – the opposite of a deprivation pattern. The single-interaction slope test in Table 1 aggregates these fluctuations to a non-significant −1.59 Rds/year.

Table 2: Event-study specification of the pre-1779 panel.

Future petitioner × year	Estimate	SE	95% CI	p
1764	34.8	127.7	$(-223.6, 293.1)$	.787
1765	−35.8	112.1	$(-261.6, 189.9)$	.751
1766	−29.6	118.1	$(-265.9, 206.6)$	.803
1767	7.9	95.4	$(-183.6, 199.4)$	.934
1768	−97.4	77.2	$(-251.7, 56.8)$	.212
1769	28.1	81.5	$(-134.6, 190.9)$	.731
1770	−173.6	81.4	$(-336.1, -11.2)$	**.037**
1771	−133.3	59.8	$(-252.2, -14.4)$	**.028**
1772	−210.6	72.8	$(-355.2, -66.0)$	**.005**
1773	−45.4	74.7	$(-193.8, 103.0)$	.545
1774	−163.7	63.4	$(-289.4, -37.9)$	**.011**
1775	−31.1	52.3	$(-134.8, 72.6)$	.553
1776	−78.4	45.1	$(-167.7, 10.8)$	.084
1777	−90.8	45.3	$(-180.5, -1.0)$	**.047**

Note: 5,558 household-year observations on 786 households (106 future petitioners, 680 non-petitioners), 1764–1778. Household fixed effects implemented through within transformation and year fixed effects included via dummies; SEs cluster-robust (CR2) at household level (`estimatr::lm_robust`).

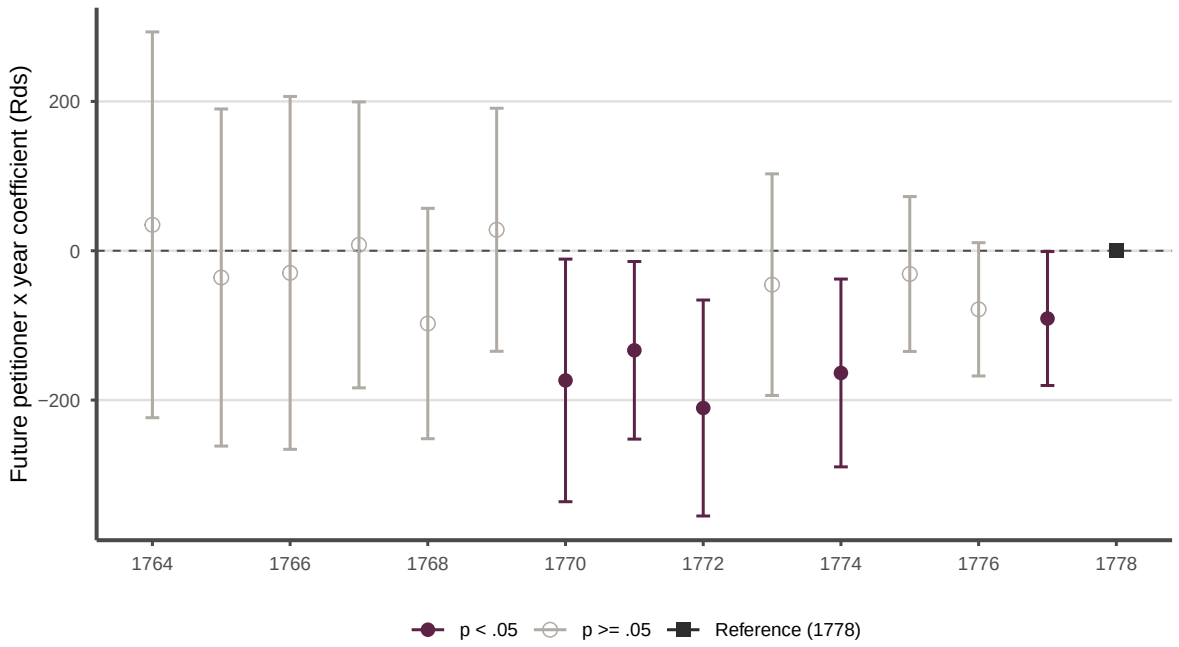

Figure 1: Event-study coefficients from the within-household pre-1779 wealth regression, 1764–1778. Points are $\hat{\beta}_t$ on Future petitioner $\times$ Year t, with 1778 as the reference year and household fixed effects absorbed through within transformation. Whiskers are 95% confidence intervals based on CR2 household-clustered standard errors. Plum points mark years at which the coefficient is individually significant at $p < .05$.

Appendix C: Network analysis

C1: Network logit

Table 3: Network logit — MLE, Firth penalisation, and residualised betweenness

	MLE				*Firth penalisation*	
	Coef.	SE	p	AME (pp)	AME (pp)	VIF
Log betweenness	0.101	0.041	.014	0.71	0.72	3.60
Log degree	0.088	0.299	.768	0.62	0.68	3.60
Residualised specification						
Betweenness $\perp$ degree (residual)	0.101	0.041	.014	0.71	—	—
Log degree (in residualised spec.)	0.790	0.161	$< .001$	5.56	—	—

Note: Rare-events regime: 172 signatories of 2,225 nodes (7.7%). Firth-penalised likelihood (Firth, 1993) yields AMEs within 0.07 pp of the MLE. The residualised specification regresses log betweenness on log degree by OLS and enters the residual alongside log degree in a logit. In the residualised specification, both degree and orthogonal-betweenness are significant predictors of signing, though log degree in this specification reflects its total association with signing, not a unique contribution net of betweenness.

C2: Network permutation tests

Table 4 reports two permutation tests on the realised auction network. The first reallocates petition labels uniformly at random across nodes (B = 2,000); the second preserves the bipartite degree sequence via a configuration-model / Stochastic Degree Sequence Model sampler (Neal, 2014) implemented as a double-edge-swap Markov chain (B = 500). The second is the stronger test: it examines if the observed petitioner–non-petitioner brokerage gap is more extreme than what would arise from degree heterogeneity alone.

Both tests reject the null, but against different benchmarks. The observed gap in mean log betweenness is 1.443. Under the uniform-label null the expected gap is essentially zero (0.002); the observed 1.443 is overwhelmingly larger than chance. Under the degree-preserving SDSM null the expected gap rises to 1.370: roughly 95 per cent of the observed gap is mechanically attributable to petitioners' higher degree. The component the SDSM test isolates is the residual 0.073 – the part of the observed gap that cannot be explained by degree heterogeneity alone – and this residual is still outside the SDSM null distribution, yielding p = .006. An analogous decomposition holds for the top-broker share: of the observed gap of 0.155, about 0.123 is attributable to degree, with a residual 0.032 surviving the SDSM null at p = .004. The SDSM analysis refines the logistic regression analysis of the main text: it highlights that most of the brokerage advantage is mechanically coupled to higher auction attendance. Yet it still identifies that bridging position net of attendance volume, althought a smaller component, survives this stricter null.

Table 4: Permutation tests on the realised auction network.

Metric	Observed gap	Perm. mean	Perm. SD	p
Panel A. Uniform-label permutation ($B = 2{,}000$)				
Mean log betweenness gap	1.443	0.002	0.284	$\leq 0.0005^{***}$
Top-broker share gap	0.155	−0.0002	0.032	$\leq 0.0005^{***}$
Panel B. Configuration-model (SDSM-equivalent) permutation ($B = 500$)				
Mean log betweenness gap	1.443	1.370	0.032	0.006^{**}
Top-broker share gap	0.155	0.123	0.012	0.004^{**}

Note: "Observed gap" is the difference in the statistic between petitioners and non-petitioners on the realised network, computed by a single procedure shared across both panels (betweenness from the bipartite one-mode projection, log-transformed as $\log(b + 1)$; top-broker share equal to membership in the top quintile). Panel A permutes petition labels uniformly at random while holding the edge set fixed; Panel B additionally preserves the bipartite degree sequence via a double-edge-swap configuration-model sampler. p-values use the North–Curtis–Sham $(B + 1)^{-1}$ convention.

Appendix D: Impact of signatories outside the common wealth sample

Of the 404 signatories, 153 enter the common 1779 wealth sample. The other 251 fall outside it for two distinct reasons: 239 were never matched to the censuses, and 12 were matched to households recording no positive quantity on any of the five priced asset series. The first group is genuinely unobserved; the second is observed at zero. Both tables below treat all 251 alike, which is conservative for the first group and generous to the second, and neither is a bound in the formal sense: they are imputation scenarios, and we label them as such.

Table D1: Imputation scenarios for the 251 signatories outside the sample

Assumption on the 251 outside the sample	Pet. mean	Non-pet. mean	Gap
Same as observed petitioners	516	282	+235
Same as non-petitioners	371	282	+89
All hold nothing	196	282	−86

Note: Common 1779 wealth sample (N = 1,139; 153 petitioners, 986 non-petitioners), the same basis as Tables 2 and 3 of the main text (and, less the four collision rows, Table 6) and as Table D2 below. Each row imputes a wealth level to the 251 signatories outside the sample and reports the resulting combined petitioner mean. The gap survives the first two scenarios and not the third. The third is the extreme case rather than a likely one, but we report it because the exercise is otherwise silent on where the premium fails. No inferential test accompanies these means.

Table D2: Break-even grid for Cape District petitioner wealth

Cape share	Assumed Cape n	Break-even mean (Rds)	Share of non-pet. mean
0.1	25.1	−1,148	−407%
0.2	50.2	−433	−153%
0.3	75.3	−195	−69%
0.4	100.4	−75	−27%
0.5	125.5	−4	−1%
0.6	150.6	44	16%
0.7	175.7	78	28%
0.8	200.8	103	37%
0.9	225.9	123	44%
1.0	251.0	139	49%

Note: Common 1779 wealth sample: 153 observed petitioners with a mean of 516 Rds, 251 signatories outside the sample, non-petitioner mean 282 Rds. Each row assumes a given share of those 251 came from the Cape District and that the remainder hold the non-petitioner mean; the break-even column reports the mean the assumed Cape signatories would need for the combined petitioner mean to fall to the non-petitioner mean. Negative entries indicate that no attainable Cape mean can close the gap. An earlier version of this grid inadvertently assigned the non-Cape remainder the observed petitioner mean rather than the non-petitioner mean, which moved the sign change from a Cape share of 0.5 to 0.7.

If the signatories outside the sample who were not from the Cape District held wealth at the non-petitioner mean, what would the Cape share have to hold to break even with the non-petitioner group? For Cape-share assumptions of 0.5 or below the break-even mean is negative, meaning no

attainable Cape wealth could neutralise the rural premium. Above that, closing the gap requires a Cape mean well below the non-petitioner average: 44 Rds at a 0.6 share, rising to 139 Rds when every signatory outside the sample is assumed to be from the Cape District. The structure of our linkage suggests that around half of the signatories outside the sample may be from the Cape District, and at that share no positive Cape mean closes the gap. The premium is therefore secure against moderate assumptions about the Cape component, though not, as Table D1 shows, against the extreme case in which those signatories held nothing at all.

Appendix E: Heemraad reconstruction and linkage under alternative specifications

The two heemraad tables in the main text use different designs. Table 5 starts from the officeholders and asks what share signed. The all-district extract contains six names in each annual cohort but omits district labels; comparison with the district-labelled source workbook shows that the first four are Stellenbosch–Drakenstein appointments and the final two are Swellendam appointments. Table 5 retains the first four. December appointments begin service the following January and last two years, so the 1777–1778 appointment cohorts give the eight men serving in 1779. Hand-adjudication consolidates spelling variants but retains common-name officeholders lacking a patronymic or *d'oude* marker as distinct. This yields 4 of 8 current officeholders, 7 of 17 men serving during 1775–1779, and 11 of 39 serving during 1763–1779. All twelve signers titled current or former heemraad can be linked to the longer Stellenbosch–Drakenstein roll; François du Toit de Oude, the one outside the seventeen-year window, appears in the 1722 cohort. The replication package supplies the 68 appointment-level records and the complete identity crosswalk.

Table 6 instead starts from matched 1779 households and asks whether petitioners were more likely than non-petitioners to have prior heemraad service. Its indicator is necessarily asymmetric. For petitioners we use the Rank field of the 1779 signatory roster, which records each signatory's civic office at the time of signing; this is direct primary-source attribution. For non-petitioners no such field exists, so service must be inferred by matching normalised names against the heemraden roll. Table 5 reports the differential under the specifications that this asymmetry makes available.

Table 5: Prior heemraad service: petitioner–non-petitioner differential under alternative linkage rules.

Specification	N	Pet. %	Non-pet. %	Diff. (pp)	p
Roster title, continuity match (Table 6)	1,135	7.2	2.4	+4.7	.004
Continuity match on both sides	1,135	6.5	2.4	+4.1	.017
Random draw within collision groups	1,139	5.3	2.8	+2.4	—
Name match, collisions ignored	1,139	14.4	5.2	+9.2	<.001
Collision cases dropped	1,081	0.8	1.5	−0.7	1.00
One-office-per-name stress test					
every collision resolved against	1,139	2.6	3.2	−0.6	1.00
every collision resolved in favour	1,139	8.5	2.3	+6.2	<.001

Note: Common 1779 wealth sample. Rows differ only in how a normalised heemraad name matching several 1779 household heads is resolved. "Continuity" assigns the office to the candidate whose panel tenure best fits the recorded service years; "collisions ignored" flags every candidate, which double-counts populous lineages such as de Villiers and du Toit; "collisions dropped" removes them from both groups. Fisher exact two-sided p; the random-draw row reports means over Monte Carlo replications and has no single p (its difference column is the mean of per-replication differences, which need not equal the difference of the two rounded shares). The final two rows are a stress test, not bounds on the reported estimate: they assume exactly one real officeholder per normalised name, an assumption the roster itself contradicts – two distinct Jan de Villiers signatories are recorded as heemraad and oud heemraad – and they discard the petitioner-side roster titles that the reported specification uses. A lineage-level reading yields +22.6pp but changes the unit of analysis from household to lineage and is not comparable with these rows.

Three points follow. First, the specification reported in the main text is not the most favourable available: resolving every ambiguous case in favour gives +6.2pp, and the naive match that ignores namesake collisions gives +9.2pp. Second, the differential is not robust to every rule. Dropping the ambiguous cases, or resolving them all against, produces small negative differences indistinguishable from zero. Both discard or misassign most of the signal: the conservative rule removes 58 households and, between exclusion and reclassification, retains only one of the eleven petitioners the roster identifies as heemraden. They are not thereby illegitimate, and we report them for that reason.

Third, the continuity rule is weaker than a summary description suggests. Of the twenty-one collision groups, eight are settled by an actual overlap between the recorded service years and a candidate's panel tenure. Eleven have no overlap at all and fall back on assigning the office to the longest-tenured candidate, a heuristic whose error direction we cannot sign. Two are unresolvable and are dropped. We prefer this rule because it uses information the alternatives throw away, but the officeholding differential rests on a linkage judgement in a way the wealth and network results do not.

The petitioner side allows one direct check. Against the twelve signatories the roster identifies as

current or former heemraden, the simple name match flags twenty-one, of which seven are correct: precision 0.33, recall 0.58. It is tempting to read the over-inclusion as conservative, on the grounds that false positives inflate the non-petitioner rate and so shrink the gap. We do not make that claim. The check validates the simple match rather than the continuity-corrected one used in Table 6; recall of 0.58 implies substantial false negatives pushing in the opposite direction; and the two Jan de Villiers show that a single normalised name can conceal two genuine officeholders, which the continuity rule will undercount. The net direction of the error is not identified.

References

Firth, David. 1993. "Bias Reduction of Maximum Likelihood Estimates." *Biometrika* 80 (1): 27–38.

Neal, Zachary P. 2014. "The Backbone of Bipartite Projections: Inferring Relationships from Co-Authorship, Co-Sponsorship, Co-Attendance and Other Co-Behaviors." *Social Networks* 39: 84–97.

Roth, Jonathan. 2022. "Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends." *American Economic Review: Insights* 4 (3): 305–322.