Online supplementary material for Democracy and Clustered Models of Global Economic Engagement

ByungKoo Kim and Iain Osgood

This draft: June 10, 2022

Contents

Appendix A: Details on Clustering **1**

Data collection, selection, and imputation 1

Clustering method 3

Appendix B: Additional Models **7**

Simple regression models 7

Models investigating cluster reassignment 8

Models fit on non-imputed data 10

List of Figures

A1 Missing values in variables by country and year. On the left is after thresholding for missings, on the right is the pre-threshold dataset. 2
A2 Log-likelihood and information criteria across clustering models 5
B1 Simple Model: Predicted proportion of clusters as function of democracy at 25%, 50% and 75% percentile 8
B2 Full model: Predicted cluster proportions as a function of democracy (unimputed data) 11

List of Tables

B1 Results from multinomial regression model with all controls 7
B2 Results examining original 7 clusters 9
B3 Results from multinomial regression model with all controls (unimputed data) 10
B4 Results examining original 7 clusters (unimputed data) 11

Appendix A: Details on Clustering

Data collection, selection, and imputation

In this section, we describe in greater detail than the main text the principles and techniques in compiling and organizing the dataset of government policies on globalization for our empirical analysis. We began by attempting to collect measures and proxies for the policy outcomes identified in the 9 categories described above across as many countries and years as possibles. We looked across a wide variety of single- and multi-use datasets. Casting as wide a net as possible, we ultimately attempted to collect data on 135 policy variables across 194 countries from 1940 to 2016.

While this initial data is large in size, it contains vast numbers of missing values which make it unusable as is. The problem of missingness is presented visually in the left hand size of Figure A1. Each cell in the figures represents a single country year (where countries fall along rows and years in columns). If the country-year's data is entirely missing across the 135 initial measures, then the cell is colored black. If the country-year's data is fully observed across those measures, it is colored very light gray. Varying levels of missingness are reflected in gradations between these poles. As can be seen on the left-hand side, missingness in the observations is overwhelming for nearly all countries before 1970, and still severe for many, particularly developing, countries up to 1990.

In response, we resorted to thresholding of years and measures followed by missing data imputation. We also did some iteration between these two steps, because in some cases imputation models did not converge or did not generate satisfactory imputations upon inspection. We also made some ad hoc choices about thresholding, retaining particular variables to preserve coverage across policy areas.

Our first major decision on thresholding was to focus on the years 1990-2016. Extending the analysis into the Cold War would only have been feasible on a much smaller set of variables and would have required pruning entire policy domains. Future scholars may wish to do so but we have opted for preserving country features across all policy domains. We also refined the variables down to 75 target variables that we felt were most clearly within the scope of our project while preserving coverage across policy areas.

We then resorted to automated thresholding of variables. First, we used a threshold of the 50th percentile of missingness across the countries, and dropped all countries below this cutoff. This translates into dropping countries with 75.1% of the variable-year observations measured. We also made an *ad hoc* decision to preserve in the analysis the following countries that otherwise would have been pruned: Cambodia, DRC, Republic of Congo, Croatia, Estonia, Hong Kong, Iran, Latvia, Lebanon, Lithuania, Saudi Arabia, and Zimbabwe.

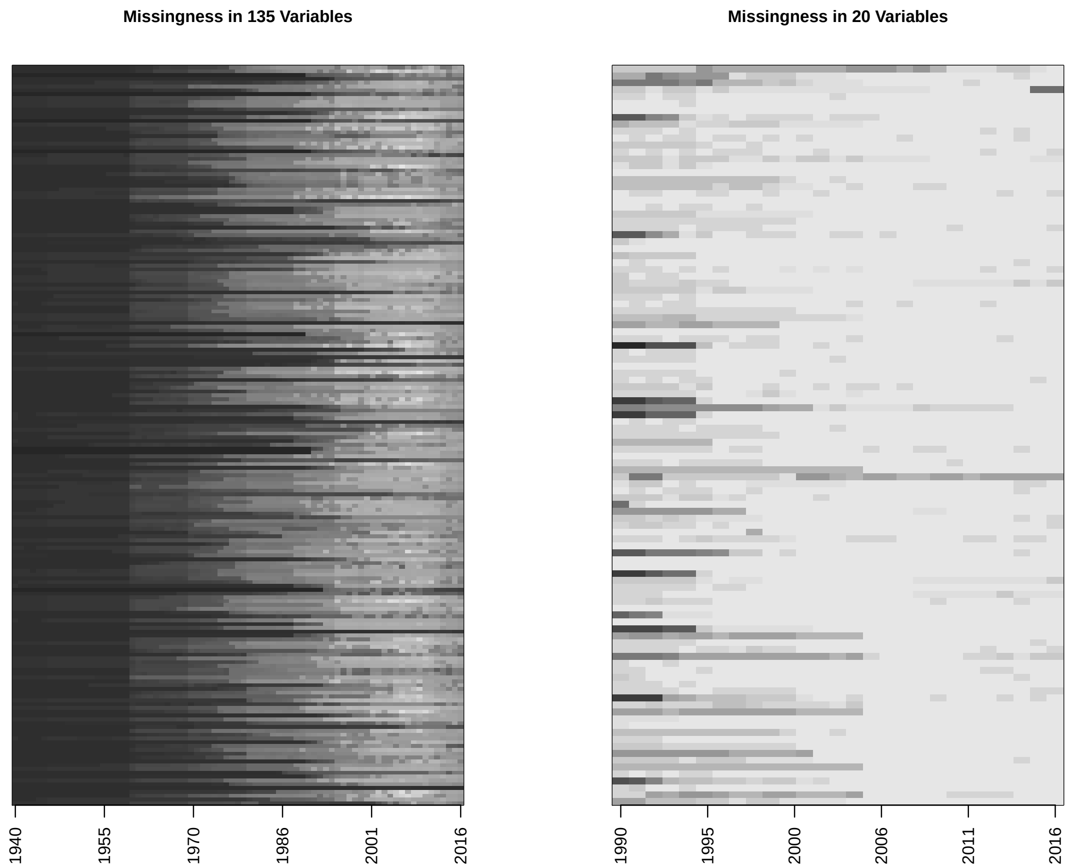

Figure A1: Missing values in variables by country and year. On the left is after thresholding for missings, on the right is the pre-threshold dataset.

We also were forced to eliminate two countries, Belize and Kyrgyzstan, that were completely missing one variable. We felt that imputations would be unreliable in these cases. This results in 107 countries remaining in our data.

We then employed a hard threshold for eliminating variables which fell below the 60th percentile of missingness in the data. This translates to dropping variables with less than 91.5% of the data observed. This left us with 23 remaining variables. Four of these (exports and imports as a share of GDP growth and exports to high-income countries and exports to low-income countries) we felt were not well-suited to our question and were dropped. Dummies for membership in the IMF and GATT/WTO hardly varied across the dataset and were also dropped (as this impeded convergence of the clustering). A final variable (trade as a % of GDP) was dropped because of multicollinearity with imports as a % of GDP and exports as a % of GDP. Finally, we added in two pairs of variables that we felt were substantively important enough to justify the extra missingness: total goods exports and imports; and average MFN tariffs on both primary products and manufacturers. We thus ended up with 20 total variables covering all of our nine topic areas.[1] The 20 variables that we use in our main analysis are presented along with their overarching policy categories in Table 1. The missingness in our remaining data is then shown on the right hand side of Figure A1.

We then employed multiple imputation proposed by **?** to fill in the remaining missing values. **?**'s multiple imputation fits a sequence of predictive models using one variable as the outcome and the rest of the variables in the data as predictors. In fitting the chains of predictive models, it makes the assumption that the full dataset follows a multivariate normal distribution such that $\mathbf{X} \sim MVN(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, in addition to the standard missing at random (MAR) assumption. We apply logarithm and square-root functions to a subset of our variables to fit the distributional assumption of the multiple imputation. The strength of **?**'s multiple imputation is that it allows researchers to model cross-sectional and time-series features of the target data. We introduce country-fixed effects in all sequences of predictive models since policy choices and policy outcomes correlate strongly with country-specific characteristics such as population, culture, and institutional features that we did not include in our data. To capture temporal dependence, we add one-year lagged dependent variables in all predictive models. We chose one-year lag over the other alternative of polynomial of degree d to avoid overflowing our predictive models with too many variables. Since each run of multiple imputation generates different prediction values for missingness due to uncertainty in the models, we run 20 different multiple imputations and use the average of the resulting 20 complete datasets for the clustering analysis.[2]

Clustering method

With an imputed data set in hand, we begin the search for coherent clusters of external economy policies. Because clusters are likely to exhibit significant temporal dependence and to evolve dynamically over time,

[1] Note that some of our trade variables (total imports and exports of both goods and services) are not represented as a share of GDP. For each of these, we regressed the outcome on a measure of current GDP and dummies for whether the country was landlocked or an island.

[2] The authors' recommendation is to run the imputation 5 times for a moderate level of missingness. The missingness in our pruned data is moderate in the conventional standard, but we chose to run 20 imputations to be safe. The resulting 20 imputed datasets did not vary much, suggesting that 20 was sufficient given the level of missingness in our data.

we employ a Gaussian Mixture Hidden Markov model to uncover latent clustering of the policy matrix over time. We describe the model here in brief. N, P, T are the total number of countries, variables, and time indices in the data, respectively. Countries' policies over time are an $N \times P \times T$ tensor called $\mathbf{X}$. $\mathbf{x}_{it}$ denotes the policy vector of length P for country i at time t. Similarly, $\mathbf{X}_t$ refers to an $N \times P$ policy matrix of countries at time t. All skewed variables were logged, and we then normalize all variables so that they follow an approximately normal distribution.

$z_{it} \in \{1, 2, ..., K\}$ is the latent cluster membership for a country i at time t where K is the number clusters. We model the policy vectors with multivariate normal distributions such that

$$\mathbf{x}_{it} | z_{it} = k \sim \text{MVN}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \quad \text{for } i = 1, 2, ..., N \text{ and for } t = 1, 2, ..., T$$

$\boldsymbol{\mu}_k$ denotes the average policy vector for cluster k and $\boldsymbol{\Sigma}_k$ indicates the covariance for development policies within k. These, along with the latent cluster membership $\mathbf{Z}$, are our quantities of interest.

The cluster membership evolves through a process given by the Hidden Markov Model:

$$z_{it} | z_{i,t-1} = k \sim \text{Categorical}(A_{k,1}, A_{k,2}, ..., A_{k,K}) \quad \text{for } i = 1, 2, ..., N \text{ and for } t = 1, 2, ..., T$$

The transition probability matrix A characterizes the evolution of latent cluster memberships. The kth row and the $k^{'}$th column of $\mathbf{A}$ ($A_{k,k'}$) refers to the transition probability of cluster membership from k to $k^{'}$. When the diagonal entries of $\mathbf{A}$ are high, cluster memberships are relatively constant over time. If, on the other hand, the off-diagonal entries of $\mathbf{A}$ are high, it indicates that countries frequently switch their cluster memberships. We use an EM algorithm to fit the above model following steps delineated in **?**.

Fitting this model requires choosing a number of clusters K. To make this model selection choice, we fit models with $K \in \{2, ..., 10\}$ clusters and then examined the Akaike and Bayesian Information Criteria. We also scrutinized the log-likelihood using an informal elbow heuristic. These are visualized in Figure A2. The Akaike Information Criteria was maximized at $K = 9$ clusters but with a second local mode around the $K = 7$ cluster. The Bayesian Information Criteria was maximized at $K = 7$ clusters, while $K = 9$ was clearly rejected in favor of several alternatives with lower numbers of clusters. In light of these results, and with a view toward parsimony, we opted to set $K = 7$.

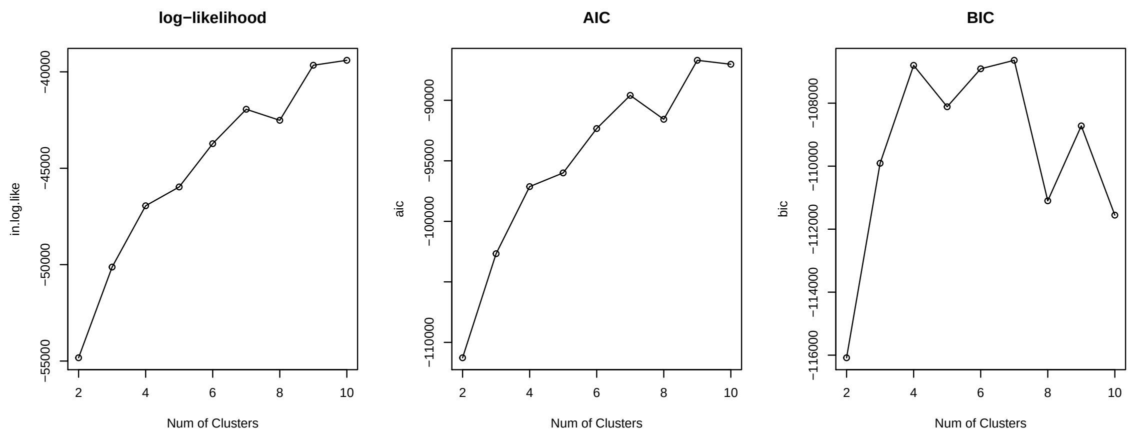

Figure A2: Log-likelihood and information criteria across clustering models

Initial clustering results and post-processing

We then began an intensive examination of our clustering results. To facilitate this, we ordered the clusters in terms of their Mahalanobis distance from a hypothetical 'most orderly' cluster which features the lowest trade barriers, highest number of trade agreements, most fixed currency, lowest capital controls etc. Clusters are then named in terms of their ranking, with cluster 1 the most orderly/open and cluster 7 the least orderly/open. We looked across each of the clusters and attempted to construct verbal descriptions, summary statistics, and emblematic countries that lie closest to the cluster's mean policy vector. We find that 5 of the clusters were easily described in words and *prima facie* represented coherent groups of countries.

Two of the clusters (numbers 4 and 7, representing a moderately orderly/open group and a closed off group) were much harder to describe. Examining the mean Mahalanobis distance for these clusters, we noticed a striking pattern. The average Mahalanobis distance from the cluster mean for the country-years in clusters 1, 2, 3, 5 and 6 were all between 12 and 17. The same figures for clusters 4 and 7 were 30 and 24, respectively. In a similar fashion, less than 5% of the observations in clusters 1/2/3/5/6 had a Mahalanobis distance of more than 40 from the cluster mean. The same figures for clusters 4 and 7 were 21% and 11%. We conclude from this (and after inspection of specific examples) that clusters 4 and 7 represent clusters of outlying countries. For example, Ireland, with its outsized FDI inflows owing to its status as a tax haven, and Zimbabwe, with its hyperinflation, were both included in the moderate cluster 4. The algorithm resisted placing them in clusters 1/2 or cluster 5/6, respectively, because their outlying features would sharply destabilize the cluster mean.

Because our descriptive goal is a parsimonious set of external economic policies that represent coherent policy bundles, we then examined different strategies of cluster reassignment including reassigning only extreme outliers (Mahalanobis distance > 40) and moderate outliers (Mahalanobis distance > 20). Ultimately, we chose to eliminate clusters 4 and 7 completely, and reassign countries in those clusters to their next most likely clusters among 1/2/3/5/6. Most of the Ireland years were reassigned to clusters 1 and 6, for example, while nearly all of the Zimbabwe years were reassigned to cluster 6. Note that our cluster reassignments preserve the dynamic coherence of the cluster assignments from our clustering model.

We thus ended up with 5 remaining clusters, which we relabeled as clusters 1, 2, 3, 4 and 5 (preserving the order from the original labels described above). We found that the resulting clusters have several good properties after reassignment. First, the average distance of cluster observations from the mean increased by less than 6% for clusters 1, 2, 3 and by less around 20% for clusters 4 and 5 (formerly, 5 and 6). Second, the number of observations with distance greater than 40 from the cluster mean remained below 5% for two clusters, below 10% for two other clusters, and just exceeded 10% for the new cluster 4. Our reassigned clusters therefore remain highly coherent, a point reinforced in our substantive investigation of the clusters in the main text.

Appendix B: Additional Models

Simple regression models

Table B1: Results from multinomial regression model with all controls

	$\log\frac{p(\text{C1})}{p(\text{C5})}$	$\log\frac{p(\text{C2})}{p(\text{C5})}$	$\log\frac{p(\text{C3})}{p(\text{C5})}$	$\log\frac{p(\text{C4})}{p(\text{C5})}$
Bivariate model:				
Democracy	22.30^{***}	1.58^{***}	4.08^{***}	2.01^{***}
	(1.39)	(0.27)	(0.30)	(0.24)
Intercept	-17.95^{***}	-1.27^{***}	-2.81^{***}	-1.02^{***}
	(1.19)	(0.15)	(0.20)	(0.14)
Simplest multivariate model:				
Democracy	7.66^{***}	0.26	4.67^{***}	1.68^{***}
	(1.21)	(0.29)	(0.40)	(0.28)
ln GDP per capita	6.39^{***}	1.20^{***}	-0.10	0.41^{***}
	(0.41)	(0.13)	(0.14)	(0.12)
Intercept	-33.35^{***}	-4.86^{***}	-2.82^{***}	-2.30^{***}
	(1.91)	(0.43)	(0.41)	(0.36)

Notes: $N = 2889$ (top half) and $N = 2682$. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.10$.

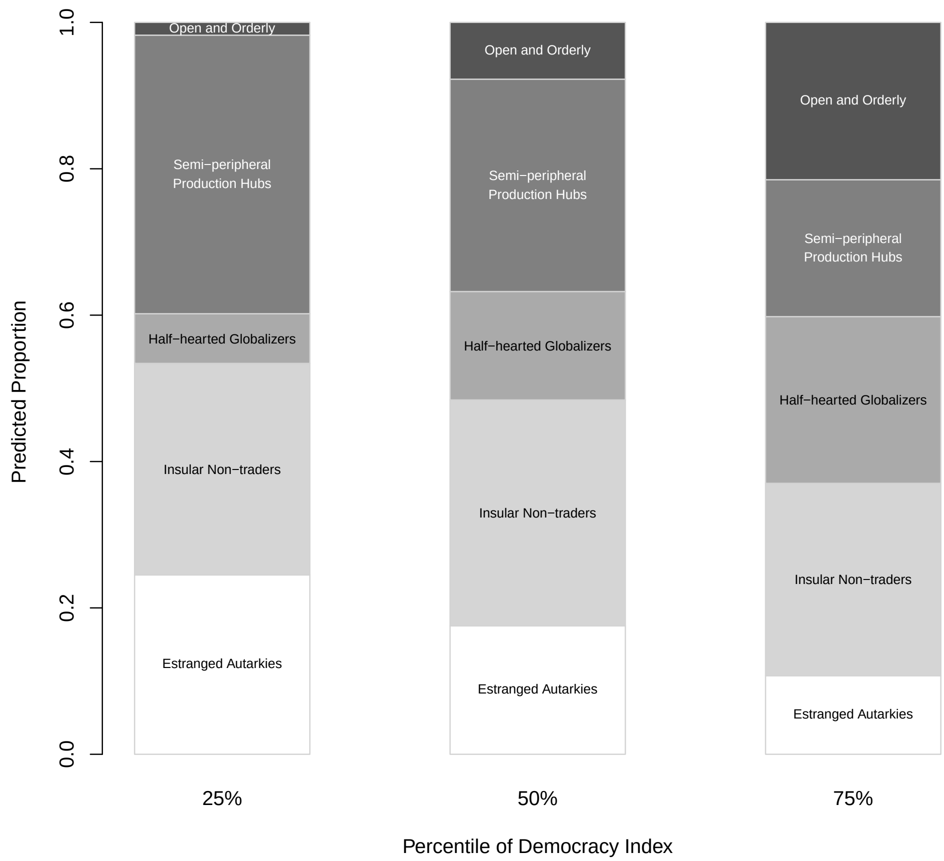

Figure B1: Simple Model: Predicted proportion of clusters as function of democracy at 25%, 50% and 75% percentile

Models investigating cluster reassignment

In this subsection, we consider whether the decision to reassign two clusters in the original 7-cluster model to their next most likely clusters has affected our substantive findings on the links between democracy and clusters 1 and 3 (and non-democracy and clusters 2 and 5). We do this in Table B4. In the middle-third of the Table, we do a simple subsample replication exercise where we fit exactly the same models as in Table B3 but we dropped all country-years who had their clusters reassigned to clusters 1-5. We see results in Table B4 that are extremely similar to the results from the full sample (i.e. the sample with reassigned clusters) from Table 3. In the lower third of Table B4, we examine the results of a model that employs the original 7-cluster clustering. We called the two dropped clusters Dropped 1 and Dropped 2. We again, we strikingly similar results across the 5 main clusters. Thus, our decision to drop some country-years and reassign their clusters is not a significant driver of our findings on the relationships between models/clusters 1-5 and democracy.

Table B2: Results examining original 7 clusters

	$\log\frac{p(\text{C1})}{p(\text{C5})}$	$\log\frac{p(\text{C2})}{p(\text{C5})}$	$\log\frac{p(\text{C3})}{p(\text{C5})}$	$\log\frac{p(\text{C4})}{p(\text{C5})}$	$\log\frac{p(\text{Dropped 1})}{p(\text{C5})}$	$\log\frac{p(\text{Dropped 2})}{p(\text{C5})}$
Estimates from Table B3, original sample:						
Democracy	7.24^{***}	-0.27	5.46^{***}	3.73^{***}		
	(1.40)	(0.52)	(0.51)	(0.37)		
Sample dropping country-years with switched clusters:						
Democracy	6.05^{***}	-1.14	6.80^{***}	6.21^{***}	–	–
	(1.75)	(0.76)	(0.62)	(0.52)	–	–
Original (pre-switch) clustering:						
Democracy	6.87^{***}	0.09	6.31^{***}	6.02^{***}	-0.47	3.19^{***}
	(1.39)	(0.62)	(0.57)	(0.48)	(0.63)	(0.74)

Notes: $N = 2889$. Model also includes continent intercepts and cubic polynomial time trend; intercepts are not reported. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.10$.

Models fit on non-imputed data

In this section, we refit the models from Table 3 and B2 on non-imputed data. Note that this reduces the sample size, for example, from 2889 in Table 3 to 2166 in Table B3. We see results that are extremely similar.

Table B3: Results from multinomial regression model with all controls (unimputed data)

	$\log\frac{p(\text{C1})}{p(\text{C5})}$	$\log\frac{p(\text{C2})}{p(\text{C5})}$	$\log\frac{p(\text{C3})}{p(\text{C5})}$	$\log\frac{p(\text{C4})}{p(\text{C5})}$
Democracy	6.86^{**}	-1.08	5.77^{***}	4.58^{***}
	(2.09)	(0.69)	(0.61)	(0.51)
Developmental controls:				
Population	-4.51^{***}	-3.29^{***}	-3.06^{***}	-0.23
	(0.45)	(0.27)	(0.27)	(0.19)
GDP pc	8.21^{***}	-1.47^{***}	-1.74^{***}	0.82^{**}
	(0.96)	(0.39)	(0.35)	(0.26)
Economic Complexity	3.20^{***}	2.41^{***}	0.35^{+}	0.33^{*}
	(0.39)	(0.24)	(0.20)	(0.14)
Conflict controls:				
Militarized dispute	-0.03	-0.10^{***}	0.05^{**}	-0.05^{***}
	(0.02)	(0.02)	(0.02)	(0.01)
Militarized threat	0.01	-0.12^{*}	0.01	0.08^{***}
	(0.03)	(0.05)	(0.03)	(0.02)
US ally	0.37	0.31	0.29	0.05
	(0.59)	(0.38)	(1.46)	(0.28)
Historical controls:				
Former Colony	-1.40^{**}	0.72^{*}	-2.03^{***}	0.11
	(0.46)	(0.33)	(0.39)	(0.27)
Former Soviet State	-0.38	-0.57	-4.48^{***}	-2.04^{***}
	(0.85)	(0.51)	(0.59)	(0.59)
Other Communist	2.12^{**}	1.03^{**}	-11.57	0.86^{**}
	(0.69)	(0.35)	(11.89)	(0.30)
Geographical controls:				
Landlocked	-3.76^{***}	-7.71^{***}	-0.86^{***}	0.01
	(0.66)	(0.81)	(0.24)	(0.24)
Islands	-3.00^{***}	-0.54^{+}	-9.49	-0.28
	(0.46)	(0.31)	(7.24)	(0.26)
Territorial contiguity	0.00	0.02	0.14^{***}	0.11^{***}
	(0.05)	(0.04)	(0.04)	(0.03)

Notes: $N = 2166$. Model also includes continent intercepts and cubic polynomial time trend; intercepts are not reported. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.10$.

Figure B2: Full model: Predicted cluster proportions as a function of democracy (unimputed data)

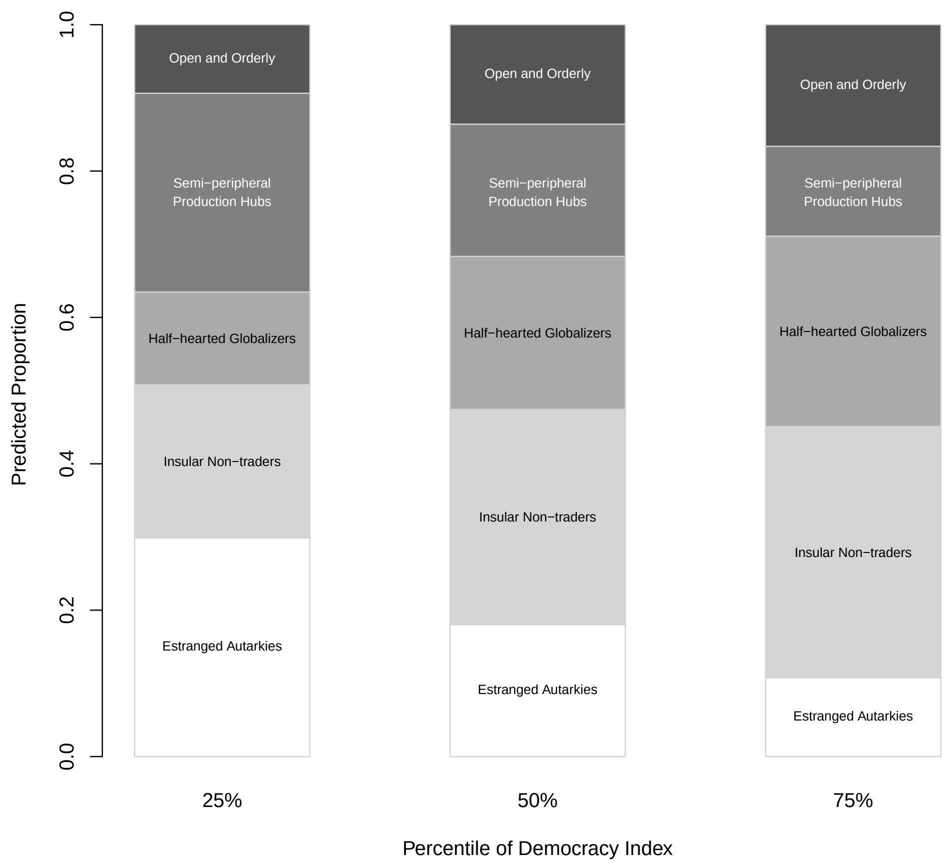

Table B4: Results examining original 7 clusters (unimputed data)

	$\log\frac{p(\text{C1})}{p(\text{C5})}$	$\log\frac{p(\text{C2})}{p(\text{C5})}$	$\log\frac{p(\text{C3})}{p(\text{C5})}$	$\log\frac{p(\text{C4})}{p(\text{C5})}$	$\log\frac{p(\text{Dropped 1})}{p(\text{C5})}$	$\log\frac{p(\text{Dropped 2})}{p(\text{C5})}$
Estimates from Table B3, original sample:						
Democracy	6.86^{**}	-1.08	5.77^{***}	4.58^{***}	–	–
	(2.09)	(0.69)	(0.61)	(0.51)	–	–
Sample dropping country-years with switched clusters:						
Democracy	6.97^{**}	-1.98^{*}	6.00^{***}	5.93^{***}	–	–
	(2.46)	(0.90)	(0.68)	(0.62)	–	–
Original (pre-switch) clustering:						
Democracy	5.03^{*}	-1.37^{+}	5.56^{***}	5.57^{***}	-2.77^{**}	1.87^{+}
	(2.12)	(0.79)	(0.65)	(0.60)	(0.89)	(1.03)

Notes: $N = 2166$. Model also includes continent intercepts and cubic polynomial time trend; intercepts are not reported. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{+}p < 0.10$.