Dynamic factor models applied to aggregate indicators of political and economic institutions

January, 2023

Abstract

1 Introduction

While the roles of economic and political institutions in economic development have been in the spotlight for many years, there remains significant weaknesses in the conceptualization, measurement and aggregation methods used by indicators that measure the quality of these institutions. In this paper, we provide a clear link between the theoretical and empirical literature by creating aggregate indicators for economic institutions and democratic political institutions that are consistent with the theoretical literature and have a sufficiently large geographical and temporal coverage to be used in panel data studies. In this paper, we particularly use dynamic factor models (**???**), which besides being a tool used to reduce the dimensionality of multivariate data, it considers the uncertainty about the aggregate indicator. It provides estimation of common latent factors and its factor loadings, that often have an interesting theoretical interpretation in real problems. Moreover, the multiple structure of the dynamic unobserved latent factors allow to characterize the temporal comovements in the cross-country panel of economic and political institutions time series, minimizing problems regarding the comparability of scores between countries and over time that are present in existing aggregate indicators.

Building on North (1990) and North (1993), we define economic institutions as the formal rules, enforcement mechanisms and informal norms that govern the functioning of property and contract rights. Comparably, we define political institutions as the formal rules, enforcement mechanisms and informal norms that constrain the ability of political actors – both citizens and rulers – to influence political decision making. In particular, democratic political institutions guarantee the rights of citizens to participate in political decision-making, and place constraints on the power of rulers.

While several studies that have aimed to measure the quality of economic institutions, there has been little work on measuring the quality of democracy using the theoretical framework on institutions developed by North and his successors. Even North (1990), which we use as the theoretical basis of our democracy indicators, did not develop the theoretical analysis to systematically set out the essential attributes of democratic political institutions, nor did it analyze what types of data could be used to assess whether countries possess these attributes.

Some motivations to consider an aggregate indicator are: (i) indicators for democracies and economic institutions are subject to considerable measurement error and one way to minimize this is to aggregate data from several different sources (Knack, 2006); (ii) when measuring democracy, it can capture data on several features that may not all be represented in any one disaggregated indicator. Many of the most commonly used disaggregated democracy indicators do not include information on the three attributes that we argue are vital for democratic political institutions – electoral self-determination, political liberties and horizontal constraints on political rulers. Several democracy indicators do not include coded data on political liberties, including Polity IV Project (Marshall, et al, 2010), Boix et al's Dataset of Political Regimes (2012) and Cheibub et al's Democracy and Dictatorship Index (2009). Meanwhile, other indicators such as the Freedom House's Freedom in the World indicators and CIRI do not specifically measure the extent to which chief executives are constrained by legislatures or other accountability groups, a key element of democracy.

There have been very few indicators for democratic political institutions and economic institutions that aggregate data from different sources, and many of them have very limited geographical and temporal coverage. The Worldwide Governance Indicators (WGIs), the most commonly cited aggregate indicators on several governance issues, capture many of the theoretical dimensions of both economic institutions and democratic political institu-

tion, and have a broad geographical and temporal coverage.However, they also have several weaknesses as proxies for these concepts. The WGI that measures many attributes of democracy is Voice and Accountability (V & A), which captures "perceptions of the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media". The WGIs only goes back to 1996, whereas our aggregate democracy indicator goes back to 1980.

The WGI's Rule of Law indicator captures many relevant dimensions of property and contract rights, but does not include corruption, (which they measure separately in the Control of Corruption indicator) and which we consider an important source of insecurity of property rights. Moreover, the Rule of Law indicator contains data regarding issues that are not manifestations of secure property and contract rights, for example the efficiency of the legal framework for challenging regulations and government policies on tax evasion. In particular, each of the WGI indicators bundle together a large number of data sources without justifying the choice of sources using theory. In contrast, we use a much more parsimonious selection of data sources, explaining how each of the variables in our aggregate indicators correspond to the theoretical dimensions of economic institutions and democratic political institutions that we have identified.

Another commonly cited criticism of the WGIs is regarding the comparability of scores across countries and over time (Arndt and Oman, 2006 and Knack 2006), which consists in our main aim in this paper. Firstly, Arndt and Oman (2006), argue that because they are constructed such that the average value of the indicator across all countries worldwide is always zero and its standard deviation is always one, the indicators "do not allow for a reliable comparison of levels of governance over time - for given countries or regional and global trends". Although Kauffman et al (2007) argues that there is little evidence upward or downward trend in governance levels, this is not the case for democracy from 1980 to 1996 (the WGIs began in 1996) and there may be significant global trends in the future. Secondly, the WGI indicator has incorporated many new data sources over time, which means that the weights assigned to each source for a given country may have changed significantly over the years, affecting intertemporal comparability of scoresThirdly, the data sources, and their respective weights, used to make the indicators vary greatly from country to country, affecting the comparability of country scores (Knack, 2006).

This paper sets out to contribute to the conceptualization, measurement

and mainly in the propose of aggregation methods used to statistically measure the quality of economic institutions and democratic political institutions by creating new aggregate indicators for these concepts.

We first discuss the improvement in the measurement of these concepts by carefully examining what types of data can be used to represent the key theoretical dimension of the concepts outlined in the literature, as well as considering other features of the data, including temporal and geographical coverage. Then, dynamic factor models with multiple factors are applied to describe the original dependence of the data sources by a smaller number of unobservable latent variables over time. In particular, we are interested in a one-dimensional aggregate indicator. The proposed dynamic factor model alows to capture dynamic fluctuations of the latent factors accross countries and for the world. Besides, considering uncertanty, the method considered in this paper addresses several of the criticisms of the WGIs regarding the comparability of scores across countries and over time. The inference is done under the Bayesian approach and we make use of Markov chain Monte Carlo (MCMC) methods to sample from the posterior distribution, since its kernel does not result in a known distribution.

The remainder of the paper is organized as follows. Section 2 highlights the essential theoretical attributes of democratic political and economic institutions discussed in the literature. Subsequently, we outline and justify the data chosen for our aggregate indicators, including the limitations of existing democracy indicators. Section 3 presents the proposed dynamic factor model with multiple latent factors. The model allows to create an aggregate indicator common for all the countries and a country-specific aggregate indicator, both varying over time. Section 4 applies the methodology to political and economic institutions. Finally, Section X discusses our main findings.

2 Data sources

2.1 Democratic Political Institutions

We define democracy in terms of the distribution of political constraints on ruler and citizens. In democracies, political institutions guarantee the rights of citizens to participate in political decision making – through free and fair elections, free speech, and freedom of association and assembly – and place constraints on the use of power of the executive, for example through checks

and balances by the legislative and judiciary.

Democratic political institutions enable a broad cross-section of society to participate in the process of political exchange by allowing citizens to express their political preferences at low cost to themselves, and ensuring that elected representatives mediate between these preferences.

Our indicator only takes into account the procedures by which political decisions are made - the extent to which citizens can participate in, and rulers are constrained from dominating, political decision making. It is not an outcome measure that assesses the government performance - the extent to which government policy fulfills the desire or needs of the people, or the manner in which the government carries out its business i.e. transparency and corruption.

One of the compelling advantages of our democracy indicator is that unlike many of the most commonly used disaggregated indicators, it contains coded data on all three components of democracies. Many democracy indicators do not include coded data on political liberties, including Polity IV Project (Marshall, et al, 2010), Boix et al's Dataset of Political Regimes (2012) and Cheibub et al's Democracy and Dictatorship Index (2009).

Our aggregate indicator for democratic political institutions indicator consists of four separate data sources. For each of them, we carefully filtered out any data that did not fit in with the theoretical properties of democratic political institutions. Additionally, we only selected data sources that have a very large geographical coverage of countries and a sufficiently long time-series. To minimize problems of intertemporal consistency, we used the same data sources for every time period covered by the aggregate indicators. Based on the criteria above, we used the following four data sources: Freedom of the Press by Freedom House; Polity IV Project; Cingranelli Human Rights Project; and V-Dem.

We use Freedom House's Freedom of the Press indicator as a measure of the extent that democratic political institutions protect free speech. The index captures the various ways in which pressure can be placed on the flow of information and the ability of print, broadcast, and internet-based media to operate freely and without fear of repercussions.

The second data source we use is the Polity IV project (**?**), which is one of the most extensively used democracy indicators in the literature. We use two of the three concept variables – Executive Recruitment (EXREC) and Executive Constraints (EXCONST). The EXREC variable measures how institutionalized, competitive and open the mechanisms are for selecting a

political leader. In order to attain the top score, countries must have free, fair and competitive elections for the chief executive. As such, the indicator corresponds very well with the electoral self-determination component of democratic political institutions. The indicator EXCONST measures the extent of institutional constraints on the decision-making powers of the chief executive. This is similar to the notion of horizontal accountability found in the democracy literature but it assumes that dictators may also be bound by certain institutional constraints. Thus, EXCONST variable fits neatly with our horizontal component of democratic political institutions.

The third data source we chose is the CIRIGHTS Database (**??**), from which we used three out of their ten indicators – Freedom of Speech, Freedom of Assembly and Association, and Electoral Self-Determination.

Our fourth data source is V-Dem Dataset Version 9 (Coppedge et al., 2019), a recent project to conceptualize and measure democracy developed in the University of Gothenburg and the Kellogg Institute, with support from the European Commission, and several research institutes and government ministry's around the world. V-Dem has taken several steps to minimize measurement error and improve cross-country and intertemporal reliability.

To represent executive constraints we added together the scores of two composite indicators that well represent the concept – judicial constraints on the executive and legislative constraints on the executive.

To represent electoral self-determination, we combined the scores of the variables share of population with suffrage and the clean elections index.

We added the scores for electoral self-determination and executive constraints to give a total score for the indicator.

We ensured that all three components of democratic political institutions are well-represented by the data we used to compile the aggregate indicator.

A more detailed explanation about the construction of each index in Appendix **??**.

It is worth noting that we did not use Freedom House's Freedom in the World reports, which are used extensively as a democracy indicator, because it contains information that does not correspond to the theoretical literature on political institutions. Freedom in the World consists of two indicators - Political Rights and Civil Liberties. We did not use the Political Rights indicator because one of the three categories, which is called "the functioning of government", measures quality of governance issues such as transparency, corruption and bureaucratic regulations, which lie outside our conception of political institutions.

2.2 Economic Institutions

Our definition of economic institutions focuses on the rules, informal norms and enforcement mechanisms governing property and contract rights. While North (1990) did not specifically define economic institutions, he defined economic rules in terms of property rights and used the terms interchangeably. Several subsequent studies have continued to define economic institutions in terms of property and contract rights, for example Borner et al (2014). However, several other studies have used a broader definition of economic institutions to incorporate other types of government policies that determine economic incentives and the functioning of markets. Many empirical studies, for example (Barro, 1997) and (Rigobon and Rodrik, 2005), use the broader concept of rule of law, defined by Guttman and Voigt (2015) as the "subordination of any government power to well-defined and established laws". In contrast, our narrower definition of economic institutions allows us to clearly focus, both theoretically and empirically, on the rights of individuals to hold property and exchange economic goods and resources. Similarly to our measurement of democratic political institutions, our indicator of economic institutions takes into account both the both the rules governing property and contract rights and the enforcement mechanisms and informal norms that complement them. The importance of the role of this holistic analysis is highlighted by the experience of several developing countries that adopted property rights laws of wealthy Western countries, but did not enjoy the same economic outcomes (North, 1990). "Although the rules are the same, the enforcement mechanisms, the way enforcement occurs, the norms of behavior and the subjective models of the actors are not. Hence, both the real incentive structure and the perceived consequence s of policies will differ as well."

There is a lack of data regarding economic institutions that have a global geographical coverage and a sufficiently long time series for dynamic panel data models. Among the data sources chosen in this study, Political Risk Services (PRS) ICRG indicator is the only one that goes back to the 1980s.We used three indicators from PRS - Investment Profile, Law and Order and Corruption. Investment Profile itself has three components - Contract viability, Profits Repatriation and Payment Delays. Contract viability measures the extent to which the government and the judicial system of the country uphold business contracts and discriminates against foreign individuals and firms. Profits repatriation measures the ability of firms to convert its profits

to hard currency and to return these profits to the investors' home country, and thus affects the security of property rights for foreign investors. Payment delays - which measures "the extent to which payments to foreign investors, whether in government contract or in direct sales are able to obtain cash payment for goods and services in a timely manner" – can be considered a proxy for the enforceability of contracts because in countries where contracts are well enforced, payments are less likely to be delayed (Howell, 2007, p65).

PRS's Law and Order indicators contains two sub components that are assessed separately: the force and impartiality of the legal system ("law") and popular observance of the law ("order"). The former measures the extent of case precedent and the consistency of legal legislation and practice, which contribute to well specified and secure property and contract rights. The latter subcomponent is an assessment of popular observance of the law, which is in part a willingness of the population to be self-regulating but also has to do with the performance of law officials to enforce the laws of the country.

PRS's Corruption indicator includes financial corruption in the form of demands for special payments in connections with business activities, which as we mention above, constitute a violation of property rights. The indicator also takes into account other forms of corruption including excessive patronage.

The second data source we use is Heritage Foundation's Property Rights indicator, which measures the degree to which a country's laws protect private property rights and the degree to which its government enforces these laws. It assesses expropriation risk, the efficiency, independence and corruptibility of the judiciary, as well as the enforceability of contracts.

Our third data source is Economist Intelligence Unit's indicators of Rule of Law and Control of Corruption, as used by WGI. The Control of Corruption indicator measures the extent of corruption among public officials. We assigned the weight of the Rule of Law indicator to be four times greater than the Control of Corruption indicator because corruption is just one of many forms of violations of private property. In contrast, the Rule of Law has seven subcomponents that are directly related to property and contract rights: Private property protection, intellectual property rights protection, expropriation risk, speediness and fairness of judicial process and enforceability of contracts. The indicator also measures the extent of violent and organized crime, which can be regarded as proxies for the extent to which the government protects private property from theft by other private agents.

Our fourth data source is Global Insight's Rule of Law indicator, as used

by WGI. The first subcomponent, Judicial Independence, measures the extent to which the government or other outside actors can influence legal decisions, a critical aspect of property and contract enforcement. The second subcomponent, Crime, measures the threat faced by businesses from crimes, including kidnapping, burglary and extortion, and as such, reflects the extent to which the government protects private property from theft by other private agents.

3 Dynamic factor model with multiple-factor

Let y_{jt} denotes a vector with $K = 4$ related data sources for country j in the year t, $j = 1, \ldots, J$ and $t = 1, \ldots, T$. Factor models are structured from a linear model that relates observable variables to the latent factors plus a random error component. The dependence structure of the original variables is explained by the matrix of factor loadings and the dynamic over time is characterized by the latent factors. The latent factors are $q-$vectors, $q << K$ and describes temporal similarities amongst the time series, such as common annual cycles or trends. We consider two types of aggregate indicators: country-specific factors varying over time, denoted by f_{jt}^c for the country j at time t and a single commom world factor, denoted by f_t^w at time t. The dimensions of f_{jt}^c and f_t^w deppends on the evolution structure assumed. By this way, let f_t^w be a q_w-vector and f_{jt}^c a q_c-country, such that $q = q_w + q_c < K$. While one aggregate indicator allows to characterize the temporal comovements in the cross-country panel of political and economic indicators time series, the other recovers a commom dynamic for all the countries.

The model assumes the K-vector of time series y_{jt} in the year t for the country j is a measurement of country and single world latent factors, which vary over time. The associations among the observed variables are wholly explained by the latent variables through the following factor model: for $j = 1, \ldots, J$ and $t = 1, \ldots, T$,

$$\begin{array}{rcl} y_{jt} & = & \beta^w f_t^w + \beta^c f_{jt}^c + \epsilon_{jt}, \epsilon_{jt} \sim N_K(0_K, \Sigma) \\ f_t^w & = & \gamma^w f_{t-1}^w + v_t^w,\ v_t^w \sim N_{q_w}(0_{q_w}, \lambda^w) \\ f_{jt}^c & = & \gamma_j^c f_{j,t-1}^c + v_{jt}^c,\ v_{jt}^c \sim N_{q_c}(0_{q_c}, \lambda^c), \end{array} \tag{1}$$

where β^c and β^w are the factor loadings matrices, with dimensions $K \times q_c$ and $K \times q_w$, respectively. Both β^c and β^w reflect the degree to which

variation in y_{jt} can be explained by the country and the world common factor, respectively. The factors f_t^w and f_{jt}^c are supposed independent for different countries, but time-dependent. The term ϵ_{jt} is an independent K-vector which follows a K-multivariate Normal distribution with vector mean 0_K and covariance matrix $\Sigma = diag(\sigma_1^2, \ldots, \sigma_K^2)$. The dynamic evolution of the world-specific latent factor and country-specific latent factor are mainly driven by the $q_w \times q_w$-matrix γ^w and the $q_c \times q_c$-matrix γ_j^c, respectively. In particular, the dynamic evolution γ_j^c allows to vary depending on the country j, $j = 1, \ldots, J$. Both γ^w and γ^c can accommodate seasonalities, stationary and non-stationary trends as well as covariate effects. Finally, λ^w and λ^c are the evolution variances with dimensions $q_w \times q_w$ and $q_c \times q_c$, respectively.

In a more general way, for $q = q_w + q_c$, let us define: $\beta = (\beta^w, \beta^c)$ a $K \times q$-matrix with the world and country factor loadings, $f_{jt} = (f_t^w, f_{jt}^c)'$ a q-vector with latent factors, $v_{jt} = (v_t^w, v_{jt}^c)'$ a q-vector with the evolution errors, $\Gamma_j = diag(\gamma^w, \gamma_j^c)$ and $\Lambda = diag(\lambda^w, \lambda^c)$, both $q \times q$-matrices. Thus, the model in (1) may be rewritten in a summarized way as:

$$\begin{array}{ll} y_{jt} & = \beta f_{jt} + \epsilon_{jt}, \epsilon_{jt} \sim N_K(0_K, \Sigma) \\ f_{jt} & = \Gamma_j f_{j,t-1} + v_{jt},\ v_{jt} \sim N_q(0_q, \Lambda), \end{array} \tag{2}$$

for $j = 1, \ldots, J$ and $t = 1, \ldots, T$.

The matrix Γ_j characterizes the evolution dynamics of the common factors for country j, while Σ and Λ are observational and evolutional variances, respectively. Stationary and non-stationary common factors could be considered. This model could be considered an extension of **?**'s model, since beside our proposed model presents a multiple-factor structure, it allows the evolution of the country-specific aggregate factor to vary by country.

Common dynamic factors can be thought of as describing temporal similarities amongst the time series, such as common annual cycles or (stationary or nonstationary) trends, while the importance of common factors in describing the measurements in a given country or in the world is captured by the components of the factor loadings matrix.

3.1 Some model properties

The dynamic factor model suffers from an identifiability problem, due to the invariance of factor models under orthogonal transformations. This is, if one considers $\beta^* = \beta\Phi^{-1}$ and $f_{jt}^* = \Phi f_{jt}$, for any nonsingular matrix Φ, the same

model defined in (2) is obtained. To deal with the factor model invariance, the alternative adopted in this paper is to constrain β to be a block lower triangular matrix, assumed to be of full rank, with diagonal equals to 1. This form provides both identification and often useful interpretation of the factor model (**?**).

Since the inference procedure is done conditional on the latent factors and then integrating them out, it is not strictly necessary to obtain marginal properties of the models. Nevertheless, from the model (2) we obtain the joint model

$$\begin{pmatrix} y_{jt} \\ f_{jt} \end{pmatrix} \sim N_{K+q}\left[\begin{pmatrix} 0_K \\ 0_q \end{pmatrix}, \begin{pmatrix} \beta\phi_j\beta' + \Sigma & \beta\phi_j \\ \phi_j'\beta' & \phi_j \end{pmatrix}\right], \tag{3}$$

where ϕ_j satisfies $\phi_j = \Lambda + \Gamma_j\phi_j\Gamma_j'$ is a block-diagonal $q \times q$-matrix. Under this specification, marginal on the latent factors f_{jt}, the components of the K-vector y_{jt} are not independent.

Thus, the marginal variance structure of y_{jt}, given by:

$$\Delta_j = var(y_{jt} \mid \beta, \Sigma, \Gamma_j, \Lambda) = \beta\phi_j\beta' + \Sigma, \tag{4}$$

is divided into a part explained by the common factors and the uniquenesses $\Sigma = diag(\sigma_1^2, \ldots, \sigma_K^2)$, which measures the residual variability in each of the data variables once that contributed by the agreggate factor is accounted for.

The models imply that, conditional on the common factors, the K time series are uncorrelated. Hence, the common latent factor explain all the dependence structure among the K variables. Then, for any elements y_{jtk} and $y_{jtk'}$ of y_{jt}, we have the following conditional moments:

$$var(y_{jtk} \mid f_{jt}) = \sigma_k^2, \;\; cov(y_{jtk}, y_{jtk'} \mid f_{jt}) = 0,$$

and marginal ones:

$$\begin{aligned} var(y_{jtk}) &= \beta_k\phi_j\beta_k' + \sigma_k^2 = \beta_k^w\phi_{j,11}\beta_k^{w\prime} + \beta_k^c\phi_{j,22}\beta_k^{c\prime} + \sigma_k^2 \\ cov(y_{jtk}, y_{jtk'}) &= \beta_k\phi_j\beta_{k'}' = \beta_k^w\phi_{j,11}\beta_{k'}^{w\prime} + \beta_k^c\phi_{j,22}\beta_{k'}^{c\,\prime}, \end{aligned}$$

where β_k, β_k^w and β_k^c are the k-lines of the matrix β, β^w and β^c, respectively, $\phi_{j,11}$ is the first $q_w \times q_w$-block matrix of ϕ_j and $\phi_{j,22}$ is the second $q_c \times q_c$-block matrix of ϕ_j.

A fairly standard way to summarize the importance of the common factor by its percentage contribution to the variability of the original dataset is given

by the variance decomposition. The variance decomposition for indicator k and country j is given by:

$$DV_{jk} = 100\frac{\beta_k \phi_j \beta_k'}{\beta_k \phi_j \beta_k' + \sigma_k^2}\% = 100\frac{\beta_k^w \phi_{j,11} \beta_k^{w\prime} + \beta_k^c \phi_{j,22} \beta_k^{c\prime}}{\beta_k^w \phi_{j,11} \beta_k^{w\prime} + \beta_k^c \phi_{j,22} \beta_k^{c\prime} + \sigma_k^2}\%. \quad (5)$$

Higher values for it enhance the result that the latent factors can explain well a variable regarding their interpretation. In this case, we can also evaluate the DV for the world latent factor separated from the country latent factor.

3.2 Inference

Let $\mathbf{y} = (y_{11}, y_{12}, \ldots, y_{1T}, \ldots, y_{J1}, \ldots, y_{JT})$ denote the vector with the observations, $\mathbf{f} = (f_{11}, f_{12}, \ldots, f_{1T}, \ldots, f_{J1}, \ldots, f_{JT})$ the vector with all the latent factors and $\mathbf{\Theta} = (\beta, \Sigma, \Gamma_1, \ldots, \Gamma_J, \Lambda)$ be the parameter vector. The inference procedure is performed under the Bayesian paradigm using data augmentation. The idea is to determine posterior distribution for the parameter vector conditional on the latent factor and then the conditional distribution of the latent factor given the observable and the other parameters. That is, the observable data are augmented by samples from the conditional distribution for the factor given the data and the parameters of the model.

The joint distribution of $\mathbf{y}$ and $\mathbf{f}$ is given by:

$$p(\mathbf{y}, \mathbf{f} \mid \mathbf{\Theta}) \propto |\Sigma|^{-TJ/2} |\Lambda|^{-TJ/2} \prod_{j=1}^{J} \prod_{t=1}^{T} \left\{ \exp\left[-\frac{1}{2} \left(y_{jt} - \beta f_{jt} \right)' \Sigma^{-1} \left(y_{jt} - \beta f_{jt} \right)' \right] \right.$$
$$\left. \times \exp\left[-\frac{1}{2} \left(f_{jt} - \Gamma_j f_{j,t-1} \right)' \Lambda^{-1} \left(f_{jt} - \Gamma_j f_{j,t-1} \right)' \right] \right\}$$

The model specification is complete after assigning the prior distribution. For simplicity, we assume $\lambda^c = \lambda^w = 1$ and that the world and country-specific latent factors to be one-dimensional, this is $q_w = q_c = 1$. The following conditionally conjugate independent prior distributions were assumed (**?**): $f_{j0}^c \sim N(0,1)$ and $f_0^w \sim N(0,1)$, $\beta^w \sim N_K(m_\beta^w, \tau_\beta^w)$ and $\beta^c \sim N_K(m_\beta^c, \tau_\beta^c)$, $\gamma^w \sim N_{(-1,1)}(m_\gamma^w, \tau_\gamma^w)$, $\gamma_j^c \sim N_{(-1,1)}(m_\gamma^c, \tau_\gamma^c)$ for $j = 1, \ldots, J$, $\sigma_k^2 \sim IG(\nu_\sigma/2, \nu_\sigma s_\sigma/2)$ for $k = 1, \ldots, K$, where $N_{(a,b)}$ denotes the truncated Normal in the interval (a, b) and $IG(\nu/2, \nu s^2/2)$ denotes the inverse Gamma distribution with mode s and degrees of freedom hyperparameter ν.

Following Bayes' theorem, the posterior distribution of $\boldsymbol{\Theta}$ is proportional to the product of the likelihood function and the prior distribution for $\boldsymbol{\Theta}$. The kernel of this distribution does not result in that of a known distribution. Specifically, the joint posterior distribution for the unknown parametersand the unobserved factors can be sampled by using a Markov Chain Monte Carlo (MCMC) procedure on the full set of conditional distributions. The MCMC algorithm was implemented in JAGS (**?**) programming language, but all the analysis were done in the software R (**?**).

3.3 Missing data

Some time series present missing values for some years and we use a Bayesian multiple imputation to handle this. In a Bayesian framework, missing data can be treated as any other parameter in the model, which means that we are able to apply MCMC algorithm to iteratively draw samples not only for the parameter vector $\boldsymbol{\Theta}$, but also for the missing outcome data from conditional distributions, and perform multiple imputation to make appropriate statistical inference. Since the estimation was done using JAGS, the missing data is automatically estimated.

4 Application to aggregate indicators

As previously mentioned, indicators for democracies and economic institutions are subject to considerable measurement error. Work with an aggregate indicator can bring several advantages. The main aim of this application is to produce an aggregate indicator that captures the evolution over the years. In particular we apply the proposed model presented in Equation (2), which allow us to produce two kind of aggregate indicators, one which varies by country and another one which is a common factor.

Four data sources were used to construct the aggregate indicators for political and economic institutions. In the inference procedure, we assumed non-informative priors for all the parameters and estimate the missing data.

We assign reasonably vague priors to model parameters, more specifically, we assume, $m_\beta^w = m_\beta^c = 0$, $\tau_\beta^w = \tau_\beta^c = 100$, $m_\gamma^w = m_\gamma^c = 0$, $\tau_\gamma^w = \tau_\gamma^c = 1$, $\nu_\sigma = 0.02$, $s_\sigma = 1$. On the MCMC algorithm, for each sample, we run two parallel chains starting from very different values, we let each chain run for 650,000 iterations, discarded the first 50,000 as burn in, and stored every

200th iteration. We used the diagnostic tools available in the package CODA (**?**) to check convergence of the chains.

4.1 Political institutions

As previously mentioned, we used the following four data sources: Freedom of the Press by Freedom House (FIW), Polity IV Project (PIV), Cingranelli Human Rights Project (CIRI), and V-Dem. Thus, the dataset is composed by $K = 4$ annual time series (from 1981 to 2016) for each one of the $J = 101$ countries from the 7 regions (Europe and Central Asia, South Asia, Middle East & North Africa, Sub-Saharan Africa, Latin America & Caribbean, East Asia & Pacific and North America). Since the four data sources are defined in the interval (0,1), we applied the logistic transformation, this is $h(x) = log(x/(1-x))$, to the four data sources. Then, the proposed model defined in Equation (2) is applied to the transformed data.

Table 1 presents the posterior mean and the 95% credibility interval (in parenthesis) of the factor loadings associated to the world and contry latent factors. All the factor loadings are positive and signifficant. The factor loadings for V-Dem are fixed in one, due to the invariance of factor models under orthogonal transformations. Although, the country and world factors has a strong association with all the sources, this association becomes less strong with the source FIW.

Table 1: Posterior distribution of the factor loadings: posterior mean and 95% credibility interval in parenthesis.

	V-Dem	PIV	CIRI	FIW
β^w	1	1.83 (1,73,1.93)	0.86 (0.81,0.92)	0.43 (0.39,0.46)
β^c	1	0.86 (0.84,0.88)	0.89 (0.87,0.92)	0.62 (0.61,0.63)

Table 2 presents the posterior mean of the other parameters of the proposed model in Equation (2).

Table 2: Posterior mean of the other model parameters.

Param	σ_1^2	σ_2^2	σ_3^2	σ_4^2	γ^w	γ^c	λ^w	λ^c
Mean	0.001	1.014	1.619	0.403	0.996	0.986	0.002	0.039

Figure 1 shows a boxplot with the variance decomposition for each type of latent factor. It presents the distribution of the percentage of the variance

of each data source, for all the years, separated by the type of factor. For all the sources the world latent factors explain a higher percentage of the variance than the country latent factors. The boxplots for the FIW source present more variability than the others, it means that for some countries and some years the variance decomposition associated to the country factor becomes larger when compared to the world latent factor.

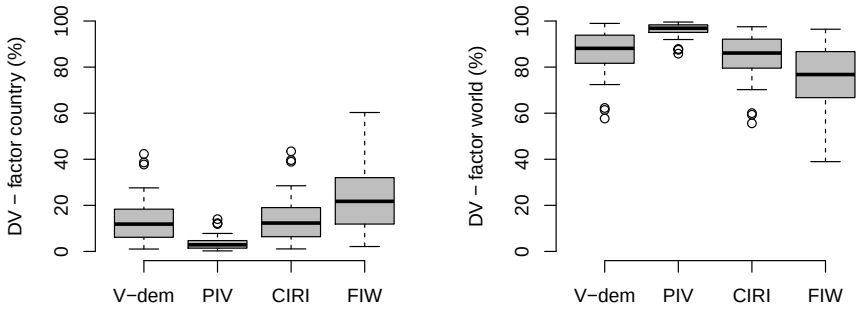

Figure 1: Boxplots with the variance decomposition for each data source divided by the world latent factor and the country-specific factor.

The advantage of the proposed methodology is that beside creating two aggregate indicators, one common for all the contries and other specific for each country, it allows to taking to investigante its behaviors over time. Figure 2 presents the posterior mean and 95% credibility interval for the country-specific latent factors for the years 1986, 1996, 2006 and 2016. Beside there are significant differences between the countries for all the years considered, some countries change position over the years.

Finally, Figure 3 displays the posterior mean of the world latent factor with the 95% credibility interval, from 1981 to 2016. The world factors seems to be significant for all the years with a increasing behavior over time. It shows the tendency for political institutions to improve during this time.

4.2 Economic institutions

As previously mentioned, we used the following four data sources: Heritage Foundation's Property Rights (HERI), the Political Risk Services ICRG indicator (PRS), the Global Insight's Rule of Law indicator (IHS) and the

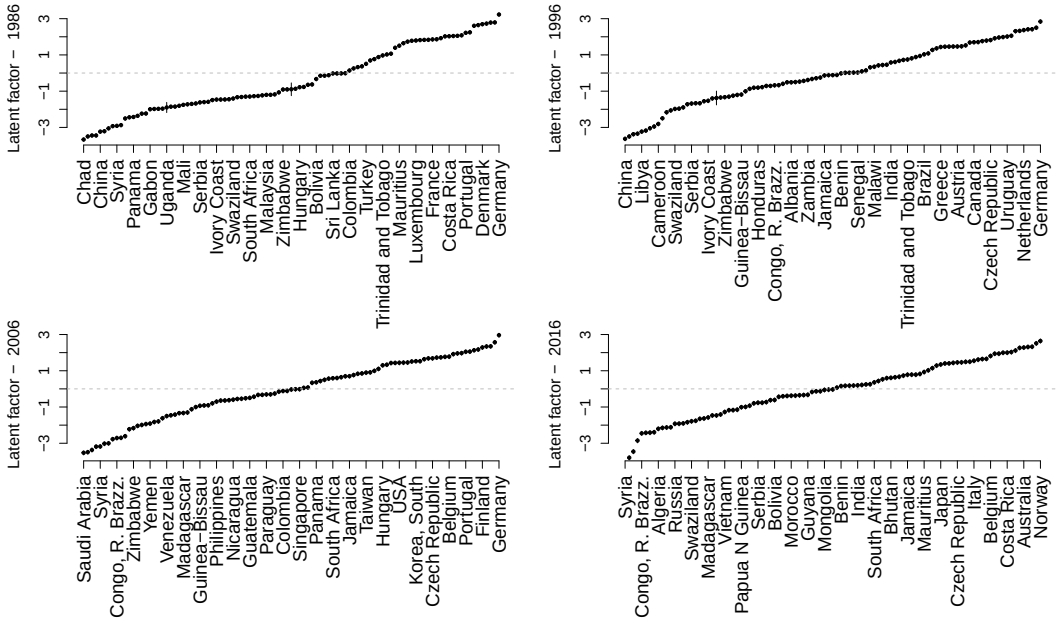

Figure 2: Posterior summary of country-specific latent factors for 1986, 1996, 2006 and 2016: point represents the posterior mean and the full line represents the 95% credibility interval for each country. The gray dashed line represents the value 0.

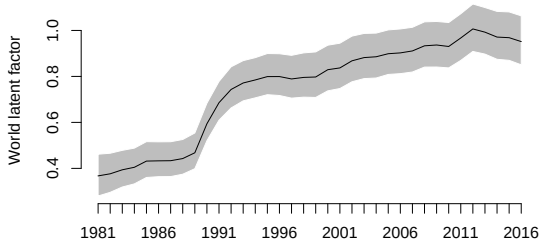

Figure 3: Posterior summary of world latent factor from 1981 to 2016. Full line represents the posterior mean and the region in gray represents the 95% credibility interval.

Economist Intelligence Unit's indicators of Rule of Law and Control of Corruption (EIU).

The dataset is composed by $K = 4$ annual time series (from 1995 to 2018) for each one of the $J = 111$ countries from the 7 regions. Since the four data sources are defined in the interval (0,1), we also applied the logistic transformation to the four data sources and fit the proposed model defined in Equation (2) to the transformed data.

Table 3 presents the posterior mean and the 95% credibility interval (in parenthesis) of the factor loadings associated to the world and contry latent factors. All the factor loadings are positive and signifficant. While the world factor has a less strong association with EIU, the country factors has a more strong association with it. Remember that the factor loadings for HERITAGE are fixed in one, due to the invariance of factor models under orthogonal transformations.

Table 3: Posterior distribution of the factor loadings: posterior mean and 95% credibility interval in parenthesis.

	HERI	PRS	IHS	EIU
β^w	1	1.69 (1.55,1.85)	2.01 (1.85,2.20)	1.00 (0.92,1.07)
β^c	1	0.73 (0.72,0.76)	0.92 (0.89,0.95)	1.29 (1.26,1.32)

Table 4 presents the posterior mean of the other parameters of the proposed model in Equation (2).

Table 4: Posterior mean of the other model parameters.

Param	σ_1^2	σ_2^2	σ_3^2	σ_4^2	γ^w	γ^c	λ^w	λ^c
Mean	0.361	0.129	0.222	0.053	0.969	0.966	0.004	0.011

Figure 4 shows a boxplot with the variance decomposition for each type of latent factor. It presents the distribution of the percentage of the variance of each data source, for all the years, separated by the type of factor. For HERI and EIU the world latent factors explain a similar percentage of the variance explained by the country latent factors. On the other hand, for the data sources PRS and IHS the world latent factor presents a higher percentage of the variance than the country latent factors.

Figure 5 presents the posterior mean and 95% credibility interval for the country-specific latent factors for the years 1997, 2001, 2005 and 2011. There are significant differences between the countries for all the years considered and some countries change position over the years.

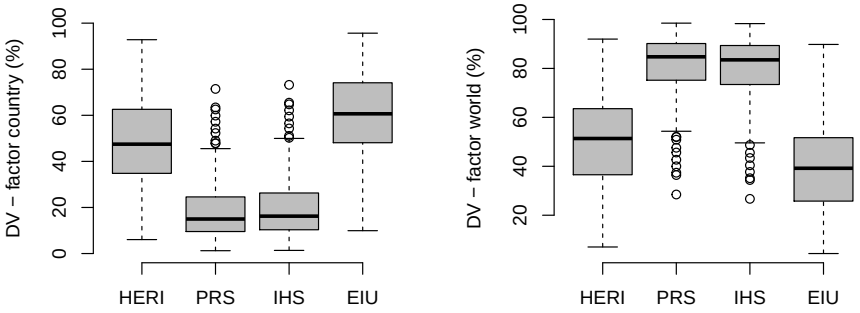

Figure 4: Boxplots with the variance decomposition for each data source divided by the world latent factor and the country-specific factor.

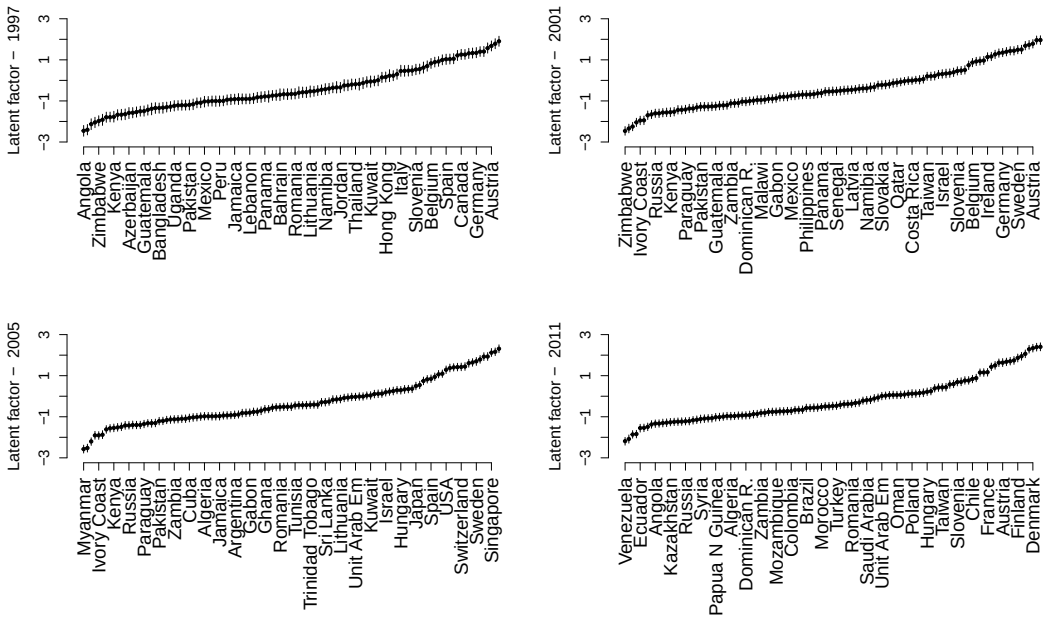

Figure 5: Posterior summary of country-specific latent factors for 1997, 2001, 2005 and 2011: point represents the posterior mean and the full line represents the 95% credibility interval for each country. The gray dashed line represents the value 0.

Finally, Figure 6 displays the posterior mean of the world latent factor

with the 95% credibility interval, from 1995 to 2018. The world factor seems to be significant for all the years with a decreasing trend until 2011 and after that the factor shows an increasing trend. It shows the tendency for economic institutions to improve after 2011.

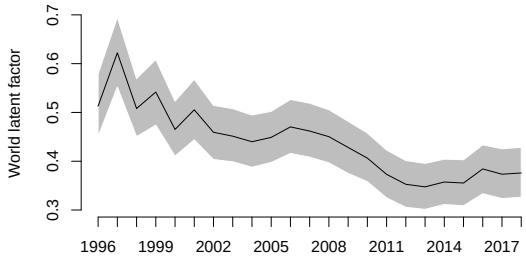

Figure 6: Posterior summary of world latent factor from 1981 to 2016. Full line represents the posterior mean and the region in gray represents the 95% credibility interval.

5 Conclusions

In this paper, we provide a clear link between the theoretical and empirical literature by creating aggregate indicators, also known as latent factors, for economic and democratic political institutions that are consistent with the theoretical literature and have a large geographical and temporal coverage. We present our own deffinition of democratic political institutions, and set out their theoretical attributes. For economic institutions, we highlight the theoretical dimensions already discussed extensively in the literature. We then clearly justify the choice of data we use to represent both political and economic institutions. Then we propose a methodology based on dynamic fator model which allows to combine multiple data sources in our indicators, minimizing problems regarding the comparability of scores between countries and over time that are present in existing aggregate indicators. The factor model create commom aggregators and country-specific aggregators, both varying over time. The inference is performed under the Bayesian paradigm and a sample from the posterior distribution is obtained using MCMC methods, in particular we used the softare JAGS. The missing data are also easily estimated, since they are naturally treated as parameters in JAGS.

Since, in this paper, the data sources are defined in the interval (0,1), the transformation of the data source was required here in order to not neglect the normal distribution assumed in the modeling. This fact encourage us to explore some extensions of this proposed model for the cases which the data sources are defined in a subinterval, without having to transform the data.

A Construction of democratic political institutions

We highlight three components of democracy that determine the distribution of constraints on rulers and citizens, and by extension, facilitate political exchange (see figure 7).

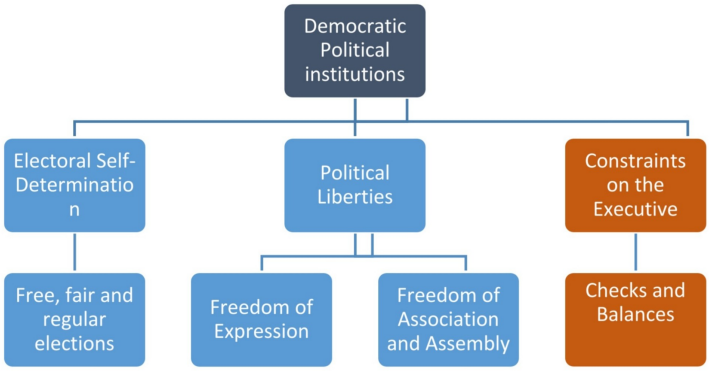

Figure 7: Components of democracy.

Firstly, the right of citizens to choose their political representatives through free and fair elections acts as a key formal mechanism through which citizens can influence political decisionmaking. Simultaneously, it acts as a critical constraint on chief executives and other elected representatives, who must then rely on a degree of popular support for election. As such, electoral self-determination facilitates political exchange by ensuring that elected representatives, both in the executive and the legislative branch, mediate, at least to some extent, the political preferences of voters.

The second vertical component of democracy is Political Liberties, which allow citizens to participate in public policy decision-making, and simultaneously constrain the power of rulers. Freedom of expression and freedom of

association and assembly complement the role of electoral self-determination by facilitating the development of rival political parties and the dissemination of information and ideas, which are key requirement for well-functioning, competitive elections.

The third component of democracy is horizontal constraints on the power of the executive through systems of checks and balances. In democracies, the executive constraints are usually imposed by the legislative and judiciary branches of government. These checks and balances play a crucial role in preventing arbitrary use of power by chief executives.

In the Freedom House's Freedom of the Press indicator, for each year since 1993, the index has given each country a score out of 100 (with 0 as the top score) based on the legal, political and economic constraints imposed on the media. Over the years, the indicators has made several modifications to its methodology to capture changes in the news and information environment without altering the comparability of data for a given country over the 35-year span of the index. For the period 1988 to 1992, the index only assigned each country a status of "free", "partly free" or "not free". To provide the option of extending the data back to this period, we assigned numerical scores based on the statuses. In cases where the status of the previous period is the same as the subsequent period for which we have a score, we assign the same score to the previous period. In cases where the status of the previous period is different to the subsequent period, we assign to the previous period the mean of the range of possible scores within its status. For the period 1981 to 1987, the index reported separate statuses for "print" and "broadcast" media. We therefore employed the same procedure as for the period 1988 to 1992, to estimate a score for "broadcast" and "print" media for each observation based on their statuses, and then calculated the mean of the two estimated scores to get an overall score.

In the CIRIGHTS Database, from which we used three out of their ten indicators – Freedom of Speech, Freedom of Assembly and Association, and Electoral Self-Determination. Each of the three indicators has a score of 0, 1 or 2, with 2 being the highest score. We double the score assigned to electoral self-determination so that it has the same combined weighting as Freedom of Speech and Press and Freedom of Assembly and Association, which both fall within our civil liberties component of democratic political institutions.

For the V-Dem Dataset, firstly, it uses a minimum of 5 country experts to code each country-year for every indicators and aggregates these various coding to arrive at point estimate and confidence interval for that rating.

Secondly, country experts submit ratings for each question over a long period of time, typically from 1900 to 2018, to improve intertemporal comparability. Thirdly, many cross-country experts code a number of other countries for the same set of questions to improve cross-country comparability. Fourthly, country experts are required to report a self-assessed level of certainty for their ratings.

To represent executive constraints we added together the scores of two composite indicators that well represent the concept – judicial constraints on the executive and legislative constraints on the executive , which both have scores between 0 and 1. The former is calculated by taking the point estimates from a Bayesian factor analysis model of five separate indicators: executive respects constitution, compliance with judiciary, compliance with high court, high court independence and lower court independence. The variable Legislative Constraints on the Executive is calculated by taking the point estimates from a Bayesian factor analysis model of four different indicators: legislature questions officials in practice, executive oversight, legislature investigates in practice and legislature opposition parties.

To represent electoral self-determination, we combined the scores of the variables share of population with suffrage (v2x_suffr) and the clean elections index (v2xel_frefair), which both have scores between 0 and 1. We multiplied the two variables together because we judged that in countries where a portion of the population is disenfranchised, the country's score for electoral democracy should be proportionately lower.

We added the scores for electoral self-determination and executive constraints to give a total score for the indicator. In table 5 we match all the indicators from each data source with the relevant theoretical component of democracy. In brackets are the scores available to each indicator as a percentage of the whole data source.

B Construction of economic institutions

NÃO SEI O QUE DEIXAR AQUI

Source	**Electoral Self-Determination**	**Political Liberties**	**Executive Constraints**
CIRIGHTS	Electoral Self-Determination (50%)	Freedom of Speech (25%) Freedom of Assembly and Association (25%)	
Freedom House		Freedom of the Press index (100%)	
Polity IV Project	Executive Recruitment (50%)		Executive Constraints (50%)
V-Dem	Share of Population with Suffrage multiplied by Clean Elections index (50%)		Judicial Constraints on the Executive (25%) Legislative Constraints on the Executive (25%)

Table 5: Disaggregated indicators, by source.

Source	**Property and Contract Rights**
PRS	Investment Profile (1/3): - Contract viability - Profits repatriation - Payment delays Law and Order (1/3): - Force and impartiality of legal system - Observance of the law Corruption in public organs (1/3)
Heritage Foundation	Property Rights (1)
Economist Intelligence Unit	Effectiveness of the Rule of Law (4/5) - Protection of private property - Protection of intellectual property - Confiscation/ misappropriation - Speed of judicial processes - Guarantees of contract enforcement - Legitimacy of judicial processes - Violent and organized crime Corruption (1/5)
Global Insight	Rule of Law (1) - Judicial independence - Crime

Table 6: Título Tabela 3