ABSTRACT

Title of dissertation: FOUR ESSAYS IN THE MEASUREMENT OF GOVERNANCE INSTITUTIONS

David Givens, Doctor of Philosophy, 2010

Dissertation directed by: Professor Peter Murrell
Department of Economics

This dissertation produces a new set of orthogonal governance measures based on expert assessment data. Chapter 1 constructs the measures using a factor model. Chapter 2 applies the measures to study comparative economic development. Chapter 3 conducts a number of robustness checks on results from the first two chapters. Chapter 4 uses Monte Carlo experiments to assess potential inaccuracy in my governance measures caused by the application of the maximum-likelihood estimator to polytomous data.

FOUR ESSAYS IN THE MEASUREMENT OF GOVERNANCE INSTITUTIONS

by

David Michael Givens

Dissertation submitted to the Faculty of the Graduate School of the University of Maryland, College Park in partial fulfillment of the requirements for the degree of Doctor of Philosophy
2010

Advisory Committee:
Professor Peter Murrell, Chair/Advisor
Professor John Chao
Professor Anton Korinek
Professor Boragan Aruoba
Professor Gregory Hancock, Dean's Representative

Acknowledgements

This work would not have been possible without the insights, dedication and encouragement of my advisors. I will always be thankful for their patience, their honest and constructive critiques, and their eagerness to see me succeed. I owe a special debt of gratitude to my committee chair, Dr. Peter Murrell, who first suggested studying the way governance institutions might best be measured. His faith in the wider importance of my research was a recurring source of strength whenever progress seemed too slow to bear.

I also thank my parents, my sister and all my dear friends for the encouragement, love and laughter they brought me throughout the six years of my graduate studies.

My most profound thank-you I reserve for my wife, Stacey. She has been my closest confidant, most ardent advocate, and most patient listener. She has sacrificed mightily to enable my advancement. She alone could show me the magnitude of my accomplishments, and she more than anyone helped me to see beyond the obstacles that graduate studies brought my way. For this and for so much more, I will be eternally grateful to her.

Contents

List of Tables

List of Figures

Introduction

Writers dating back to at least Adam Smith have advanced the importance of good governance for investment, innovation and growth. More recently, empirical work has quantified the economic benefit of good governance, using various governance proxies.[1] A limitation of these analyses is that a country's governance is typically represented unidimensionally.[2] This dissertation constructs a multi-dimensional assessment of governance quality and estimates each dimension's importance to long-run economic performance.

The economic contributions of good governance clearly come from a variety of different directions. Casual observation suggests that countries can excel in some governance areas (like regulation) even as they lag in others (like civil liberties). Relying on a single, catch-all measure of governance quality makes it impossible to estimate the marginal economic contributions of different categories of governance and may obscure a country's underlying strengths and weaknesses.

One solution is to replace the catch-all measure with a vector of measures that each capture a different aspect of governance. However, this approach creates several challenges. First, all governance measures tend to be highly correlated. In a regression setting, this poses the problem of multicollinearity. A second challenge is variable

[1] Over the last decade, the empirical study of governance and economic performance has taken off. Part of the reason for this is the novel application of expert assessments from political risk consulting companies. These perceptions-based variables are often available for a broad cross section of countries; they touch more directly than previous measures on the very governance concepts thought vital to economic decisions; and, they have proven capable of explaining much of the otherwise unaccounted-for cross-country variation in income.

[2] See e.g. - Mauro (QJE, 1995), Knack and Keefer (1995), Hall and Jones (QJE 1999), Acemoglu, Johnson and Robinson (AER, 2001), Sachs (2003), Easterly and Levine (JME, 2003), Rodrik et al. (JEG, 2004)

selection: how does one choose the best subset of governance variables from the many that are available? A third challenge is signal extraction: how do we control for extraneous information (e.g. - measurement error, biases, idiosyncratic measurement methodology, etc.) in governance variables - especially perceptions-based ones?

This dissertation addresses the above concerns by using a factor model to construct four new orthogonal indices of governance quality. In Chapter 1, I estimate the factor model using a diverse dataset of 45 governance-related variables (primarily expert assessments) from twelve different data sources. I interpret the four factors, discuss their robustness to rotation and method of extraction, and speculate on the potential biases created by the large amount of missing data. I also compare and contrast my governance factors with another popular set of measures produced using the same expert assessment variables - the Worldwide Governance Indicators (WGI) of Kaufmann et al (1999a).

Chapter 2 applies the factor score predictions developed in the first chapter to the study of comparative development. Using instrumental variables regressions, I estimate the causal impacts of each area of governance on per capita income. Across a wide variety of specifications, I find consistently that two of my four factors (market infrastructure and civil liberties) emerge as highly statistically and economically significant. Consistent with some recent work (Easterly and Levine, 2003; Rodrik et al., 2004) I find that controlling for the quality of governance, neither latitude nor trade share of GDP has any direct effect on income. Unlike earlier work, however, I am able to decompose the overall contribution of governance into distinct components.

Chapter 3 performs a robustness check - a monte-carlo-based examination of factor score prediction using polytomous data. Many of the 45 variables used to estimate the factor model parameters in Chapter 1 are polytomous in nature (e.g. - "rate country j on a scale of 1 to 7 in the independence of the judiciary"). In addition, many have distributions with distinctly non-normal shapes - including bimodal, skewed left, skewed right and U-shaped. Strictly speaking, the maximum likelihood (ML) factor model estimator used in Chapter 1 is valid only for continuous, normally distributed variables. The question addressed in Chapter 3 is the extent to which a country's predicted factor scores from Chapter 1 (which were based on the ML estimator) may deviate from that country's true factor scores due to the polytomous and non-normal characteristics of observed governance data.

Chapter 1

1 Introduction

This paper uses a factor model to construct four orthogonal categories of governance from a diverse dataset of 45 governance-related variables. The data are primarily expert assessments and were drawn from twelve different data sources.

In contrast to previous research, I use a unified statistical framework to determine the number of governance categories to create, the conceptual content of each category, and the relative importance each has in accounting for the observable data. I label my categories market infrastructure, civil liberties, downside governance risk, and order.[3]

Because my measures are orthogonal, they can be used jointly as regressors without the problem of multicollinearity. Orthogonality also makes the measures conceptually distinct from one another, and it helps them to accentuate strengths and weaknesses in a country more sharply than a set of highly correlated measures can.[4] Furthermore, as I demonstrate in the body of the paper, the factor model also addresses the variable selection and signal extraction challenges in appealing ways.

With respect to variable selection, country scores in my governance categories are constructed from linear combinations of all the variables in the dataset, alleviating

[3] While the precise wording we use to label our categories can be debated, we will show in the body of the paper that the substantive interpretation of these latent constructs is quite straightforward and unambiguous.

[4] Given that observable governance variables tend to be so highly intercorrelated, one legitimately might wonder how closely a set of orthogonal measures like ours can capture the concepts they purport to measure. However, we will show (see, e.g., Tables 6 and 7) that each of our measures is in fact highly positively correlated with popular observable benchmarks.

the need for the researcher to make a priori judgments about which observable variables are the most informative, or which observable variables represent which latent governance categories. All such relationships are estimated simultaneously within the factor model. Variables found to be more highly correlated with a governance category are assigned greater weight in its prediction.

With respect to signal extraction, the model decomposes the variation of each variable into governance-related and idiosyncratic components, thereby isolating signal from noise. Variables estimated to have smaller idiosyncratic components are given greater weight in the prediction of country governance scores.

My work is most closely related to the work of Kaufmann, Kraay and Zoido-Lobaton (e.g. - 1999b). Kaufmann et al.'s Worldwide Governance Indicators (WGI), now updated annually, are a popular set of governance measures - used widely in economic research and among policymakers.[5] In my analysis, I use largely the same dataset used by Kaufmann et al. for their WGI.[6]

There are three aspects of my work that are shared with Kaufmann et al. First, I produce aggregated governance measures - that is, measures constructed through lin-

[5]The WGI have lately become popular as a measure of governance in cross-country growth studies (Glaeser et al., 2004; Djankov et al. 2003; Easterly and Levine, 2003; Rodrik et al. 2004). A recent download statistic from SSRN provides some quantitative evidence of the measures' popularity: *Governance Matters VII* (the most recent version available) was the most downloaded article in SSRN's Economic Research Network (June 26 - August 25, 2008), with more than three times the number of downloads as the second most popular article for that period. The previous version, *Governance Matters VI,* was the second most frequently downloaded paper for the twelve months ending April 16, 2008, with twice as many downloads as the paper one notch below. The WGI have also gained influence among policymakers; for example, the U.S. Millenium Challenge Corporation uses the WGI as one criterion for distributing billions of dollars in foreign aid.

[6]Our paper uses a substantial subset of the perceptions-based governance variables collected by Kaufmann et al. for estimation of the 2005 version of their WGI. Our analysis excludes only the following types of variables used by Kaufmann et al: i.) variables available for only small, non-representative country samples, e.g. - Afrobarometer and Latinobarometro variables; and ii.) variables not publicly available outside the World Bank , e.g. - the World Bank's Country Policy and Institutional Assessments.

ear combinations of dozens of observable variables. Second, my conceptual approach is to treat governance as a latent variable, reflected noisily through the observable data. Finally, the statistical technique used for constructing the latent variables is a factor model.

However, I apply the factor model to governance data in a new manner, advancing the Kaufmann et al. results in several important ways. One way is my ability to force orthogonality on the governance measures that emerge. Another is my ability to quantify the relative importance of each governance measure in explaining the data. A third is my reliance on objective, statistical evidence to determine the optimal number of governance measures to extract from the data. Finally, a fourth is my ability to purge governance measures of data-provider effects.

The difference in my methodology that generates these advancements is straightforward. I estimate a single multi-factor model on the pooled set of 45 governance variables, whereas Kaufmann et al. divide the variables into six categories and estimate six independent one-factor models on the non-overlapping subsets of data.[7] This departure highlights another relative merit of my approach. It can be replicated easily by different analysts using different governance data because I do not rely on subjective categorizations of the observed data.

Each of my governance measures is defined according to what makes it unique from all the others. Governance, under this perspective, emerges not as a tangle of overlapping characteristics, but as an array of strongly differentiated capacities.[8]

[7] The Kaufmann et al. governance categories are: Rule of Law (RL), Control of Corruption (CC), Government Effectiveness (GE), Regulatory Quality (RQ), Voice and Accountability (VA), and Political Stability/No Violence (PSNV).

[8] For 2005, the minimum pairwise correlation among the six Kaufmann et al. governance cate-

The paper proceeds as follows. Part two provides an overview of the factor model that I use for extracting latent variables from governance data. Part three presents the data, motivates my model specification and presents estimation results. I interpret the governance measures that emerge from my model and contrast them with Kaufmann et al.'s WGI. Part four discusses potential biases due to missing data, and part five concludes.

2 The Factor Model and Governance

2.1 Motivation

The perspective of the factor model is that observable data reflect the systematic influence of latent variables called *factors*. An obvious reason to use the factor model is that governance capacities as typically envisioned are inherently *latent* - one cannot observe them directly. For instance, the growth literature routinely posits broad intangibles like "social infrastructure" (Hall and Jones, 1999), "protection against expropriation" (Acemoglu et al., 2001), "institutions" (Easterly and Levine, 2003).

To measure such concepts, one needs to infer the general from the particular - to translate observable attributes like the frequency with which contracts are enforced or the speed at which one can obtain a building permit into more fundamental qualities, like market infrastructure. A factor model is one of the most well-established methods for doing this (Anderson, 2003; Stewart, 1981).

To be sure, broad governance constructs aren't necessary or even appropriate for

gories is $\rho_{\min} = 0.7$; the maximum is $\rho^{\max} = 0.96$; the mean is $\overline{\rho} = 0.83$.

all governance-related questions. Where interest centers on a particular manifestation of governance (e.g. - freedom of the press), individual variables exist that measure specifically that institution. But for analysis of the overall governance environment facing consumers, businesses and investors, treating governance as latent is a sensible way forward.

It should also be noted that the use of latent variables is not new in economics. As mentioned, the *WGI* are latent variables estimated via a factor model. Other economic applications have included using factors as predictors of macro time series (Stock and Watson 1998, 2002; Cristadoro et al., 2005; Bernanke and Boivin, 2003; Boivin and Ng, 2005), and as proxies for systematic risk in asset pricing (e.g. - Roll and Ross, 1980; Chamberlain and Rothschild 1983; Connor and Korajczyk 1993).

2.2 Overview of the Model

At its core, the factor model is a tool for decomposing the intercorrelation of p random variables into systematic and idiosyncratic sources (Rao, 1955; Anderson, 2003). The systematic forces are called common factors, or simply, factors. An appealing characteristic of the model is that the number of factors required to summarize the systematic comovement of p variables is generally far fewer than p. From this perspective, the factor model is a data-reduction tool. Most importantly, the factor structure so derived suggests relationships both between the variables and between the subjects (i.e. - countries) that would otherwise be impossible to perceive from the raw data .

In my case, I show that just a few common factors can explain a great deal of the comovement in the 45 governance-related variables in the dataset. Furthermore, interpretations of the most important factors are straightforward and illuminating. I then go on to estimate each country's scores in each of the factors, enabling us to rank countries in the various dimensions of governance, and to investigate the economic role of good governance.

I begin with a brief overview of the factor model itself, including consideration of the rotation problem. Essentially, the rotation problem is the challenge of identifying parameters in a latent structure. I then proceed to issues of estimation and finally, to a common method for predicting the factor scores. For more detailed treatments of the factor model presented here, see Anderson and Rubin (1956), Harman (1976), Anderson (2003), or Wansbeek and Meijer (2000).

Let $X = [X_1, X_2, ...X_k, ...X_p]'$ represent a vector of normally distributed random variables. The factor analysis model decomposes this p-dimensional vector in the following way:

$$\underset{p\times 1}{X} = \underset{p\times m}{\Lambda}\underset{m\times 1}{f} + \underset{p\times 1}{u} \tag{1}$$

Matrix Λ is called the loadings matrix; its (i, j)-th entry is the covariance between the i-th variable in X and the j-th common factor. When λ_{ij} is large and positive, variable X_i is said to load heavily on factor f_j. Vector f is the vector of $m < p$ (unobserved) common factors, modeled as a random vector with density $f \sim N(\mathbf{0}, \Phi)$. Random disturbance vector u is *iid* with density $N(\mathbf{0}, \Psi)$, Ψ diagonal.

For clarity, I assume all elements of X have been centered and scaled such that $E(X_k) = 0$ and $sd(X_k) = 1, \forall k \in \{1, 2, ..., p\}$. It is assumed that all m factors are independent of all p disturbances (otherwise, disturbance u_k would not represent variable X_k's idiosyncratic variation).

Given equation (1) and the stochastic assumptions stated above, one can write Σ, the VC matrix of X, as follows:

$$\Sigma = \Lambda\Phi\Lambda' + \Psi \tag{2}$$

Equation (2) is the fundamental hypothesis of the factor model. The factor model posits that Σ can be decomposed into the sum of a symmetric, positive definite matrix of rank $m < p$ (i.e. - $\Lambda\Phi\Lambda'$), and a diagonal, positive definite matrix of rank p (i.e. - Ψ). This hypothesis is testable; i.e. - a given $p \times p$ population covariance matrix Σ may or may not be decomposable in this way for a chosen value of m (Anderson and Rubin, 1956; Lawley and Maxwell, 1971; Browne, 1969).[9] Anderson and Rubin (1956) provide some conditions on Σ such that a solution to (2) exists.

Matrix $\Lambda\Phi\Lambda'$ represents the variation in X that is due to factors common to all elements of X, while matrix Ψ captures the variation in each $X_k \in \{X_1, ..., X_p\}$ that is idiosyncratic.

2.3 Indeterminacy, rotation and simple structure

The model in (1) and (2) with no further assumptions is underidentified. Assuming that a solution to (2) exists for the chosen value of m, matrices $\Lambda\Phi\Lambda'$ and Ψ are iden-

[9] Lawley and Maxwell (1971) construct a simple example of three highly correlated variables, and one factor. Given the values they specify for the off-diagonal terms of $\Sigma = V(X)$, the factor model implies at least one $\psi_{ii}^2 < 0$, an obvious violation of the spirit of the factor model, since ψ_{ii}^2 represent variances.

tified uniquely; but to identify Λ, Φ individually, one needs to impose an additional m^2 independent conditions on them (Ledermann, 1937; Anderson and Rubin, 1956; Lawley and Maxwell, 1971; Anderson, 2003; Hayashi and Marcoulides, 2006).

In this paper, I make headway toward identification by imposing $\Phi = I$, so that the m factors are required to be orthogonal. Because my goal is to construct governance measures that are sharply delineated from one another conceptually, orthogonality is a natural constraint.[10]

When $\Phi = I$, (2) reduces to $\Sigma = \Lambda\Lambda' + \Psi$, and the number of additional restrictions needed for identification falls to $\frac{1}{2}m(m-1)$. A common tactic is then to require $\Lambda'\Psi^{-1}\Lambda$ diagonal (Lawley, 1940). This additional requirement just identifies the parameters and produces what is referred to in the literature as the unrotated solution. I designate the unrotated solution with Λ_0. Imposing $\Phi = I$ and $\Lambda'\Psi^{-1}\Lambda$ diagonal in no way restricts the $p \times p$ rank-m matrix $\Lambda\Phi\Lambda'$, meaning one has not ruled out any solution to (2) through these conditions (Anderson and Rubin, 1956; Anderson, 2003).

In truth though, the $\Lambda'\Psi^{-1}\Lambda$ restriction - while common in practice - is an arbitrary technique for gaining identification.[11] It selects but one of an infinity of admissible solutions to (2). In particular, it can be shown that post-multiplying Λ_0 by any $m \times m$ nonsingular and orthogonal matrix T produces a rotated solution, $\widetilde{\Lambda} = \Lambda_0 T$, that will

[10] Neither the identities of our four governance factors, nor the estimated factor scores for the countries, nor the growth regression results shown in section 4 are changed significantly by moving to an oblique rotation. So the practical consequences of this restriction turn out to be minimal.

[11] Another tactic for identifying Λ, Φ is the imposition of zero-loading constraints on individual elements, λ_{ij}, of Λ. For example, *a priori* theory may suggest to the analyst that variable X_k should be unaffiliated with factor f_j, implying the restriction $\lambda_{kj} = 0$ (Anderson, 2003). We do not pursue identification by zero-loading restrictions here because the point of our analysis is to impose as little *a priori* sorting to the data as possible.

generate observable moments of the data identical to those generated by Λ_0.[12]

Matrix $\widetilde{\Lambda} = \Lambda_0 T$ is called an orthogonal rotation of Λ_0.[13] The question is: on what basis should one select one rotation over another? The answer provided by the factor model literature is "simplicity". That is, some rotations of Λ generate factor structures that are simpler to interpret than others; since ultimately one is interested in solutions one can interpret, only the simplest, most interpretable rotations should be selected for analysis.

Motivated by this identification philosophy, Thurstone (1935, 1947) codified rules of simple structure - i.e., descriptive guidelines for extracting a more easily interpretable Λ. Literally, these rules seek a loadings matrix with a small number of large, positive loadings, and a large number of small, near-zero loadings (Harman, 1976; Lawley and Maxwell, 1971). Conceptually, Thurstone's rules aim at two objectives. Objective one is to clearly define and delineate factors from one another (by associating each factor with a small subset of the variables in X, and by forcing that subset to differ sufficiently from the subsets associated with the other factors). Objective two is to simplify the factorial characterization of each variable in the dataset (by minimizing the number of factors with which each variable is correlated). Objective one is sometimes called *factor simplicity,* and objective two is sometimes called *variable*

[12] Take any $m \times m$ orthgonal matrix T and construct $\widetilde{\Lambda} = \Lambda_0 T$, yielding $\widetilde{\Lambda}\widetilde{\Lambda}' = (\Lambda_0 T)(\Lambda_0 T)' = \Lambda_0 T T' \Lambda_0' = \Lambda_0 \Lambda_0'$, since $T' = T^{-1}$for T orthogonal. The observable data does not distinguish between a model with Λ_0 and a model with $\widetilde{\Lambda}$. See Anderson (2003), p. 571-573 for a thorough summary of the rotation problem in factor models.

[13] Rotations can also be *oblique*, meaning they allow for correlated factors. In an oblique rotation, fewer constraints are imposed on the rotation matrix T than in an orthogonal rotation. In particular, any $m \times m$ nonsingular matrix (not just the orthogonal ones) may constitute an oblique rotation matrix. For any hypothesized model (Λ, Φ, Ψ, f), the obliquely rotated model is $(\Lambda T, T^{-1}\Phi(T^{-1})', \Psi, T^{-1}f)$.

simplicity.

Beginning in the 1950s, psychometricians and statisticians started to operationalize Thurstone's simple structure guidelines. This was accomplished by translating them into mathematical objective functions that score the simplicity of any rotated loadings matrix $\widetilde{\Lambda} = \Lambda_0 T$. Simple structure is then achieved by finding the unique rotation, $\widetilde{\Lambda}^* = \Lambda_0 T^*$, that maximizes (or more often, minimizes) the chosen objective function. The optimization approach to rotation has provided an objective alternative for what used to be an arduous and subjective procedure for selecting one rotation over all others (Browne, 2001).

There is, however, no universally accepted mathematical measure of the simplicity of a loadings matrix. Thurstone's guidelines, while intuitive, are too vague to pin down a single objective function embodying all the characteristics of simple structure. In response, the literature has formulated dozens of rotations over the years (Harman, 1976; Brown, 2001). In fact, though, nearly every rotation pursues variable simplicity, factor simplicity, or some weighted average of the two objectives.

Two of the earliest, and still most commonly used orthogonal rotations (I limit discussion to orthogonal rotations for the reasons stated above) are called the *varimax* rotation and the *quartimax* rotation (Harman, 1976; Browne, 2001; Bernaards and Jennrich, 2003).[14] The varimax rotation pursues factor simplicity (objective one), by maximizing the variation in squared loadings within each *column* of Λ. The goal for each column is a few ones, and many zeros. The quartimax rotation pursues variable

[14]For example, the widely used orthomax family of orthogonal rotations, - which includes special cases parsimax and equamax - are defined as weighted averages of the quartimax and varimax criterion functions (Harman, 1976). One rotational criterion that is not a linear combination of quartimax and varimax is the orthogonal minimum entropy criterion, due to McCammon (1966).

simplicity (objective two), by maximizing the variation in squared loadings within each *row* of Λ. The goal for each row is a few ones and many zeros.

Let $\Lambda_v^* = \Lambda_0 T_v^*$ and $\Lambda_q^* = \Lambda_0 T_v^*$ be the varimax-rotated[15] and quartimax-rotated[16] versions of Λ_0, respectively. Under certain conditions (that is, for certain Λ_0), the varimax and the quartimax criteria push in the same direction - that is, $\Lambda^* = \Lambda_v^* = \Lambda_q^*$. In general, however, $\Lambda_v^* \neq \Lambda_q^*$ because in general, the objectives described in i.) and ii.) are not coincident. One can understand this better by considering a few examples.

First, consider a Λ_0 for which $\Lambda_v^* \neq \Lambda_q^*$. Let $\lambda_{i1} \approx 1, \forall i \in \{1, 2, ..., p\}$ and $\lambda_{ij} \approx 0, \forall i \in \{1, 2, ..., p\}, \forall j \in \{2, 3, ..., m\}$. In other words, first-column entries of Λ_0 are all near one, while entries in all other columns are near zero. Under this Λ_0, f_1 is often called a "general factor" because it is the most influential factor for all variables in the dataset (i.e. - all variables load most heavily on f_1). A general-factor configuration is consistent with maximization of the quartimax criterion (since a general-factor

[15] The maximand for the varimax rotation, developed by Kaiser, is (Harman, 1976 - p. 290):

$$g_v\left(\widetilde{\Lambda}\right) = \sum_{j=1}^{m} s_j^2$$

where, $s_j^2 \equiv \frac{1}{p}\sum_{i=1}^{p} \widetilde{\lambda}_{ij}^4 - \frac{1}{p^2}\left(\sum_{i=1}^{p} \widetilde{\lambda}_{ij}^2\right)^2$, and where $\widetilde{\lambda}_{ij}$ is the (i, j)-th element of rotated loading matrix $\widetilde{\Lambda} \equiv \Lambda_0 T$.

[16] Various maximands have been proposed for quartimax, but the simplest turns out to be:

$$g_q\left(\widetilde{\Lambda}\right) = \sum_{i=1}^{p}\sum_{j=1}^{m} \widetilde{\lambda}_{ij}^4$$

where again $\widetilde{\lambda}_{ij}$ is the (i, j)-th element of rotated loading matrix $\widetilde{\Lambda} \equiv \Lambda_0 T$ (Harman, 1976 - p. 282).

The quartimax criterion acheives its parsimony objective because a transformation, T^*, that maximizes the objective function above turns out to minimize $2\sum_{i=1}^{p}\sum_{j<k=1}^{m} \widetilde{\lambda}_{ij}^2 \widetilde{\lambda}_{ik}^2$, the sum of within-variable (within-row) products of squared loadings. Minimizing the sume of these products ($\widetilde{\lambda}_{ij}^2 \widetilde{\lambda}_{ik}^2$) within the ith row is equivalent to forcing many of variable i's loadings lower towards zero and i's remaining loadings upwards towards one - i.e., maximizing the within-row variation in squared loadings (variable i's communality, $\Sigma_{j=1}^{m}\lambda_{ij}^2$, is held constant under rotation, meaning a decrease in some λ_{ij}^2 requires an increase in some other λ_{ik}^2).

configuration characterizes each variable primarily in terms of a single factor) but is inconsistent with maximization of the varimax criterion: factors $2, ..., m$ tend not to be strongly differentiated from one another. In general, under the varimax rotation a subset of loadings in column one of Λ_0 will go down, and a subset of loadings in columns $2, ..., m$ (a different subset for each column) will go up.[17] This is a crude characterization, but precise enough for my purposes.

Next, consider a Λ_0 for which $\Lambda_v^* = \Lambda_q^*$. Suppose that Λ_0 admits perfect simple structure. Perfect simple structure means that each row of Λ contains exactly one non-zero entry. Bernaards and Jennrich (2003) show that if there exists an orthogonal rotation T_{pss} that can rotate Λ_0 to a perfect simple structure, then $\Lambda^* = \Lambda_0 T_{pss}$ will be the unique argmax of both the quartimax and varimax criteria. Put simply, if perfect simple structure exists, then both quartimax and varimax rotations will find it.[18]

Perfect simple structure represents the rare coincidence of *factor simplicity* and *variable simplicity*; factors can be cleanly characterized by mutually exclusive subsets of variables in X, and each variable is indicative of exactly one factor. In general, however - and especially in datasets like ours where every variable is highly positively correlated with almost every other variable - one is unlikely to encounter a factor structure that admits orthogonal rotation to perfect simple structure. The reason is straightforward. Under perfect simple structure, only a small fraction of the total

[17] For this reason, varimax is not an optimal rotation when one suspects a dataset of having a single meaningful factor (Lawley and Maxwell, 1971).

[18] Perfect simple structure isn't the only set of sufficient conditions guaranteeing that $\Lambda_v^* = \Lambda_q^*$. An alternative set of sufficient conditions is offered by Jennrich (*Psychometrika*, 2004a). They are slightly less intuitive and hinge on a characteristic of Λ_q^*. Whenever the sum of squared loadings in each column of Λ_q^* is the same for *all* columns in Λ_q^*, then $\Lambda_v^* = \Lambda_q^*$.

number of pairwise combinations of variables in X are explained through a common factor. This means the model must attribute most of the observed off-diagonal values in $V(X)$ to sampling variation, which becomes less and less tenable as the values become large.

As further illustration of the effects of varimax vs. quartimax, consider the first factor I extract from the governance data, f_1. Under the quartimax rotation of Λ_0, f_1 emerges as a general factor. This means that nearly all 45 variables in the dataset load heavily on it - including variables covering such diverse topics as civil liberties, regulatory and bureaucratic quality, and political violence. However, under a varimax rotation of Λ_0, the loadings on f_1 for civil liberties variables fall relative to the loadings on f_1 for regulatory and bureaucratic quality. Under varimax, the identity of f_1 becomes more clearly delineated, and factors 2 through 4 assume greater relative importance in explaining $V(X)$.

What has not been mentioned so far is the fact that rotating Λ can alter the definition of factors. These definitional changes may be drastic or they may be quite subtle. In the data, it turns out that the four factors of governance are nearly equivalently defined in the quartimax and varimax rotations (and in a number of other rotations tried for robustness but not presented in this paper). For reasons outlined below, however, I settle ultimately on varimax-rotated factors as my governance measures.

2.4 Estimating the model parameters

Assuming $\Lambda\Psi^{-1}\Lambda'$ diagonal and $\Phi = I$ just identifies the parameters in (1) and (2). Given the assumption of normality, one can then estimate the model (i.e., the $pm + \frac{1}{2}m\left(m+1\right)+p$ unique parameters in Λ, Φ, Ψ) by maximum likelihood. One maximizes the following (simplified) log-likelihood equation (Anderson, 2003; Wansbeek and Meijer, 2000):

$$\ln L = -\ln\left|\Sigma\right| - tr\left(\Sigma^{-1}S\right) \tag{3}$$

subject to constraints $\Lambda\Psi^{-1}\Lambda'$ diagonal and $\Phi = I$, where $\Sigma \equiv \Lambda\Phi\Lambda' + \Psi$, and $\underset{p\times p}{S} = \frac{1}{N}X'X$. The resulting estimate of Λ will correspond to the unrotated solution, Λ_0. One is then free to impose rotations on Λ_0.

The model presented in (1)-(2) assumes that the number of factors, m, is known. Of course, in practice one has to estimate m. Call the estimate $\widehat{m}$. Common tactics for picking $\widehat{m}$ include a likelihood-ratio test, information criteria (e.g. - Schwarz, 1978; Akaike, 1987) parallel analysis (Horn, 1965), the scree test (Cattell, 1966) and the eigenvalue $>$ 1 rule (Kaiser, 1960). All methods have the same objective of pinpointing the *minimum* number of common factors sufficient to replicate the observed $V\left(X\right)$. Monte carlo studies (e.g. - Cattell and Vogelmann, 1977; Hakstian et al., 1982; Zwick and Velicer, 1982 & 1986; Thompson, 2004) have documented that the performance of any particular rule can vary significantly with the factor structure, number of variables and number of factors in the data-generating process. Since different methods can produce different $\widehat{m}$ for a given dataset, it is therefore prudent to examine the robustness of factor identities by raising and lowering $\widehat{m}$.

2.5 Predicting the factor scores

The preceding section does not address how to estimate the factor scores themselves - i.e., the f's. Of course, since f is a random variate, one cannot estimate it *per se*. Put simply, even if one knew population values of Λ, Φ, Ψ, there would still be no way of separately identifying the levels of the u's and the f's in (1).

The approach I take is to predict f using its conditional mean, given data vector X. Such a predictor is called the regression predictor and is due to Thomson (1951). Anderson (2003) and Lawley and Maxwell (1971) provide full derivations of the regression predictor. Here I only reproduce the final result, noting that it follows directly from the joint distribution of f and X implied by (1), (2) and accompanying stochastic assumptions.

The $m \times 1$ population conditional mean vector of f, given observed data X can be written:

$$E[f|X] = \Phi\Lambda'\Sigma^{-1}X \tag{4}$$

In the orthogonal case ($\Phi = I$), the predictor simplifies to $\Lambda'\Sigma^{-1}X$. It will be helpful at this point to present the population conditional variance formula as well:

$$V[f|X] = \Phi - \Phi\Lambda'\Sigma^{-1}\Lambda\Phi \tag{5}$$

which in the orthogonal factors case simplifies to $I - \Lambda'\Sigma^{-1}\Lambda$.

As the term implies, the regression predictor conditions on an optimally weighted linear combination of the X's when predicting f. The intuition is most easily appreciated in the $m = 1$ case (the case of a *single* factor), where the matrix equation for $E[f|X]$ can be written in scalar terms as:

$$\underset{1\times 1}{E[f|X]} = c\sum_{k=1}^{p}\frac{\lambda_k}{\psi_{kk}}X_k \tag{6}$$

where,

$$c = \frac{1}{1+\sum_{k=1}^{p}\frac{\lambda_k^2}{\psi_{kk}}} \tag{7}$$

and where $\psi_{kk} \equiv V(u_k)$ is the kth diagonal element of matrix Ψ. One can regard the X's as noisy signals of the underlying f's. The scalar expressions in (6) and (7) then clearly demonstrate the factor model's desirable features as a signal extraction technique. Variables with greater sensitivity to the factor f (as measured by $\lambda_k = Cov(X_k, f)$) and/or smaller measurement error variance (as measured by ψ_{kk}) are weighted most heavily.

Since in practice the parameters Λ, Φ, Ψ are unknown, their ML estimates are used. Designate the estimated conditional mean and variance of f as:

$$\widehat{E}[f|X] = \widehat{\Phi}\widehat{\Lambda}'\widehat{\Sigma}^{-1}X \tag{8}$$

$$\widehat{V}[f|X] = \widehat{\Phi} - \widehat{\Phi}\widehat{\Lambda}'\widehat{\Sigma}^{-1}\widehat{\Lambda}\widehat{\Phi} \tag{9}$$

Finally, I note that when the parameters in (6) and (7) are translated into their conceptual counterparts using the notation from Kaufmann et al. (e.g. - 1999a, 1999b), the resulting expression is precisely the governance predictor proposed by those authors.[19] This equivalence demonstrates that the unobserved components model presented by Kaufmann et al. is identical to the random factor model presented here, with $m = 1$. In addition, because the Kaufmann et al. framework always restricts $m = 1$, the question of rotational indeterminacy does not arise. This does not mean, of course, that Kaufmann et al. have solved the underlying identification problem posed by rotation. Rather, they have implicitly imposed identification conditions at an earlier stage, through an *a priori* categorization of the data.

3 Results

In this section I present and interpret ML estimates of Λ, the loadings matrices embodying the latent structure of governance data. I discuss at length the interpretation of each governance factor and compare their varimax and quartimax-rotated versions. I then argue that, despite the robustness of factor identities to rotation, nevertheless a few characteristics of the varimax-rotated factors make them superior measures. My argument rests on both a comparison of the respective loading matrices, Λ_q and Λ_v, and on the the way the rotations differentially rank particular countries. Finally, I present a conceptual mapping between my new measures and the Kaufmann et al. WGI and show that in a very real sense, my measures are more fundamental. I begin with a discussion of my model specification and data.

[19] Kaufmann et al. predict governance for the jth country (g_j) as: $E[g_j|y_j] = \frac{1}{\left(1+\sum_{i=1}^{K}\sigma_i^{-2}\right)}\sum_{k=1}^{K}\sigma_k^{-2}\left(\frac{y_{jk}-\alpha_k}{\beta_k}\right)$

3.1 Model specification

I estimate an orthogonal factor model on $p = 45$ variables. I specify the model with $\widehat{m} = 5$ factors, as recommended by the Bayesian Information Criteria (BIC).[20] Extracting too many factors can reduce the precision of the loading estimates on the most important factors, whereas extracting too few factors risks obscuring important systematic forces of correlation in the data and clouding the identities of the factors that are retained (Thompson, 2004; Hayton et al., 2004). In the end, my decision to extract five factors rests on a mixture of statistical evidence, the durability of factor identities to values of $\widehat{m}$ greater than five, and the fact that factors higher than the fifth have minimal explanatory power and are simply uninterpretable.[21]

Two orthogonal rotations of Λ are presented and analyzed - the quartimax and varimax. The quartimax and varimax rotations were selected because each pursues exclusively one of the two fundamental notions of simple structure put forth by by

[20] The BIC actually recommends six factors, but we reject the model with $m = 6$ because it turns out to be a Heywood case. In a Heywood case, the ML estimate of at least one ψ_{kk} is a corner solution, i.e. - the constraint that $\psi_{kk} \geq 0, \forall k$ is binding. A Heywood case can result from i.) sampling variation (meaning that in the population, the offending ψ_{kk} is actually positive); ii.) non-uniqueness of the solution Λ, Φ, Ψ; or iii.) non-existence of an admissable solution. See Van Driel (1978), Anderson and Gerbing (1984), Gerbing and Anderson (1987), and Dillon, Kumar and Mulani (1987) for discussions. The next-lowest value of the BIC is reached with $m = 5$, which - combined with the fact that the sixth and higher factors are indecipherable - constitutes our grounds for retaining five factors.

[21] To be sure, the various number-of-factor rules do not all agree on m for our data. The Akaike Information Criteria (AIC) and likelihood-ratio statistic (LR) recommend twelve and thirteen factors, respectively. Horn's (1965) parrallel analysis (PA) and Kaiser's eigenvalue >1 rule (K) both seem to recommend 4-7 factors. Monte carlo studies have shown that the PA is one of the most accurate rules under a broad variety of factor structure DGP's (with a slight tendency to extract too many factors), while the K rule has a pervasive tendency to extract too many factors (Humphreys and Montanelli, 1975; Zwick and Velicer, 1982, 1986; Hayton et al., 2004; Horn, 1965). The LR test has been found to over and underextract, depending on the DGP (Hakstian et al., 1982;). It is difficult to extrapolate much from monte carlo results because so many of the conclusions depend on features of the DGP which in practice are unknowable (although the tendency for the K rule to overextract seems well documented). Perhaps the most important message from such studies is that one must consider many pieces of evidence before settling on $\widehat{m}$.

Thurstone (1935, 1947). If any pair of rotations is likely to generate differing interpretations of the factors, it is this pair.

Quartimax and varimax are both *orthogonal* rotations. I prefer to look at only orthogonal rotations because, on the conceptual level, orthogonality implies that the factors I uncover are basic and irreducible. If two factors were allowed to be correlated, this could indicate either a causal relationship between the two, or an omitted variable affecting both. In either case, I would want to dig deeper. Oblique rotations of Λ produce largely the same five factors of governance, but they can cause conceptual differences between factors to become less pronounced. Excessive intercorrelation constitutes one of my main critiques of the WGI, which are so highly correlated that they appear to be largely measuring the same latent forces.

3.2 Data

The dataset consists of 45 country-level variables assessing the quality of governance in various economic, political and legal contexts, in 2005. The data were taken from an online database of governance variables constructed by Kaufmann et al. and available at the World Bank's *Worldwide Governance Indicators* (WGI) website. Detailed documentation for each variable is taken directly from *Governance Matters V: Appendices* (Kaufmann et al., 2006). All variables represent subjective assessments - primarily by country experts at for-profit risk consulting firms (e.g. - Economist Intelligence Unit, Political Risk Services, Global Insight) but also by researchers at NGOs, universities and think tanks (e.g. - Freedom House, Amnesty International, Brown University, Heritage Foundation).

All variables are standardized by Kaufmann et al. such that realizations fall within the zero-one interval and such that higher scores correspond to better outcomes. This makes scores comparable across sources while preserving the ordinal properties of the original data.

While select variables in the dataset are observed for over 200 countries, the model parameters were estimated using the largest balanced dataset available, 73 countries.[22] I say more about missing data issues at the end of this Chapter.

3.3 Interpreting the Five Factors

I now interpret maximum likelihood estimates of Λ from the quartimax and varimax rotations of Λ_0. I refer to these loadings matrices as Λ_q and Λ_v, respectively. I extract five factors, and each factor has a straightforward definition derived from the concepts common to the variables most closely associated with it.

A word about variable names is in order. There are 45 closely related variables in the governance dataset, making it challenging to find short, distinctive and descriptive names for each. I have opted for the following convention. I label each variable with its Kaufmann et al. governance category, followed by an acronym indicating the data provider - e.g., $\text{corrupt}_{\text{GRS}}$ (Global Insight's Global Risk Service variable classified as

[22] Since the actual input required for estimating Λ, Φ, Ψ of a factor model is just a correlation matrix (and not the raw data, per se), it is theoretically possible to include all observations (complete and incomplete) by substituting the matrix of pairwise correlations (based on all available observations for all variables) in place of the complete-case correlation matrix that's based on the 73 observations of the balanced panel. We do not take this approach, however. Aside from its questionable statistical properties (different cells of the pairwise correlation matrix will be based on different numbers of observations), it turns out in our case to be impossible for another reason: the matrix of pairwise correlations using all available observations is not positive definite.

Control of Corruption by Kaufmann et al.), or $\text{voice}_{\text{EIU}}$ (Economist Intelligence Unit variable classified as Voice and Accountability).[23] The Kaufmann et al. categories are a way to communicate succinctly the putative content of a variable, but they in no way inhibit my model's determination of a variable's factorial content.

To interpret or define a factor, one looks to the variables with which that factor is highly correlated - i.e., the variables that load most heavily on it. To that end, for each factor, I list in descending order the ten variables that load most heavily on it under the given rotation. For example, looking at Table 1 below, one sees that in the quartimax rotation, variable $\text{goveffect}_{\text{BCRI}}$ loads more heavily on f_1 than any other variable in the dataset ($\lambda_k = 0.960$); variable $\text{corrupt}_{\text{BCRI}}$ has the second largest f_1 loading ($\lambda_k = 0.956$), etc. Examining the content of these two variables enables us to begin characterizing f_1 under the quartimax solution. These lists constitute the primary input used to define each factor. I create for each factor a label, based on the ideas common to its highest-loading variables. One can quibble with the precise wording of the labels, but the intent is simply to characterize succinctly the concepts at the core of each factor.

3.3.1 Factor one

I begin with the first factor, f_1. This factor is a broad metric of the overall legal and bureaucratic environment faced by firms and investors. It measures the state's

[23] Data provider acronyms are deciphered in the appendix. We retain the Kaufmann et al. governance categories as our variable descriptors even as we advocate against the *a priori* organizing principle used to develop them. While at first glance this may seem contradictory, the position taken in this paper is not that *a priori* governance categories make no sense at all (on the contrary - they can be quite intuitive), but that the approach behind them is open to certain shortcomings, laid out in the introduction.

tendency to treat all investors and entrepreneurs equally under the law, to uphold contracts, to discourage predation (including predation by its own agents), and to minimize disruptive policy shifts. To the extent that f_1 can be said to measure the rule of law, the focus is on the protection of certain economic rights (primarily property and contracting rights) and on the suppression of corruption.

Some common phrases in the documentation of variables that characterize f_1 include "enforceability of contracts", "property rights", "corruption among public officials", "nepotism", "intrusiveness of the country's bureaucracy", "quality of the country's bureaucracy", "whether the necessary business laws are in place", "enforced consistently", and "competence of public sector personnel".

MARKET INFRASTRUCTURE			
Quartimax		Varimax	
variable	λ_k	variable	λ_k
$\text{goveffect}_{\text{BCRI}}$	0.960	$\text{rulelaw}_{\text{QLM}}$	0.892
$\text{corrupt}_{\text{BCRI}}$	0.956	$\text{corrupt}_{\text{QLM}}$	0.887
$\text{regulation}_{\text{BCRI}}$	0.955	$\text{rulelaw}_{\text{HER}}$	0.811
$\text{rulelaw}_{\text{EIU}}$	0.953	$\text{corrupt}_{\text{GCS}}$	0.808
$\text{corrupt}_{\text{EIU}}$	0.946	$\text{rulelaw}_{\text{GCS}}$	0.803
$\text{rulelaw}_{\text{HER}}$	0.941	$\text{corrupt}_{\text{EIU}}$	0.792
$\text{corrupt}_{\text{QLM}}$	0.937	$\text{goveffect}_{\text{EIU}}$	0.774
$\text{rulelaw}_{\text{QLM}}$	0.936	$\text{goveffect}_{\text{GCS}}$	0.751
$\text{voice}_{\text{BCRI}}$	0.926	$\text{corrupt}_{\text{BCRI}}$	0.742
$\text{rulelaw}_{\text{BCRI}}$	0.919	$\text{corrupt}_{\text{PRS}}$	0.722

Table 1

The unifying theme in f_1 is the state's ability to reduce uncertainty and lubricate exchange by providing a legal, regulatory and judicial infrastructure for markets.

Based on this role, I refer to f_1 as simply *market infrastructure*. The name is intended to envelop the full scope of government intervention in markets - be it bureaucratic, legislative, tax or judicial.

The quartimax and varimax rotations of f_1 have much in common with each other. They share five variables in their respective lists of the ten most important, and the conceptual similarities of the variables on which the lists differ allow us to apply comfortably the label market infrastructure to both rotations.

Nevertheless, there are two ways in which the quartimax and varimax characterizations of market infrastructure differ. First,market infrastructure emerges as a *general factor* under the quartimax rotation, but not under varimax (to see this, compare column one from the complete $\widehat{\Lambda}_q$ and $\widehat{\Lambda}_v$ matrices in the appendix). Second, civil liberties play much less of a role in the varimax market infrastructure than in the quartimax version. These two differences are related. That is, varimax avoids creating a general factor precisely because it allocates civil liberties content away from market infrastructure and onto f_2.[24]

3.3.2 Factor two

The second factor is a measure of democracy and civil liberties.

[24] There is a third difference. The quartimax f_1 is dominated by data from Global Insight's Business Conditions and Risk Indicators (BCRI), whereas the varimax f_1 is not. Five of the quartimax f_1's top ten variables come from BCRI, but only one of the varimax's top ten comes from that source.This difference may be rooted in the varimax f_1's de-emphasis of civil liberties. BCRI variables tend to be more strongly correlated with civil liberties indicators in our datasets than QLM and GCS variables - which dominate the varimax f_1.

CIVIL LIBERTIES			
Quartimax		Varimax	
variable	λ_k	variable	λ_k
$\text{voice}_{\text{HUM}}$	0.695	$\text{voice}_{\text{FRH}}$	0.906
$\text{voice}_{\text{FRH}}$	0.670	$\text{voice}_{\text{HUM}}$	0.901
$\text{voice}_{\text{RSF}}$	0.553	$\text{voice}_{\text{RSF}}$	0.783
$\text{voice}_{\text{PRS}}$	0.419	$\text{voice}_{\text{EIU}}$	0.740
$\text{voice}_{\text{EIU}}$	0.379	$\text{voice}_{\text{PRS}}$	0.711
$\text{rulelaw}_{\text{HUM}}$	0.370	$\text{rulelaw}_{\text{HUM}}$	0.685
$\text{regulation}_{\text{HER}}$	0.281	$\text{voice}_{\text{BCRI}}$	0.622
$\text{voice}_{\text{BCRI}}$	0.234	$\text{regulation}_{\text{HER}}$	0.601
$\text{regulation}_{\text{EIU}}$	0.166	$\text{regulation}_{\text{EIU}}$	0.517
$\text{stability}_{\text{EIU}}$	0.152	$\text{regulation}_{\text{BCRI}}$	0.494

Table 2

Variables with the highest loadings on f_2 come from rights-oriented organizations like Freedom House, Reporters Without Borders and Amnesty International. These variables convey unequivocally their content in phrases such as: "restrictions on domestic and foreign travel", "imprisonments because of ethnicity...", "freedom of assembly", "protection from political terror", "press freedom index", "free and fair elections", and "right to freely organize in different political parties".[25]

The interpretation of f_2 is robust to rotation.[26]

[25] Other phrases characteristic of f_2 include: "free religious institutions", "fair electoral laws", "free from domination by the military", and "accountability of public officials".

[26] There is even greater across-rotation uniformity in the loading pattern for f_2 than for f_1. The quartimax and varimax rotations of f_2 share nine out of ten variables in their respective lists, and they rank these common variables in a very similar order.

3.3.3 Factors three and four

The factor I label downside governance risk emerges as f_3 under the varimax rotation and f_4 under the quartimax rotation.[27]

DOWNSIDE GOVERNANCE RISK			
Quartimax		**Varimax**	
variable	λ_k	variable	λ_k
$\text{regulation}_{\text{GRS}}$	0.558	$\text{goveffect}_{\text{GRS}}$	0.742
$\text{goveffect}_{\text{GRS}}$	0.448	$\text{regulation}_{\text{GRS}}$	0.740
$\text{rulelaw}_{\text{GRS}}$	0.418	$\text{rulelaw}_{\text{GRS}}$	0.709
$\text{stability}_{\text{GRS}}$	0.384	$\text{stability}_{\text{GRS}}$	0.630
$\text{corrupt}_{\text{GRS}}$	0.294	$\text{corrupt}_{\text{GRS}}$	0.615
$\text{regulation}_{\text{PRS}}$	0.288	$\text{regulation}_{\text{PRS}}$	0.550
$\text{voice}_{\text{PRS}}$	0.120	$\text{regulation}_{\text{BCRI}}$	0.460
$\text{regulation}_{\text{BCRI}}$	0.114	$\text{goveffect}_{\text{GCS}}$	0.420
$\text{regulation}_{\text{GCS}}$	0.090	$\text{regulation}_{\text{GCS}}$	0.415
$\text{voice}_{\text{GCS}}$	0.086	$\text{voice}_{\text{GCS}}$	0.405

Table 3

Unlike market infrastructure (f_1) or civil liberties (f_2), the most obvious common feature of variables associated with downside governance risk is their common data source (Global Insight's Global Risk Service (GRS)), not their common content. The GRS variables in the dataset all forecast the likelihood of a deterioration in governance sufficient to cause economic contraction of a specified severity. For example, $\text{stability}_{\text{GRS}}$ measures the likelihood that personnel problems within the government cause a 1% decline in the GDP growth rate over any 12-month period. The variables in Table 3 assess different areas of governance (e.g. - political stability, corruption,

[27] The reason for the reversal is that in our data the relative explanatory powers of f_3 and f_4 are nearly equivalent. Explanatory power of a factor is proxied by its corresponding eigenvalue.

regulatory burdens), but they share a focus on potential economic fallout of institutional decay. I hypothesize that it is this focus on economic fallout which generates systematic correlation among GRS variables that persists even after the effects of the other four factors have been controlled for.

Implicit in predictions made by GRS variables is a hypothesis concerning the relationship between elements of governance and their effects on macroeconomic performance. But that hypothesis is not spelled out in the variable descriptions (see Appendix). As a result, a detailed interpretation of these variables is problematic. For example, a 1% decline in GDP growth rate may require massive government turnover in one country but only a small disruption in others. As a result, cross-country variation in $\text{stability}_{\text{GRS}}$ may be measuring differences in the fragility of governance, differences in the sensitivity of output growth to governance shocks, or both.

Given the uncertainty surrounding the precise meaning of GRS variables, it seems prudent to infer only that they are broad measures of governance volatility. Countries with high scores are less likely to experience abrupt, negative governance shocks over the forecast horizon; countries with low scores are more likely to experience such shocks. For this reason, I characterize the factor dominated by GRS variables as downside governance risk.[28]

The next factor is a measure of political, social and ethnic turmoil, which I label order. Order emerges as f_4 under the varimax rotation and f_3 under the quartimax

[28] The emergence of a data-source specific factor is interesting. We've argued above in effect that, it's not that GRS country assessments are inconsistent with those of other data providers but rather that GRS simply measures a phenomenon (economic fallout) that's slightly unique. Strictly speaking, this claim is not verifiable statistically since the factor model can't tell us why certain groups of variables covary in the ways they do. In any case, this phenomenon appears to be common. For example, MIG also generates a factor unto itself (f_5).

rotation.

ORDER			
Quartimax		Varimax	
variable	λ_k	variable	λ_k
$\text{stability}_{\text{GCS}}$	0.649	$\text{stability}_{\text{BCRI}}$	0.785
$\text{stability}_{\text{BCRI}}$	0.621	$\text{stability}_{\text{GCS}}$	0.732
$\text{stability}_{\text{PRS}}$	0.505	$\text{stability}_{\text{PRS}}$	0.615
$\text{stability}_{\text{MIG}}$	0.497	$\text{stability}_{\text{HUM}}$	0.613
$\text{stability}_{\text{HUM}}$	0.428	$\text{stability}_{\text{MIG}}$	0.603
$\text{stability}_{\text{GRS}}$	0.296	$\text{stability}_{\text{GRS}}$	0.458
$\text{stability}_{\text{EIU}}$	0.252	$\text{stability}_{\text{EIU}}$	0.446
$\text{regulation}_{\text{MIG}}$	0.132	$\text{rulelaw}_{\text{BCRI}}$	0.345
$\text{rulelaw}_{\text{GRS}}$	0.124	$\text{rulelaw}_{\text{PRS}}$	0.337
$\text{rulelaw}_{\text{PRS}}$	0.123	$\text{rulelaw}_{\text{GRS}}$	0.317

Table 4

Characteristic phrases associated with order include: "cohesion of the government and governing party or parties", "racial and nationality tensions", "political unrest", "tribal conflict", "government coups", "armed opposition", "frequency of political killings", "foreign-supported insurgency", and "sustained terrorist threat".[29]

Broadly speaking, order captures the extent to which the political system minimizes threats to stability and social cohesion. A wide variety of threats from a wide variety of sources are addressed: military coups, foreign invasions, breakdowns of governing coalitions, civil unrest, violent popular demonstrations, violent ethnic clashes, and terrorism. Impacts on the government and on businesses and investors are both considered. Order is also robust to rotation: the quartimax and varimax versions

[29] Other characteristic phrases for the order include: "extremism", "arbitrary violence" by the state, and "frequency of torture".

share nine of ten top variables and rank these in a very similar order of importance.

Nothing in the variables that define order distinguishes between the tranquil civility of a democracy and the cowed silence of a police state. Order displays a preference for stability, however it may be accomplished. Stability may stem from cultural, geographic or demographic factors (e.g. - Mongolia or Samoa), or it may stem from a repressive political regime (e.g. - North Korea, Belarus).

To understand order's content more, I look at how it ranks countries. In the scatterplot below, each country's horizontal coordinate equals its order score under the quartimax rotation; a country's vertical coordinate is its order score under varimax. (The diagonal line is the 45-degree line.)

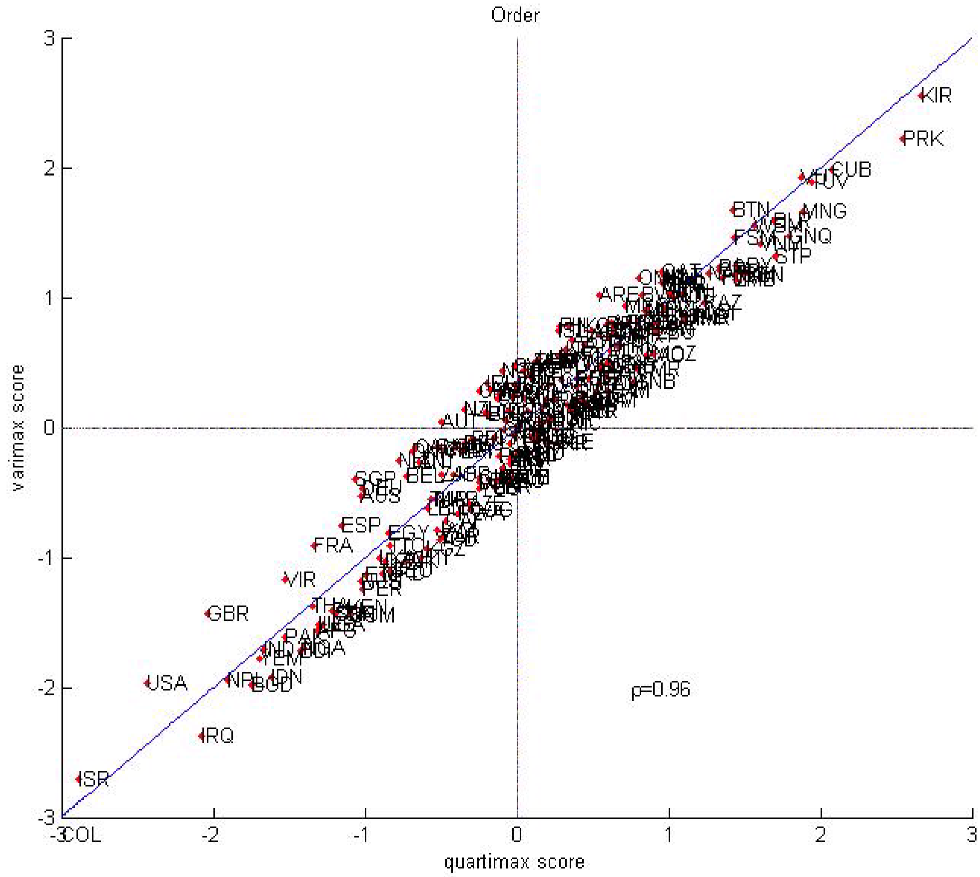

Figure 1

At $\rho = 0.96$, the correlation between varimax and quartimax order is nearly perfect. I will bring out some more subtle differences between the rotations later in the paper, but for now I focus exclusively on what their versions of order have in common.

Order presents a curious but ultimately explainable juxtaposition of countries. The highest-scoring countries contain a sizable number of authoritarian regimes. For example, Cuba (CUB) and North Korea (PRK) rank second and third, respectively under varimax. And countries such as Belarus (BLR), Turkmenistan (TKM), Bhutan

(BTN), Lybia (LBY), and Myanmar (MMR) all fall within the top quintile. Typically what places countries like these near the top is an above-average score in $\text{stability}_{\text{BCRI}}$ (risk of civil unrest and terrorism), combined with very poor scores in three human rights measures: $\text{voice}_{\text{BCRI}}$ (representativeness of political system), $\text{voice}_{\text{FRH}}$ (political rights, civil liberties, freedom of the press), and $\text{voice}_{\text{EIU}}$ (accountability of public officials, human rights, freedom of association).

At the opposite end of the spectrum, the very lowest-scoring countries contain a sizable number of large, diverse democracies. Colombia (COL) and Israel (ISR) score lowest and second-lowest, respectively under varimax. And the United States (USA), India (IND), Indonesia (IDN), Nigeria (NGA), the Philippines (PHL), the U.K. (GBR) and France (FRA) all score in the bottom quintile. Typically, what places countries like these near the bottom is below-average scores in three stability variables: $\text{stability}_{\text{BCRI}}$ (risk of civil unrest and terrorism), $\text{stability}_{\text{HUM}}$ (killings, disappearances, torture), and $\text{stability}_{\text{GCS}}$ (business costs of terrorist threat). In the case of the U.S. and U.K., high rankings in $\text{voice}_{\text{FRH}}$ (political rights, civil liberties, freedom of the press) and $\text{voice}_{\text{EIU}}$ (accountability of public officials, human rights, freedom of association) also worked against them.

What emerges from examination of the country rankings, variable definitions and loading patterns is a factor that - while not definitively anti-democratic - nevertheless tends to penalize the decentralization of political power. For all their potential virtues, the enshrinement of certain civil liberties in a country's political culture; the allocation of government power from the federal to the state and local levels; and a more open electoral system may also sew seeds of upheaval.

Order's other defining characteristic is the strong penalty it puts on perceived risks of civil unrest and/or terrorism. Since social instability can stem from a wide variety of causes - ethnic, religious, political, cultural, historical - a great diversity of countries can be found at any given level of order.

3.3.4 Factor five

I label the fifth factor MIG after Merchant International Group, the data provider from which f_5's most important variables are drawn. Its identity is robust to rotation, as Table 5 below shows.

MERCHANT INT'L. GROUP			
Quartimax		Varimax	
variable	λ_k	variable	λ_k
$\text{corrupt}_{\text{MIG}}$	0.459	$\text{corrupt}_{\text{MIG}}$	0.653
$\text{goveffect}_{\text{MIG}}$	0.422	$\text{goveffect}_{\text{MIG}}$	0.617
$\text{regulation}_{\text{MIG}}$	0.406	$\text{regulation}_{\text{MIG}}$	0.583
$\text{rulelaw}_{\text{MIG}}$	0.382	$\text{rulelaw}_{\text{MIG}}$	0.581
$\text{stability}_{\text{MIG}}$	0.213	$\text{corrupt}_{\text{PRS}}$	0.375
$\text{corrupt}_{\text{PRS}}$	0.164	$\text{stability}_{\text{MIG}}$	0.337
$\text{goveffect}_{\text{PRS}}$	0.110	$\text{goveffect}_{\text{PRS}}$	0.305
$\text{regulation}_{\text{HER}}$	0.081	$\text{voice}_{\text{GCS}}$	0.292
$\text{stability}_{\text{HUM}}$	0.077	$\text{rulelaw}_{\text{HER}}$	0.279
$\text{voice}_{\text{GCS}}$	0.077	$\text{regulation}_{\text{HER}}$	0.271

Table 5

Like downside governance risk, MIG is another data-source-specific factor. However, unlike downside governance risk, very little is known about what makes MIG's country assessments unique.[30] This factor accounts for a mere 8% of the explained variation in X under the varimax rotation and for less than 3% under the quartimax rotation. While I find the emergence of a second provider-specific factor interesting, its modest explanatory power combined with its inscrutability (these characteristics are no doubt related) mean I will have littler more to say about f_5.

3.3.5 Summing up

I have extracted five orthogonal factors from a body of governance data. The identities of all factors are robust to rotation. The first four factors are readily interpretable as governance concepts: *market infrastructure, civil liberties, downside governance risk* and *order*. Although each of the first four factors individually comprises a broad range of ideas, each is also starkly delineated from all the others. Taken together, they encompass most of the legal and political institutions commonly implied by the term "governance", including business regulation, property and contracting rights, judicial independence, bureaucratic efficiency, corruption, political stability, and civil liberties.

The fifth factor appears not to measure a recognizable feature of governance but is instead defined by a methodological or conceptual idiosyncrasy of data provider, MIG. Whatever its origin, the fifth factor is of negligible importance in explaining $V(X)$. The first four factors alone account for 92% of explained variation under the varimax rotation, and for over 97% under the quartimax rotation. For these reasons,

[30] World Bank documentation of MIG's Grey Area Dynamics (the MIG variables in our model) are brief and lack detail. MIG's website does provide more detail about their assessments, but nothing there suggests a unique analytical approach.

I focus exclusively on factors one through four henceforth.

3.4 Varimax: the preferred rotation

In this section, I argue that the varimax rotation of Λ_0 produces the most plausible and intuitive factor structure, and that therefore the varimax-rotated factor scores should serve as my governance indicators.

As mentioned earlier, simplicity is the traditional criterion for selecting a rotation. The idea is that one should select whichever rotation yields the most plausible factor interpretations. The complication in my case is that, for the most part, the broad outlines of each factor have been found to be robust to rotation.

On the other hand, the objectives of each rotation are manifestly different, and these differences, for example, have been shown to cause civil liberties concepts to play a much smaller role in the varimax version of market infrastructure than in the quartimax version. Therefore, even when multiple rotations produce broadly similar interpretations of the data, potential grounds for preferring a particular rotation may still exist.

Along these lines, I believe there are some compelling reasons to prefer varimax. First, I believe varimax's factor simplicity objective is the most appealing approach, given that a main purpose of my inquiry is to define the factors of governance. Factor simplicity is the ideal rotational objective for bringing out the unique essence of each factor. Although factors - linear combinations of the variables in X - can be made orthogonal in an infinity of ways, some ways are more intuitive than others.

Figure 2 below is a scatter plot of country scores in my market infrastructure factor. Each country's horizontal coordinate again equals its score under the quartimax rotation, and its vertical coordinate is its score under varimax.

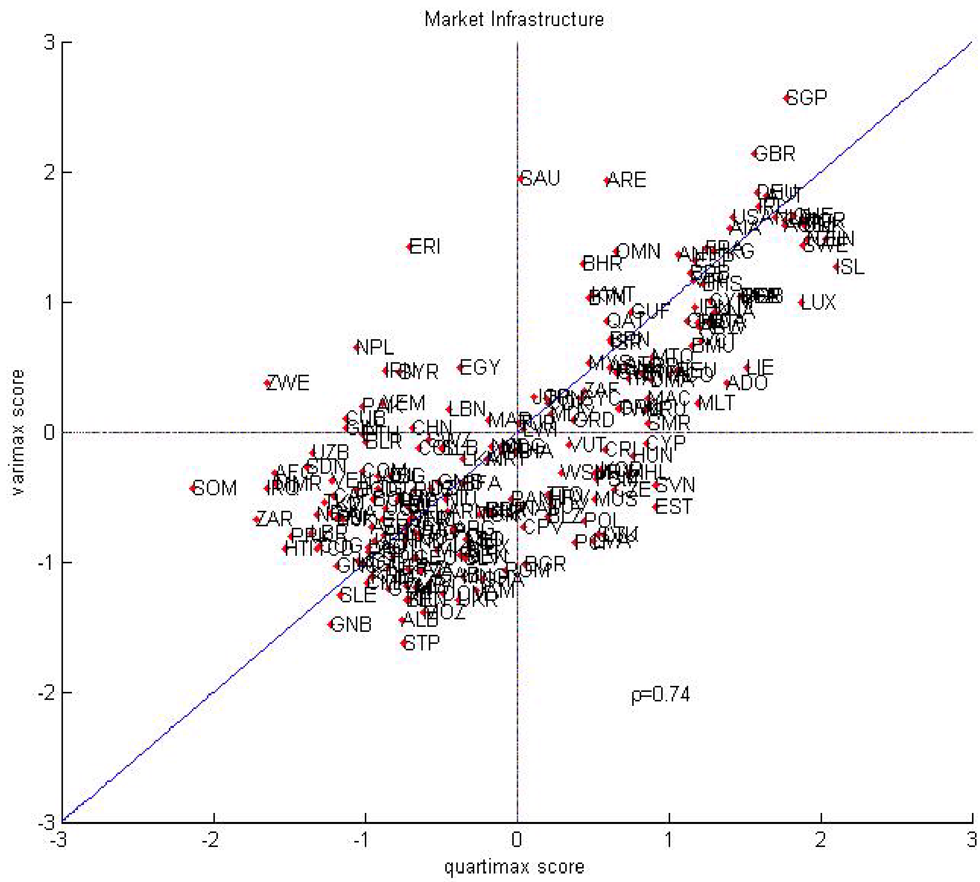

Figure 2

One can see that on the whole, scores under the two rotations are highly correlated, at $\rho = 0.74$. However, there is also a sizable number of countries that score substantially better under one rotation than the other. These discrepancies are attributable to the way in which quartimax and varimax define market infrastructure differently.

Consider countries lying farthest above the 45-degree line, e.g. - Eritrea (ERI), Saudi Arabia (SAU), United Arab Emirates (ARE),Nepal (NPL), Zimbabone (ZWE), Iran (IRN), Syria (SYR), Yemen (YEM), Cuba (CUB) and Uzbekistan (UZB). The market infrastructure scores for these countries are one standard deviation or more higher under varimax than under quartimax. This is exactly what one would expect for countries that have poor human rights records (as these do) because civil liberties concepts are relatively less important to varimax's market infrastructure than to the quartimax version. In effect, when ranked under varimax, human rights abuses do not count against these countries, whereas when ranked under quartimax, they do.

Conversely, countries lying farthest below the 45-degree line are more highly ranked under the quartimax version of market infrastructure than under the varimax version. This group is dominated by Eastern European countries like Latvia (LVA), Lithuania (LTU), Ukraine (UKR), Bulgaria (BGR), the Slovak Republic (SVK) and Slovenia (SVN). Just as one would expect given the distinctions pointed out earlier, these are countries that score relatively well in human rights (i.e. - own scores in human rights-related variables are high compared to own scores in other governance variables). When scored according to varimax's market infrastructure, their virtuous human rights records do not benefit these countries, but under the quartimax measure, they do.[31]

I prefer the varimax rendering of market infrastructure on the grounds that it puts relatively little emphasis on civil liberties content. Both *a priori* ideas about

[31] Civil liberties information is not lost under varimax; it simply shows up in f_2 instead of f_1 and f_2.

governance, as well as the results of this factor model, suggest that notions of civil liberty (electoral democracy, protection of human rights, etc.) are sufficiently distinct from the other content in f_1 that they constitute a separate factor. The emergence of f_2 is proof of this.

By including civil liberties content, the quartimax rotation casts f_1 as a general factor rather than a sharply defined facet of governance. A general factor goes against the spirit of this analysis. I am not trying to condense all of governance into a single number. Such an approach carries data reduction too far, glossing over important fundamental distinctions between governance concepts.

Next I examine how the two rotations render *civil liberties.* Factor scores are plotted in Figure 3 below.

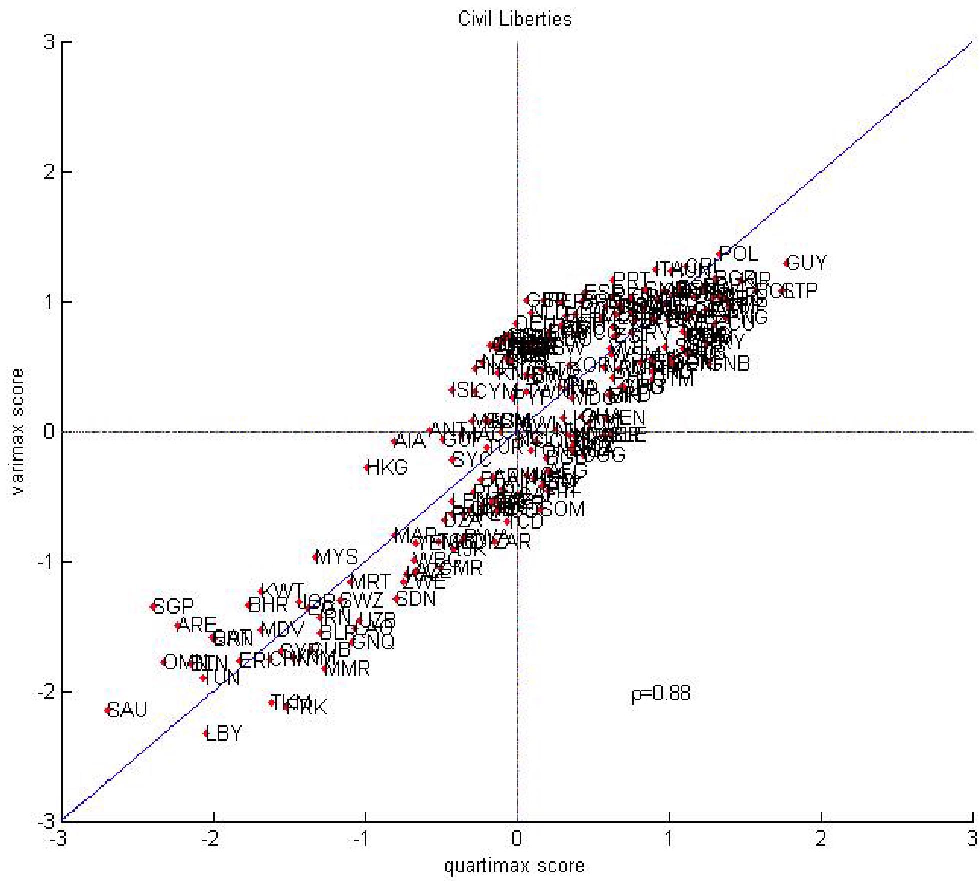

Figure 3

I concluded in an earlier section that quartimax and varimax rotations characterize *civil liberties* in a nearly identical manner, positioning the same set of variables as the most important. Figure 3 above bears this out; the quartimax and varimax *civil liberties* scores are very highly correlated ($\rho = 0.88$).

On the other hand, the varimax rendering of civil liberties bears a much closer association with observable indicators of human rights, as can be seen from Table 6 below.

Pairwise correlations			
variable	varimax f_2	quartimax f_2	N
$\text{voice}_{\text{EIU}}$	0.741	0.379	119
$\text{voice}_{\text{FRH}}$	0.915	0.667	194
$\text{voice}_{\text{GCS}}$	0.269	-0.205	114
$\text{voice}_{\text{HUM}}$	0.902	0.718	182
$\text{voice}_{\text{PRS}}$	0.695	0.375	134
$\text{voice}_{\text{RSF}}$	0.810	0.632	158
$\text{voice}_{\text{BCRI}}$	0.596	0.179	192

Table 6

Table 6 above shows pairwise correlations between f_2 scores (i.e. - $\widehat{E}\left[f_2|X\right]$) and the key civil liberties / human rights variables in the dataset. By a wide margin, the varimax-rotated scores are more strongly correlated with every such variable. Although one ought not to expect factors (especially orthogonal ones) to correspond neatly with preconceived governance notions in this way, one also shouldn't be afraid to embrace those factors that do. In weighing competing measures of civil liberties like the quartimax and varimax versions of f_2, concordance with observable civil liberties benchmarks (such as the variables in Table 6) should count as an attribute, all else equal.

A similar argument can also be made using order. Quartimax and varimax scores in order were seen to be tightly correlated ($\rho = 0.96$).[32] However - as with civil liberties (f_2) - it is the varimax version that's noticeably more congruent with benchmark

[32] When the two rotations do assess a country differently in order, it is generally not by much. Only sixteen of 215 observations have quartimax and varimax scores that differ by more than one-half a standard deviation. And the maximum discrepancy (Singapore) is just 0.7 standard deviations. The largest discrepancies are frequently cases where the varimax order ranks a Western European nation more favorably than quartimax.

measures of political and social stability, as Table 7 shows.

Pairwise correlations			
variable	varimax f_4	quartimax f_3	N
$\text{stability}_{\text{GRS}}$	0.462	0.293	120
$\text{stability}_{\text{EIU}}$	0.443	0.252	118
$\text{stability}_{\text{GCS}}$	0.740	0.631	117
$\text{stability}_{\text{HUM}}$	0.616	0.470	190
$\text{stability}_{\text{IJT}}$	0.475	0.258	180
$\text{stability}_{\text{MIG}}$	0.609	0.506	152
$\text{stability}_{\text{PRS}}$	0.602	0.471	138
$\text{stability}_{\text{BCRI}}$	0.815	0.650	198

Table 7

The correlations in the varimax column are larger than those in the quartimax column because the variables in the table load more heavily on order under the varimax rotation than under the quartimax. This difference in loadings is a direct consequence of the fact that varimax defines market infrastructure more narrowly than quartimax does.[33]

The final two factors, downside governance risk and MIG, provide no compelling evidence that favors one rotation over another. They are therefore not reviewed here.

In summary, I have shown that, to the extent that the two rotations differ in their definition of individual factors $f_1, ..., f_4$, the varimax perspective weakly dominates the quartimax for all factors. I believe that the primary reason for the varimax's superiority is its penchant for resisting a general factor. I advocate using governance scores from the varimax rotation.

[33] To see this, compare quartimax and varimax-rotated loadings on Factor 1 for stability variables using the complete Λ tables in the appendix.

3.5 A comparison to Kaufmann et al.'s WGI

I now compare my governance indicators directly to the WGI of Kaufmann et al. I perform a simple exercise aimed at demonstrating how well they can explain each other. First, I run OLS regressions of each of the six WGI on all five of my factors. I then reverse the specifications and regress each of my five factors on all six WGI. Table 8 below shows the results.

Comparison with Kaufmann et al.'s WGI - OLS Regressions

	(DEPENDENT VARIABLES)										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	CC	*RL*	*RQ*	*GE*	*VA*	*PS*	*mrkt. infr.*	*civ. lib's.*	*d-side. gov. risk*	*order*	*MIG*
CC							1.025***	-0.0649	-0.975***	-0.0341	-0.122
							(0.0858)	(0.0507)	(0.161)	(0.101)	(0.206)
RL							0.577***	-0.322***	-0.220	-0.0373	0.291
							(0.0941)	(0.0556)	(0.177)	(0.110)	(0.226)
RQ							-0.263***	0.0679	1.139***	-0.336***	-0.0464
							(0.0850)	(0.0502)	(0.160)	(0.0996)	(0.204)
GE							-0.0235	-0.218***	0.838***	-0.329**	0.0170
							(0.121)	(0.0713)	(0.227)	(0.141)	(0.290)
VA							-0.462***	1.323***	-0.484***	-0.189***	0.107
							(0.0405)	(0.0239)	(0.0760)	(0.0474)	(0.0972)
PS							-0.322***	-0.142***	0.165**	1.332***	0.0620
							(0.0412)	(0.0244)	(0.0775)	(0.0483)	(0.0990)
mrkt. infr.	0.766***	0.690***	0.528***	0.642***	0.303***	0.380***					
	(0.0151)	(0.0158)	(0.0190)	(0.0166)	(0.00883)	(0.0154)					
civil lib's.	0.368***	0.379***	0.465***	0.417***	0.846***	0.332***					
	(0.0149)	(0.0157)	(0.0189)	(0.0165)	(0.00875)	(0.0153)					
d-side. gov. risk	0.250***	0.311***	0.493***	0.419***	0.213***	0.305***					
	(0.0151)	(0.0158)	(0.0190)	(0.0166)	(0.00883)	(0.0154)					
order	0.194***	0.277***	0.110***	0.156***	0.156***	0.681***					
	(0.0148)	(0.0156)	(0.0187)	(0.0163)	(0.00869)	(0.0152)					
MIG	0.188***	0.216***	0.186***	0.199***	0.160***	0.172***					
	(0.0150)	(0.0157)	(0.0189)	(0.0165)	(0.00878)	(0.0153)					
Constant	0.00943	-4.13e-05	0.00469	0.0206	-0.0117	-0.0408***	-0.0552**	0.0456***	-0.104**	0.0404	0.152***
	(0.0147)	(0.0155)	(0.0186)	(0.0162)	(0.00863)	(0.0151)	(0.0231)	(0.0137)	(0.0434)	(0.0271)	(0.0555)
Observations	202	202	202	202	202	202	202	202	202	202	202
R-squared	0.957	0.953	0.932	0.949	0.985	0.953	0.872	0.957	0.602	0.840	0.127
F-stat.	878.9	796.9	536.6	733.1	2618	795.6	221.3	726.7	49.09	170.8	4.741

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 8

The point of the exercise is to take two sets of indicators, project each onto the other, and use what is known about each set's construction to draw inferences about the other set.[34] All regressors have been standardized to have zero mean, unit

[34]Standard errors for coefficient estimates have not been adjusted (e.g. - see Wooldridge, 2002, Ch. 6) to account for the fact that all regressors are estimated quantitites. On the other hand, it's

standard deviation to facilitate comparison.

Consider first columns (1)-(6), where the dependent variable in each specification is one of the Kaufmann et al. WGIs. The explanatory variables (1)-(6) are my five varimax-rotated governance factors.

A first observation is that the R^2's are all very high - 0.93 and above. Clearly, my five factors can replicate the variation in the WGI very well. Of course, it would be strange if this were not the case given that all the underlying data used to construct the dependent variable (and much more) are also used to construct each of the regressors.

A second observation is that every one of my five factors enters positively and highly significant (1% level) in all six regressions. This indicates that each Kaufmann et al. measure is itself a hybrid of five sharply delineated (i.e. - orthogonal) influences. This evidence supports (though it doesn't prove) the claim that my measures are deeper, more fundamental than the Kaufmann et al. WGI. To reiterate, my claim is not that my measures cause the Kaufmann et al. figures; rather, I simply claim to produce a more illuminating categorization of the governance data.

A third observation is that the size and pattern of coefficient estimates are extremely similar for columns (1)-(4). All four columns have much in common, with market infrastructure and civil liberties generally dominant, followed next by downside governance risk. Order and MIG are generally least important. Thus, loosely speaking, not only do these four Kaufmann et al. measures (CC, RL, RQ and GE)

not clear such an adjustment is necessary in this exercise, if one views the regressions as modeling the relationships between two sets of estimates contain the same ingredients (my five factors), they contain them in the roughly the same proportions.

It was clear from the high intercorrelation of CC, RL, RQ and GE in sample that they assess highly collinear phenomena. However distinctly these four concepts may have been defined by Kaufmann et al., in the data they are barely distinguishable. What one gains from columns (1)-(6), however, is a rigorous decomposition of that correlation: CC, RL, RQ and GE contain equal measures of market infrastructure, civil liberties, downside governance risk and order (and little else, as the high R^2 attests).

Not all the WGI are equivalent mixtures of my factors. Columns (5) and (6) (dependent variables are VA and PS, respectively) stand apart from the first four. Market infrastructure is no longer the dominant explanatory variable, displaced by civil liberties in column (5) and by order in column (6). However, just as with CC, RL, RQ and GE, each of my five factors has a nontrivial influence on the dependent variables VA and PS. Columns (5) and (6) again demonstrate that a single WGI category is actually a complicated hybrid of orthogonal concepts.

I now move to columns (7)-(11), where the dependent variable in each specification is now one of my five governance factors. The explanatory variables in columns (7)-(11) are the six Kaufmann et al. WGI measures.

Regressors enter highly significant (1% level) roughly half as often in columns (7)-(11) as they do in columns (1)-(6). Not surprisingly then, one sees that the R^2's are generally lower compared with specifications (1)-(6). The R^2 is lowest in (9) and (11), where the dependent variable is one of my data-source-specific governance factors. There is no parallel to data-source-specific factors in the WGI because the Kaufmann et al. approach never estimates a model with more than one variable from each source. The lower R^2 values in (7)-(11) generally reflect the fact the WGI are more similar to one another than my measures are.

Another contrast with columns (1)-(6) is that more than half the coefficient estimates in (7)-(11) are negative; of the 18 negative coefficient estimates, ten are significant at the 1% level. What explains the abundance of negative and significant estimates? Evidently, the variation unique to each of the explanatory variables is often negatively correlated with my governance factors.

A further difference evident in (7)-(11) compared to (1)-(6) is that the coefficient patterns found in (7)-(11) are unique; no two columns in (7)-(11) look remotely alike. This is not surprising given the mutual orthogonality of the dependent variables in (7)-(11).

In summary, four general findings emerge. First, my governance factors explain the Kaufmann et al. WGI better than the latter explain my governance factors. Second, each Kaufmann et al. WGI category is a composite of five very different (i.e. - orthogonal) components. Third, the makeup of four of the WGI categories (CC, RL, RQ, GE) is virtually identical in terms of these five components. Fourth, my governance factors are deeper, or more fundamental, than the WGI in the sense that each of my measures exerts a significant marginal impact on all WGI categories, whereas each WGI category is found to exert a small and statistically insignificant impact on at least one of my factors.[35]

4 Other Issues

4.1 Factors or Components?

The factor model employed in this paper implicitly assumes a causal relationship between f and X: an observation's scores in the factors (f) cause observable variables (X) to take on certain values. I have also spoken about the factors I uncover as latent capacities, further emphasizing their role as causal forces behind the observable governance data.

This approach is open to question. The factor model, after all, was developed by psychologists to model cognitive capacities of human beings. Can countries be said to possess capacities in a way that's at all analogous to the ways humans possess intelligence? If so, where do these capacities reside, and how are they determined? Crucially for policymakers, is a country's governance capacity fixed forever in the same way that an individual's intelligence is fixed by her genetic inheritance? If it is fixed, how does one reconcile that conclusion with the widespread belief that countries can, through reforms and hard work, improve their governance?

At its core, the issue is whether it is valid to interpret the linear combinations of governance variables produced by my factor model as anything more meaningful than complicated averages satisfying certain statistical properties. Consider, for com-

[35] Four of the Kaufmann et al. WGI (CC, RL, RQ and GE) are insignificant for two or more of our governance factors.

parison, factor model applications in macroeconomics, where hundreds of time series variables are reduced to, say, a half-dozen factors. Such factors have proven useful for forecasting other time series, for dating recessions and expansions, etc. But in that research, authors do not maintain that the factors so extracted exist on their own, somewhere beyond the data, exerting a causal effect on the data. The factors are simply a lower-dimensional version of the data - indices of economic activity. Should factors extracted from governance data be interpreted any differently? Are they really more than the sum of their parts?

From a philosophical point of view, I believe that causal factors at the country level are just as defensible as causal factors at the individual person level. A person's core intellectual and psychological attributes are forged by her genetic inheritance and by her environment. Psychometricians factor analyze the correlation of test items to uncover those core attributes. Similarly, a country's tendency to govern well or govern poorly is shaped by its history - conquest, colonialism, religious upheaval - and its environment - climate, natural resources, coastlines, etc. Why is it not equally valid to factor analyze correlations of expert assessments to uncover those tendencies?

From a practical point of view, this debate matters for my results only if the two approaches define governance dimensions and/or rank countries in substantially different ways. To embody the alternative view that governance dimensions are properly regarded as indexes rather than causal factors, one should use the principal components (PC) model. The components are linear combinations of the variables, constructed such that each successive component explains the maximum possible amount of the data's variation, subject to being orthogonal to all other components.

A standard distinction that is drawn between the two models is to say that the PC model explains variance whereas the factor model explains covariances. There is no hypothesis that components cause the data, nor is there any partitioning of the data's variation into systematic and idiosyncratic parts. The PC model is data reduction, pure and simple.

To test the practical import of the factors vs. components distinction, I estimate a PC model on the governance data and compare component loadings to factor loadings. I specify a five-component model and rotate to the varimax solution for comparability with earlier results. The ten largest-loading variables for each component are displayed in Table 9 below.

Principal Components Model Loadings Estimates (varimax rotation) - Dominant Variables for Each Component

component 1		component 2		component 3		component 4		component 5	
rulelaw_{GCS}	0.2879	voice_{HUM}	0.4136	regulation_{MIG}	0.4108	regulation_{GRS}	0.5567	stability_{GCS}	0.5161
rulelaw_{QLM}	0.2877	voice_{FRH}	0.3942	corrupt_{MIG}	0.3941	goveffect_{GRS}	0.3597	stability_{BCRI}	0.3933
corrupt_{GCS}	0.2769	voice_{RSF}	0.3608	goveffect_{MIG}	0.3818	rulelaw_{GRS}	0.3369	stability_{PRS}	0.3708
corrupt_{QLM}	0.2727	voice_{PRS}	0.2945	rulelaw_{MIG}	0.3591	regulation_{PRS}	0.3103	stability_{MIG}	0.3111
regulation_{GCS}	0.272	rulelaw_{HUM}	0.2938	stability_{MIG}	0.2654	stability_{GRS}	0.3024	stability_{HUM}	0.2741
goveffect_{GCS}	0.2449	voice_{EIU}	0.2557	goveffect_{EGV}	0.2305	corrupt_{GRS}	0.2325	stability_{GRS}	0.2168
goveffect_{EIU}	0.2311	regulation_{HER}	0.1836	corrupt_{PRS}	0.2182	regulation_{BCRI}	0.1358	stability_{EIU}	0.1762
rulelaw_{HER}	0.218	voice_{BCRI}	0.1776	goveffect_{PRS}	0.1736	regulation_{GCS}	0.1055	rulelaw_{PRS}	0.1102
corrupt_{EIU}	0.2175	regulation_{EIU}	0.1354	regulation_{HER}	0.1678	regulation_{HER}	0.1002	rulelaw_{GRS}	0.0897
rulelaw_{BCRI}	0.2099	stability_{EIU}	0.1188	stability_{HUM}	0.1386	goveffect_{EGV}	0.0993	goveffect_{GRS}	0.0851

Table 9

Each governance factor has a clear parralel component in the PC model. Factors one and two (market infrastructure, civil liberties) correspond to components one and two, respectively. Factor three (downside governance risk) corresponds to component four. Factor four (order) corresponds to component five. Finally, factor five (MIG factor) corresponds to component three. At least eight of the ten variables with highest loadings on each factor are also among the ten variables with highest loadings on the corresponding component. Evidently, the PC model and the factor model define the dimensions of governance very similarly.

For that reason, countries are ranked in very similar ways, whether I use their component scores or factor scores. Table 10 below displays correlations of each factor with its corresponding component in the balanced dataset sample ($n = 73$).

Correlation of Factor Scores with Component Scores - Balanced Dataset (n=73)

factor/component definition	ρ(factor score, component score)
market infrastructure	0.824
civil liberties	0.886
downside governance risk	0.799
order	0.860
MIG	0.616

Table 10

I conclude based on the high correlation of factor and component scores for governance data that inferences based on the factor model will not be seriously altered by switching to the PC model. My results are robust to the method of extraction.

4.2 Reliability: Factor Scores as Scales

The particular linear combination of the X's used to construct each factor score can be looked upon as a test or a rating scale, comprising 45 items. The items in this case are not the raw variables in X, but rather the variables multiplied by their corresponding score coefficients. Let these modified versions of the variables be denoted $\widetilde{X}$. Note there are five versions of $\widetilde{X}$ - one version corresponding to each factor (since each factor has a unique vector of score coefficients). Cronbach's alpha estimated on the 45 variables in the ith version of $\widetilde{X}$ measures the reliability, or internal consistency, of the ith factor. Table 11 below displays Cronbach's alpha for my five factors:

Scale Reliability Coefficient (Cronbach's α) for Each Factor

factor	Cronbach's α
market infrastructure	0.870
civil liberties	0.891
downside governance risk	0.928
order	0.910
MIG	0.932

Table 11

A common rule of thumb is that Cronbach's alpha should be at least 0.7 if a scale is truly measuring a single construct. All my factors have values well above α, indicating good reliability.

5 Conclusion

In this chapter, I have applied an established statistical technique to the problem of measuring governance. While the technique itself (factor analysis) is not new, it has not, to my knowledge, been applied in this way to such an exhaustive dataset of governance variables. The result of my analysis is a new set of governance indicators that have a number of advantages over both individual expert assessments of governance, and the widely used WGI of Kaufmann et al., which aggregate multiple expert assessments.

By changing the way governance is defined and measured, this chapter also necessarily changes the way one views individual countries. In particular, I contend that my approach defines the strengths and weaknesses of a country's governance capacities more sharply than the WGI of Kaufmann et al. A simple indication of this is the variation in a single country's scores across governance categories.

For example, consider the U.K., which ranks highly in five of six Kaufmann et al. measures. The U.K.'s scores across the six Kaufmann et al. governance categories range from a minimum of 0.34 (Political Stability/No Violence) to a maximum of 1.94 (control of corruption). This compares to a minimum in my rankings of -1.4 (order) and a maximum in my rankings of 2.14 (market infrastructure).[36] Or, consider Colombia. Its Kaufmann et al. scores range from -1.79 (political stability/no violence) to 0.05 (regulatory quality); under my measures, Colombia's minimum and maximum scores are -3.14 (order) and 0.54 (downside governance risk). Wide within-country variation in scores is in fact a general characteristic of my measures, and it stems from my decision to characterize each of the four governance measures exclusively in terms of what makes it unique from all the others.

My growth regression results (Chapter 2) demonstrate one practical advantage of this approach, but it is natural to wonder also about the drawbacks. Moving from a framework of highly correlated governance indicators to a framework of orthogonal governance indicators entails a tradeoff. What I gain in terms of a sharper delineation between concepts has to be weighed against what I give up in terms of the interpretability of country rankings. Some country scores in some of my governance factors will seem surprising in light of pre-existing beliefs about those countries.

[36] Recall that the marginal density of each Kaufmann et al. governance indciator is $N(0,1)$, just like our governance factors.

On the other hand, I regard this power to surprise as another benefit of my approach. I have taken pains to motivate my methodology and contend that the rankings that emerge should not be dismissed because they clash in some instances with existing measures. Rather, my measures should serve as inputs into the refinement of the very governance concepts under examination.

One practical next step in my research is to explore ways of incorporating governance data on countries with incomplete observations. my factor model is estimated on the largest possible balanced dataset of 73 countries, thereby excluding the incomplete observations of 142 countries. Another direction for future research is to look for external sources of validation for my governance measures. I demonstrate their explanatory power for the cross-country variation in income in Chapter 2, but more work needs to be done in evaluating my measures through their relationships with other, independent economic and governance benchmarks. It would also be helpful to compare the content of my four factors with the output of factor models estimated on entirely different sets of data - either different governance variables, or the same variables observed at different points in time.

Chapter 2

1 Introduction

In this chapter, I use four of the governance measures developed in Chapter 1 ($f_1 - f_4$) to assess the relative economic importance of governance, trade and geography - three potentially primal ingredients of long-run prosperity. Because my four governance measures are sufficiently distinct from one another, I can characterize the contribution of governance to growth at an unprecedented level of detail.

I know of no previous work that has used multiple governance measures jointly in a growth regression, along with trade and geography controls. I suspect that the reason may be that no one has gone to the lengths that I have to develop conceptually distinct measures. The struggle to find appropriate instruments may have been an obstacle as well.

Besides the use of multiple governance controls as explanatory variables, what sets my growth regressions apart from earlier work is the breadth of information contained in each measure. While Kaufmann et al.'s Rule of Law measure - the preferred governance indicator in Rodrik et al. (2004) - is comprised of nine of the variables in the governance dataset, market infrastructure, civil liberties, downside governance risk, and order measures are each linear combinations of 45 underlying variables. I do not contend that more inputs necessarily equate to better governance measures, but it does seem important to understand how the incorporation of a wider variety of governance perceptions validates or contradicts existing results.

To preview my results, I find that trade and geography have no statistically significant direct effect on growth once governance is accounted for. I find furthermore that from the perspective of growth, the most important components of governance are market infrastructure, civil liberties and order. The impact of downside governance risk is never statistically significant.

My presentation is modeled on that of Rodrik, Subramanian and Trebbi's "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development" (2004). Here, as in that paper, I perform instrumental variables regressions of per capita income on measures of governance, trade and geography. I adopt Rodrik et al.'s notation to make my results as comparable as possible to theirs, and many of the variables used here were acquired through personal correspondence with those authors. However, many of my specifications have no direct parallel in Rodrik et al. - either because I employ different governance measure(s), different instrument(s), or slightly different samples. At the end of this section, I compare and contrast the implications of my results with those of other authors.

For robustness, I estimate all specifications on four different samples. The constitution of each sample is motivated by a different concern. The smallest sample comprises the 64 countries for which all 45 variables in the governance dataset are observed; 73 countries actually satisfy this requirement, but nine are missing at least one other variable required in the growth regression, such as the Frankel-Romer predicted trade share (see below). The rationale for this sample is simply that factor score predictions for these countries are based on the fullest possible set of data. The second sample comprises the 79 countries for which Acemoglu et al.'s (2001) colonial settler mortality variable is available .[37] The settler mortality variable has proven to work well in instrumenting for perceptions-based governance measures such as Kaufmann et al.'s Rule of Law (Acemoglu et al., 2001; Rodrik et al, 2004). The third sample mirrors the "large sample" of Rodrik et al. (2004) and comprises 138 countries.[38] I utilize this sample for its broad country coverage and for the sake of comparability to Rodrik et al.'s results. My final and largest sample comprises 155 countries and is the largest possible sample, given the constraints of the data. My qualitative results are very similar across samples. I offer explanations where they differ.

2 Model

I begin the formal analysis by presenting the equation of interest,

$$\log y = \mu + \boldsymbol{\alpha}'\mathbf{GOV} + \beta INT + \gamma GEO + \varepsilon \tag{10}$$

where y is per capita output, **GOV** is governance, INT is integration (trade/GDP), and GEO is a geography measure (distance from the equator).[39] Rodrik et al. measure **GOV** with Rule of Law, whereas I measure **GOV** with one, two, three or all four of my varimax-rotated governance factors ($f_1 - f_4$). Variable $\mathbf{GOV} = [GOV_1, ..., GOV_K]'$

[37] Settler mortality actually exists for 81 countries, but Myanmar lacks income data and Ivory Coast lacks governance data.

[38] Rodrik et al.'s published results (2004) are for 137 countries, but data provided by personal correspondence with those authors enables estimation for 138 countries. We have not been able to ascertain which single country was excluded from the 137-ctry. sample used in the published results.

[39] Variable y is measured using real (PPP-adjusted) 2005 per capita output and integration, INT, is a country's 2005 trade-to-GDP ratio (in logs) - both are taken from *Penn World Tables*, Mark 6.3. Variable GEO is (the absolute value of) the country's latitude (measured at the capital city).

can therefore be vector-valued (correspondingly, $\boldsymbol{\alpha} = [\alpha_1, ..., \alpha_K]'$).

Table 12 below presents summary statistics for all the variables in (1), by sample. Standard deviations are in parentheses beneath the means.

Summary statistics - OLS regressions				
	Balanced dataset (N=64)	Settler mortality sample (N=79)	Rodrik et al. large sample (N=138)	Full sample (N=155)
log y	9.318	8.356	8.839	8.788
	(1.039)	(1.157)	(1.213)	(1.259)
GEO	29.531	15.582	23.783	22.884
	(17.52)	(11.363)	(16.34)	(16.01)
INT	4.349	4.322	4.409	4.407
	(0.513)	(0.534)	(0.499)	(0.577)
mrkt. infrastructure	0.317	-0.357	-0.057	-0.037
	(1.009)	(0.811)	(0.97)	(0.945)
civil liberties	0.21	0.072	0.088	0.048
	(0.905)	(0.809)	(0.894)	(0.926)
d-side. gov. risk	0.258	-0.273	-0.081	-0.138
	(0.905)	(0.903)	(0.916)	(0.974)
order	-0.427	-0.303	-0.095	-0.068
	(0.912)	(0.877)	(0.909)	(0.940)

log y = natural log of 2005 real (PPP-adjusted) per capita output (Penn World Tables, Mark 6.3)

INT = natural log of [(imports+exports)/GDP] (Penn World Tables, Mark 6.3)

GEO = absolute value of capital city's latitude

Table 12

I seek estimates of the causal impact of each of the right-hand-side variables on income. Equation (1) represents a horse race of sorts - an experiment to find out which deep determinant matters most for long-run prosperity. It is a parsimonious specification, and no estimates of α, β, γ will settle definitively the debate over the complex process of economic development. But estimating (1) can lend support to one hypothesis or another, helping to guide policy discussions, foreign aid strategies, and future research.

For reasons discussed below, OLS estimates of $\boldsymbol{\alpha}, \beta, \gamma$ in (1) are unlikely to measure causal impacts. I nevertheless begin my analysis by presenting the OLS results in Table 13 as a benchmark.

OLS Results *(Dependent variable is log 2005 per capita income.)*				
	Balanced dataset (N=64)	Settler mortality sample (N=79)	Rodrik et al. large sample (N=138)	Full sample (N=155)
	(1)	(2)	(3)	(4)
GEO	0.00395	0.00947	0.00901*	0.0102**
	(0.00528)	(0.00840)	(0.00482)	(0.00453)
INT	0.0738	0.375**	0.235*	0.205*
	(0.160)	(0.175)	(0.129)	(0.110)
mrkt. infrastructure	0.715***	0.718***	0.738***	0.720***
	(0.0819)	(0.118)	(0.0763)	(0.0732)
civil liberties	0.380***	0.582***	0.326***	0.289***
	(0.0792)	(0.101)	(0.0681)	(0.0645)
d-side. gov. risk	0.272***	0.377***	0.343***	0.409***
	(0.0787)	(0.0924)	(0.0672)	(0.0630)
order	0.216**	0.0738	0.224***	0.241***
	(0.0888)	(0.103)	(0.0710)	(0.0655)
Constant	8.595***	6.928***	7.651***	7.738***
	(0.772)	(0.841)	(0.612)	(0.515)
Observations	64	79	138	155
R-squared	0.784	0.673	0.706	0.700
F-stat.	34.41	24.64	52.54	57.48

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 13

The OLS estimates suggest that all the explanatory variables play a direct role in determining income. Estimates reflect the fact that the simple and partial correlation of income with all the covariates is positive. Trade and geography, in particular, always enter positively, frequently at a 10% or lower significance level.

If one were to rely solely on the OLS estimates, one would have to conclude from Table 10 that each and every deep determinant makes an important marginal contribution to growth. However, as discussed at length in Rodrik et al. (2004) and elsewhere in the literature, the OLS estimates of $\boldsymbol{\alpha}, \beta, \gamma$ in (1) are unlikely to capture causal impacts because there is good reason to expect that governance and integration measures are correlated with the disturbance, ε.[40] For this reason, I estimate an IV model using two-stage least squares. I instrument for **GOV** and INT using a variety of variables, to which I now turn.

3 Instrumental Variables Specification

My instrument for endogenous regressor INT is the Frankel-Romer (1999) constructed trade share measure (denoted $TRADESHARE$), computed based on a gravity model of trade. Vector **GOV** represents up to four endogenous regressors, and so I review my instrument choice for each component of **GOV** individually.

For market infrastructure, one instrument I employ is colonial settler mortality, from Acemoglu et al. (2001) (denoted SETMORT). Variable SETMORT has been used numerous times in the literature, provides solid first-stage fit for its intended regressor, and has an intuitive story motivating its use. I forego a detailed justification

[40] For instance, higher incomes may purchase better governance (reverse causality); or, shocks to income may simultaneously affect the quality of governance (omitted variable). Either effect would cause the OLS estimate of $\boldsymbol{\alpha}$ to overstate the causal impact of **GOV** on $\log y$ (OLS estimates of β, γ will also be inconsistent, though in unknown directions). Alternatively, measurement error in our proxy variable for governance (measurement error $= f - \widehat{E}\left[f|X\right]$) would cause $\widehat{\boldsymbol{\alpha}}_{ols}$ to *under*state **GOV**'s causal impact.

Similar problems likely afflict our measure of integration. Richer countries may on average prefer to trade relatively more than poorer countries (reverse causality); or, countries that are rich for some reason not accounted for by (10) may also naturally trade more for that reason (omitted variable). (Frankel and Romer (1999) find evidence against the latter possibility.)

of this instrument and refer readers to Acemoglu et al. (2001) or Rodrik et al. (2004) for a more in-depth discussion. SETMORT's availability (79 countries) defines my second-smallest sample.

For the balanced dataset sample (N=64) and for the two larger samples - Rodrik et al.'s large sample (N=138) and the full sample of (N=155) - SETMORT is unavailable; I replace it with U.K., German and Scandinavian legal origin dummies (denoted LEGOR_E, LEGOR_{GE}, and LEGOR_{SC}, respectively), taken from La Porta et al. (1999).[41]

The use of legal origin to explain the quality of governance is not new in the economics literature. In La Porta et al. (1999), the authors regress measures of governance on legal origin and cultural and economic controls to test competing theories of the determinants of good government.[42] Legal origin has good explanatory power for market infrastructure in all the samples and may be regarded as exogenous in this context so long as it is not correlated with today's level of income through some channel other than trade, geography and the included measures of governance. I move ahead under the assumption that legal origin is a valid instrument.

Civil liberties is the second potential component of **GOV**. I instrument for civil liberties using the variable EURFRAC, the fraction of a country's population speaking one of five major European languages as their first language.[43] EURFRAC has been

[41] The left-out groups are French and socialist legal origins.

[42] The basic premise for using legal origin in that paper, and in this one, is that legal origin proxies for "the relative power of the State vis-a-vis property owners." (La Porta et al., 1999). Governments in common law countries (U.K. legal origin) are generally regarded as less inclined to, or less capable of, impinging on the rights of private property than governments in civil law countries (German, French, Scandanavian legal origin).

[43] English, French, German, Portuguese, Spanish; see Hall and Jones (QJE, 1999) for another example of this use of eurfrac. Our source for the eurfrac variable is Rodrik et al. (2004).

used an instrument for governance in circumstances similar to ours (Hall and Jones, 1999; Rodrik et al. 2004). A country's value of EURFRAC may be considered exogenous in this context so long as it is not correlated with today's level of income through some channel other than trade, geography and the included measures of governance.

For the remaining two potential components of **GOV** - downside governance risk, and order - I rely on the legal origin dummies introduced above and two measures of religious affiliation: CATHO80 and MUSLIM80, measuring the fraction of the population that is Roman Catholic and Muslim, respectively. Both religious variables are taken from La Porta et al. (1999), which used them to test the power of culture (as proxied by religion) to explain government quality. I operate under the assumption that a country's religious composition can be considered exogenous.[44]

The general form of the first-stage equations is thus,

$$
\begin{aligned}
INT &= \theta + \sigma TRADESHARE + \omega GEO + \boldsymbol{\tau}'\mathbf{Z} + \nu \qquad (11) \\
GOV_1 &= \lambda_1 + \phi_1 TRADESHARE + \chi_1 GEO + \boldsymbol{\delta}_1'\mathbf{Z} + \xi_1 \\
&\ldots \\
GOV_K &= \lambda_K + \phi_K TRADESHARE + \chi_K GEO + \boldsymbol{\delta}_K'\mathbf{Z} + \xi_K \qquad (12)
\end{aligned}
$$

where $\mathbf{Z} \equiv [Z_1, ..., Z_M]'$, $M \geq K$ is the instrument vector for **GOV**. Vector **Z** com-

[44] Endogeneity of religion by reverse causality or omitted variables both seem implausible. The only other possibility is that religion affects income through some channel other than trade, geography and governance, which also seems unlikely.

prises one or more of the excluded instruments for governance, as described above.

Finally, $\mathbf{GOV} \equiv [GOV_1, ..., GOV_K]', K \in \{1, ..., 4\}$ is made up of one or more of my governance factors: GOV_1 =market infrastructure; GOV_2 =civil liberties; GOV_3 =downside governance risk; and GOV_4 =order.

Table 14 below presents summary statistics, by sample, for the excluded instruments. Standard deviations are in parentheses beneath the means.[45]

Summary statistics - excluded instruments

	Balanced dataset (N=64)	Settler mortality sample (N=79)	Rodrik et al. large sample (N=138)	Full sample (N=155)
EURFRAC	0.337	0.304	0.251	0.25
	(0.43)	(0.413)	(0.392)	(0.395)
TRADESHARE	2.615	2.76	2.945	3.004
	(0.809)	(0.765)	(0.818)	(0.816)
SETMORT	-	4.647	-	-
	(-)	(1.201)	(-)	(-)
$LEGOR_E$	0.375	0.354	0.319	0.355
	(0.488)	(0.481)	(0.468)	(0.48)
$LEGOR_{GE}$	0.063	0	0.043	0.039
	(0.244)	(0)	(0.205)	(0.194)
$LEGOR_{SC}$	0.047	0	0.036	0.032
	(0.213)	(0)	(0.188)	(0.177)
CATHO80	0.364	0.376	0.341	0.33
	(0.383)	(0.364)	(0.363)	(0.363)
MUSLIM80	0.194	0.241	0.212	0.239
	(0.341)	(0.345)	(0.343)	(0.367)

TRADESHARE = natural log of Frankel-Romer (1999) constructed trade share; SETMORT = natural log of Acemoglu et al. (2001) settler mortality

EURFRAC = fraction of pop. speaking one of five major European languages as first language

$LEGOR_E$, $LEGOR_{GE}$, $LEGOR_{SC}$ = English, German and Scandinavian legal origin dummies, respectively

CATHO80, MUSLIM80 = fraction of pop. that is Roman Catholic and Muslim, respectively

Table 14

4 Estimation

I now present my main results - two-stage least squares estimates of $\boldsymbol{\alpha}, \beta, \gamma$ from (1). I begin with a few observations on the first-stage results presented in Panel B of

[45] The zeros for $\mathrm{LEGOR_{GE}}$ and $\mathrm{LEGOR_{SC}}$ in the N=79 column reflect the fact that no country in the 79-country settler mortality sample has German or Scandinavian legal origin.

Tables 15-18.

IV (2SLS) Results - balanced dataset

Panel A: second-stage results (Dependent variable is log 2005 per capita income.)

	(1)	(2)	(3)	(4)	(5)	(6)
GEO	0.0188***	0.0316***	0.0152***	0.0117*	0.00237	-0.0235
	(0.00593)	(0.00576)	(0.00546)	(0.00693)	(0.0110)	(0.0238)
INT	0.162	0.539**	0.284	0.211	-0.195	-0.921
	(0.209)	(0.255)	(0.192)	(0.205)	(0.402)	(0.750)
mrkt. infrastructure	0.603***		0.567***	0.604***	0.654***	0.883***
	(0.135)		(0.122)	(0.125)	(0.166)	(0.277)
civil liberties		0.358**	0.311***	0.378***	0.229	0.424*
		(0.143)	(0.104)	(0.132)	(0.144)	(0.238)
d-side. gov. risk				0.173		0.697
				(0.226)		(0.506)
order					0.680	1.294*
					(0.450)	(0.746)
Observations	64	64	64	64	64	64
R-squared	0.662	0.478	0.727	0.755	0.553	0.210
Sargan p-val	0.00404	0.00317	0.0143	0.00412	0.150	0.573
Cragg-Donald Wald p-val	0.0000	0.0000	0.0000	0.1285	0.3858	0.3961

Panel B: first-stage results - regressing INT and governance on the instruments

	(1)	(2)	(3)	(4)	(5)
	INT	mrkt. infrastructure	civil liberties	d-side. gov. risk	order
GEO	-0.00417	0.0279***	0.0143***	0.01000	0.00736
	(0.00330)	(0.00603)	(0.00520)	(0.00734)	(0.00762)
TRADESHARE	0.460***	0.342***	0.0566	0.123	0.320**
	(0.0650)	(0.119)	(0.102)	(0.144)	(0.150)
EURFRAC	-0.171	1.015***	0.0740	-0.0271	-0.000132
	(0.164)	(0.300)	(0.259)	(0.366)	(0.379)
$LEGOR_E$	-0.0408	0.836***	0.223	-0.353	-0.393
	(0.123)	(0.225)	(0.194)	(0.273)	(0.284)
$LEGOR_{GE}$	-0.0777	0.415	-0.0743	-0.292	-0.0286
	(0.219)	(0.401)	(0.346)	(0.488)	(0.506)
$LEGOR_{SC}$	-0.271	0.988**	0.157	-1.675***	0.392
	(0.264)	(0.482)	(0.416)	(0.587)	(0.609)
CATHO80	-0.210	-0.119	0.911***	-1.210**	-0.134
	(0.214)	(0.391)	(0.337)	(0.475)	(0.493)
MUSLIM80	-0.326*	0.426	-1.280***	-0.274	-0.303
	(0.179)	(0.326)	(0.281)	(0.397)	(0.412)
R-squared	0.536	0.599	0.629	0.261	0.218
F-stat.	7.937	10.28	11.67	2.431	1.912

Constants suppressed; standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 15

IV (2SLS) Results - settler mortality sample

Panel A: second-stage results (Dependent variable is log 2005 per capita income.)

	(1)	(2)	(3)	(4)	(5)	(6)
GEO	0.00618	0.0395***	0.00674	0.0123	0.0341	0.0326
	(0.0141)	(0.0110)	(0.0128)	(0.0248)	(0.0278)	(0.0299)
INT	0.389	1.026***	0.504	0.367	1.995	1.757
	(0.351)	(0.361)	(0.320)	(0.621)	(1.215)	(1.564)
mrkt. infrastructure	1.217***		0.946***	1.464*	0.388	0.642
	(0.295)		(0.273)	(0.871)	(0.579)	(1.134)
civil liberties		1.129***	1.008***	0.973***	1.160***	1.128***
		(0.237)	(0.188)	(0.352)	(0.298)	(0.336)
d-side. gov. risk				-1.117		-0.395
				(1.528)		(1.485)
order					-0.975	-0.851
					(0.734)	(0.904)
Observations	79	79	79	79	79	79
R-squared	0.360	0.134	0.471			
Sargan p-val	0.0000	0.0208	0.1392	0.4333	0.7790	NA
Cragg-Donald Wald p-val	0.0000	0.0000	0.0001	0.5078	0.1429	0.3601

Panel B: first-stage results - regressing INT and governance on the instruments

	(1)	(2)	(3)	(4)	(5)
	INT	mrkt. infrastructure	civil liberties	d-side. gov. risk	order
GEO	-0.00176	0.0154*	0.00386	0.0134	0.0259**
(0.00558)	(0.00809)	(0.00837)	(0.0115)	(0.00981)	
TRADESHARE	0.405***	0.0140	0.0177	-0.0133	0.628***
	(0.0670)	(0.0970)	(0.100)	(0.138)	(0.118)
SETMORT	-0.129**	-0.271***	0.0148	-0.0374	0.0857
	(0.0498)	(0.0721)	(0.0746)	(0.102)	(0.0874)
EURFRAC	-0.0663	0.371	0.981***	0.398	-0.167
	(0.160)	(0.232)	(0.240)	(0.329)	(0.281)
CATHO80	-0.201	-0.917***	0.0696	-0.697*	-0.268
	(0.187)	(0.270)	(0.280)	(0.384)	(0.328)
MUSLIM80	-0.236	-0.294	-0.475	-0.0963	-0.907**
	(0.200)	(0.289)	(0.299)	(0.411)	(0.351)
R-squared	0.405	0.463	0.416	0.127	0.317
F-stat.	8.171	10.06	8.294	1.701	5.409

Constants suppressed; standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 16

IV (2SLS) Results - Rodrik et al.'s large sample

Panel A: second-stage results (Dependent variable is log 2005 per capita income.)

	(1)	(2)	(3)	(4)	(5)	(6)
GEO	0.0220***	0.0372***	0.0160**	0.0152*	0.0102	0.00739
	(0.00670)	(0.00547)	(0.00674)	(0.00898)	(0.00786)	(0.0101)
INT	0.522**	0.572**	0.451**	0.446**	0.00869	-0.0293
	(0.212)	(0.263)	(0.210)	(0.210)	(0.364)	(0.359)
mrkt. infrastructure	0.688***		0.643***	0.642***	0.759***	0.760***
	(0.159)		(0.157)	(0.156)	(0.177)	(0.169)
civil liberties		0.584***	0.553***	0.561***	0.530***	0.554***
		(0.152)	(0.120)	(0.132)	(0.123)	(0.130)
d-side. gov. risk				0.0456		0.142
				(0.338)		(0.338)
order					0.416	0.436
					(0.278)	(0.270)
Observations	138	138	138	138	138	138
R-squared	0.590	0.356	0.600	0.612	0.589	0.625
Sargan p-val	0.0001	0.0128	0.1551	0.0773	0.2366	0.1068
Cragg-Donald Wald p-val	0.0000	0.0000	0.0000	0.1482	0.0194	0.0797

Panel B: first-stage results - regressing INT and governance on the instruments

	(1)	(2)	(3)	(4)	(5)
	INT	mrkt. infrastructure	civil liberties	d-side. gov. risk	order
GEO	.0018	0.0287***	0.0121***	0.0199***	0.00186
	(0.00241)	(0.00429)	(0.00423)	(0.00534)	(0.00506)
TRADESHARE	0.363***	0.130*	0.0380	0.0415	0.436***
	(0.0424)	(0.0751)	(0.0741)	(0.0934)	(0.0885)
EURFRAC	-0.187*	0.660***	0.483**	0.0198	0.0150
	(0.110)	(0.195)	(0.192)	(0.243)	(0.230)
$LEGOR_E$	0.1458	0.616***	0.161	-0.00886	-0.470**
	(0.0873)	(0.155)	(0.153)	(0.193)	(0.183)
$LEGOR_{GE}$	-0.0578	0.893***	0.0691	0.134	-0.115
	(0.181)	(0.320)	(0.315)	(0.398)	(0.377)
$LEGOR_{SC}$	-0.240	1.080***	0.168	-1.036**	-0.126
	(0.218)	(0.385)	(0.380)	(0.479)	(0.454)
CATHO80	-0.0045	0.0501	0.431*	-0.458	-0.540*
	(0.147)	(0.261)	(0.257)	(0.324)	(0.307)
MUSLIM80	-0.2642**	0.693***	-0.956***	-0.0697	-0.705***
	(0.128)	(0.227)	(0.224)	(0.282)	(0.267)
R-squared	0.40	0.506	0.434	0.142	0.218
F-stat.	10.77	16.49	12.34	2.669	4.504

Constants suppressed; standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 17

IV (2SLS) Results - full sample

Panel A: second-stage results (Dependent variable is log 2005 per capita income.)

	(1)	(2)	(3)	(4)	(5)	(6)
GEO	0.0215***	0.0367***	0.0166***	0.0156*	0.0152**	0.0122
	(0.00625)	(0.00555)	(0.00642)	(0.00892)	(0.00656)	(0.00917)
INT	0.597***	0.626**	0.477**	0.466**	0.120	0.0318
	(0.212)	(0.266)	(0.216)	(0.221)	(0.496)	(0.505)
mrkt. infrastructure	0.714***		0.672***	0.668***	0.707***	0.700***
	(0.153)		(0.154)	(0.153)	(0.159)	(0.149)
civil liberties		0.542***	0.511***	0.515***	0.501***	0.513***
		(0.146)	(0.118)	(0.119)	(0.117)	(0.112)
d-side. gov. risk				0.0594		0.170
				(0.390)		(0.388)
order					0.287	0.335
					(0.361)	(0.354)
Observations	155	155	155	155	155	155
R-squared	0.565	0.312	0.556	0.578	0.567	0.622
Sargan p-val	0.0001	0.00478	0.127	0.0570	0.0809	0.0234
Cragg-Donald Wald p-val	0.0000	0.0000	0.0000	0.2425	0.2069	0.2035

Panel B: first-stage results - regressing INT and governance on the instruments

	(1)	(2)	(3)	(4)	(5)
	INT	mrkt. infrastructure	civil liberties	d-side. gov. risk	order
GEO	0.005*	0.0283***	0.0107**	0.0215***	0.00173
	(0.00289)	(0.00407)	(0.00422)	(0.00554)	(0.00505)
TRADESHARE	0.373***	0.146**	0.0219	0.0695	0.500***
	(0.0500)	(0.0701)	(0.0725)	(0.0952)	(0.0868)
EURFRAC	-0.159	0.632***	0.515***	0.0819	-0.0211
	(0.132)	(0.185)	(0.192)	(0.252)	(0.230)
$LEGOR_E$	0.054	0.655***	0.110	0.0482	-0.310*
	(0.0999)	(0.141)	(0.146)	(0.191)	(0.174)
$LEGOR_{GE}$	-0.161	0.900***	0.0774	0.167	-0.0713
	(0.222)	(0.311)	(0.322)	(0.423)	(0.386)
$LEGOR_{SC}$	-0.401	1.080***	0.223	-0.971*	-0.110
	(0.266)	(0.372)	(0.385)	(0.506)	(0.461)
CATHO80	-0.054	0.0563	0.427	-0.313	-0.517*
	(0.178)	(0.250)	(0.258)	(0.339)	(0.309)
MUSLIM80	-0.345**	0.621***	-0.944***	-0.0832	-0.709***
	(0.142)	(0.199)	(0.206)	(0.270)	(0.246)
R-squared	0.307	0.495	0.437	0.122	0.217
F-stat.	8.10	17.87	14.14	2.545	5.052

Constants suppressed; standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 18

Across the four samples, R^2's are reasonable, and a different set of instruments is dominant in each of columns (1)-(5), suggesting the instruments do a good job at bringing out the unique variation in each endogenous regressor (Shea, 1997). Where I had strong priors, the instruments generally enter as expected, but there are also a couple of interesting surprises. The negative coefficient on LEGOR_E in the order equations (column (5), Panel B of Tables 15, 17, 18) may stem from the more decentralized political and legal landscape in common law countries. The negative coefficients on CATHO80 and MUSLIM80 in those same equations may reflect a higher perceived risk of terrorism in predominantly Catholic or Muslim countries, relative to Protestant countries.

It is somewhat surprising that settler mortality is so strongly predictive of market infrastructure (column (2), Panel B, Table 13) but insignificant for civil liberties (column (3), Panel B, Table 13). The implication from comparing columns (2) and (3) is that countries with disease environments less hostile to early European colonizers could expect to better regulation and less corruption, but not stronger protections of human rights. The importance of EURFRAC for civil liberties (columns (3), Panel B, Tables 13-15) suggests that it is the extent of Europeanization over centuries - not the survival rates of the earliest European colonists - that ultimately has determined civil liberties protections.

Finally, MUSLIM80's effect on market infrastructure (column (2), Panel B, Tables 15,17,18) is positive and significant in all but the settler mortality sample, where it enters negative. The predominantly Muslim Gulf states (e.g. - Saudi Arabia, Oman, Kuwait) are not included in the 79-country settler mortality sample. The Gulf states tend to be much wealthier, and score much better in market infrastructure, than the predominantly Muslim states in the settler mortality sample (e.g. - Bangladesh, Indonesia and Pakistan).

I now address the second-stage results, presented in Panel A of Tables 15-18. The dependent variable in all specifications is the log of 2005 per capita income. I make some general observations about the 2SLS results, focusing first on the broad area of agreement between the two samples.

First, broadly speaking, my findings support the "institutions rule" result of Rodrik et al. (2004): controlling for the quality of governance, trade and geography have no statistically significant direct effect on income. INT and GEO generally lose their significance once market infrastructure, civil liberties and order are included in the specification (i.e. - columns (5) and (6) of Panel A).[46]

Second, within the realm of governance, market infrastructure and civil liberties consistently exhibit the largest and most precisely measured effects on income. One or the other (usually both) enters highly significant in every specification, for every sample. In my two largest samples, market infrastructure and civil liberties both enter highly significant under every specification, and their coefficient magnitudes change minimally as additional explanatory variables are included.

Third, in all but one instance (market infrastructure in the full sample) IV estimates of the coefficients on market infrastructure and civil liberties exceed the OLS

[46] INT, especially, is neutralized by the introduction of order. The ability of order to displace INT so thoroughly points to a shared source of variation between the two. The most likely candidate, given our first-stage results, is population. Both INT and order are strongly negatively correlated with population, and their mutual reliance on the instrument TRADESHARE in the first-stage causes this population-related component of their variation to be passed on to the second stage.

estimates, suggesting that measurement error rather than simultaneity or omitted variables is the source of the inconsistency of OLS.[47] This result is consistent with the findings of earlier authors, e.g. - Rodrik et al., and Hall and Jones (1999).

Failure of the overidentification tests in columns (1)-(4) of Panel A is frequent and stems from my decision to use a fixed set of instruments irrespective of the number of endogenous regressors. These failures would be more troubling if I thought that each specification needed to stand on its own. However, I interpret the failures, where they occur, as evidence of omitted variables in the structural equation, a problem that disappears when I include all four of my governance factors in column (6).[48]

To test for the potential impact of weak instruments on my IV estimates, I re-estimate the specifications in Tables 15-18 using the limited-information maximum likelihood (LIML) IV estimator. I present the LIML results in the Appendix. [49]

5 Conclusion

My results show for the first time which areas of governance are the most important for long-term prosperity. Market infrastructure and civil liberties are the sine qua non of economic growth. While the proposition that both of these areas of governance are vital for long-term prosperity may not come as a great surprise, nevertheless precise,

[47] Orders's IV coefficient estimate exceeds its OLS estimate in all but the settler mortality sample.

[48] The one exception is column (6), Panel A of Table 15. Here the failure of the overidentification restrictions appears to stem from the explanatory power of MUSLIM80 in the full sample. Judging from the full sample, MUSLIM80 belongs in the structural equation. When included, MUSLIM80 has a positive and significant (10%) impact on per capita income. This is not true for any of the the three samples. The finding suggests a role for culture alongside governance, trade and geography.

[49] In each of the four samples, LIML estimates are qualitatively very similar to the 2SLS estimates. One exception is in the fullest ($n = 155$) sample, where the LIML coefficient estimates for market infrastructure and civil liberties have similar magnitudes as the 2SLS estimates but are insignificant.

robust estimates of their marginal economic impacts are a novel contribution of this paper.

I also note that I could not have carried out this exercise without sufficiently independent measures of governance. For example, running the equivalent to (6) using the six highly intercorrelated Kaufmann et al. WGI in place of my four governance factors produces statistical insignificance for every regressor - trade, geography and all components of governance! There is sufficient independent evidence from the literature to reject the substantive implication of that result. Rather, the cause of the insignificance result is the fact that the fitted values of the six WGI variables from the first-stage are so highly correlated with one another[50] that the resulting collinearity in the second stage makes it impossible to precisely estimate the marginal effects of each variable.[51]

[50] The mean correlation between the fitted values of two WGI variables (e.g., between $\widehat{RL}$ and $\widehat{CC}$) is $\overline{\rho} = 0.93$ in our full sample of 155 countries.

[51] This result holds true for a variety of alternative instrument vectors.

Chapter 3

1 Oblique Rotations

In chapters one and two, I rely on a particular orthogonal rotation of Λ_0 to conduct inference on governance and growth. In this section, I check the sensitivity of those results to allowing governance factors to be correlated. I examine factor content, factor score predictions, and growth regression results under three oblique rotations of Λ_0, and I find very little difference from the varimax results of chapters one and two. The differences that do emerge are discussed, as are the costs and benefits of forcing orthogonality on the factors.

Results are presented in Tables 19-21 below. Each table shows results for one oblique rotation. The three rotations presented below were selected because they are widely used in the psychometric literature and because they each pursue the goals of simple structure in a straightforward way (Harman, 1976). Three other oblique rotations were tried in addition to these (two oblimin rotations, and the biquartimin rotation) and they yielded very similar results. For the sake of brevity, they are not presented.

Table 19 presents results from the quartimin rotation. The quartimin is equivalent to the orthogonal quartimax but without the orthogonality constraint on the rotation matrix. Like the quartimax criterion, quartimin seeks to maximize the variation in squared loadings across factors for each variable. The aim is to summarize each variable in terms of the fewest factors factors possible.

ROTATION NAME: quartimin

PARAMETERS: none

Factor loadings (ten variables with highest loadings on each factor):

F1		F2		F3		F4		F5	
variable	loading	variable	loading	variable	loading	variable	loading	variable	loading
$\text{rulelaw}_{\text{qlm}}$	0.949	$\text{corrupt}_{\text{mig}}$	0.823	$\text{regulati}_{\text{grs}}$	0.870	$\text{voice}_{\text{hum}}$	0.908	$\text{stability}_{\text{bcri}}$	0.791
$\text{corrupt}_{\text{qlm}}$	0.924	$\text{goveffect}_{\text{mig}}$	0.768	$\text{goveffect}_{\text{grs}}$	0.773	$\text{voice}_{\text{frh}}$	0.888	$\text{stability}_{\text{gcs}}$	0.790
$\text{corrupt}_{\text{gcs}}$	0.738	$\text{rulelaw}_{\text{mig}}$	0.725	$\text{rulelaw}_{\text{grs}}$	0.719	$\text{voice}_{\text{rsf}}$	0.740	$\text{stability}_{\text{prs}}$	0.625
$\text{corrupt}_{\text{eiu}}$	0.727	$\text{regulation}_{\text{mig}}$	0.724	$\text{stability}_{\text{grs}}$	0.623	$\text{voice}_{\text{prs}}$	0.627	$\text{stability}_{\text{mig}}$	0.575
$\text{rulelaw}_{\text{her}}$	0.700	$\text{corrupt}_{\text{prs}}$	0.423	$\text{corrupt}_{\text{grs}}$	0.580	$\text{voice}_{\text{eiu}}$	0.610	$\text{stability}_{\text{hum}}$	0.569
$\text{rulelaw}_{\text{gcs}}$	0.700	$\text{stability}_{\text{mig}}$	0.377	$\text{regulation}_{\text{prs}}$	0.542	$\text{rulelaw}_{\text{hum}}$	0.595	$\text{stability}_{\text{grs}}$	0.386
$\text{goveffect}_{\text{eiu}}$	0.698	$\text{goveffect}_{\text{prs}}$	0.324	$\text{regulation}_{\text{bcri}}$	0.360	$\text{regulation}_{\text{her}}$	0.454	$\text{stability}_{\text{eiu}}$	0.370
$\text{corrupt}_{\text{bcri}}$	0.643	$\text{voice}_{\text{gcs}}$	0.287	$\text{regulation}_{\text{gcs}}$	0.349	$\text{voice}_{\text{bcri}}$	0.453	$\text{rulelaw}_{\text{prs}}$	0.239
$\text{regulation}_{\text{gcs}}$	0.627	$\text{rulelaw}_{\text{her}}$	0.268	$\text{voice}_{\text{gcs}}$	0.315	$\text{regulation}_{\text{eiu}}$	0.352	$\text{rulelaw}_{\text{bcri}}$	0.237
$\text{rulelaw}_{\text{bcri}}$	0.603	$\text{regulation}_{\text{her}}$	0.256	$\text{voice}_{\text{bcri}}$	0.299	$\text{goveffect}_{\text{prs}}$	0.315	$\text{rulelaw}_{\text{grs}}$	0.194

Factor covariance matrix (Phi):

	F1	F2	F3	F4	F5
F1	1.000				
F2	0.689	1.000			
F3	0.620	0.533	1.000		
F4	0.466	0.457	0.397	1.000	
F5	0.423	0.456	0.433	0.348	1.000

Correlation of quartimin factor score predictions with varimax-rotated predictions:

(n=138)	rho
market infrastructure	0.940
civil liberties	0.964
downside governance risk	0.886
order	0.919
MIG	0.773

Growth regressions:

IV (2SLS) Results using quartimin factor scores - Rodrik et al.'s large sample

	(1)	(2)	(3)	(4)	(5)	(6)
	Second-stage results; dependent variable is log 2005 per capita income.					
GEO	0.0113*	0.0250***	0.00691	0.00812	0.00255	0.00399
	(0.00620)	(0.00658)	(0.00659)	(0.00794)	(0.00810)	(0.00925)
INT	0.232	0.418*	0.186	0.145	-0.0682	-0.132
	(0.192)	(0.247)	(0.200)	(0.250)	(0.338)	(0.391)
mrkt. infrastructure	0.857***		0.658***	0.578*	0.647***	0.541
	(0.130)		(0.150)	(0.326)	(0.151)	(0.336)
civil liberties		0.799***	0.486***	0.332	0.472***	0.268
		(0.186)	(0.162)	(0.579)	(0.163)	(0.597)
downside governance risk				0.253		0.334
				(0.912)		(0.938)
order					0.233	0.242
					(0.250)	(0.258)
Observations	138	138	138	138	138	138
R-squared	0.693	0.460	0.670	0.664	0.667	0.648
Sargan p-val.	0.00345	0.00606	0.117	0.0663	0.0911	0.0503
Cragg-Donald Wald p-val	0.0000	0.0000	0.0000	0.5967	0.0082	0.4456

Constants not shown; standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1.

Table 19

Table 20 presents results from the covarimin rotation. The covarimin rotation is the oblique analogue to the orthogonal varimax rotation. Like the varimax, the covarimin seeks to maximize the variation in squared loadings across variables for each factor, but without restricting the rotation matrix to be orthogonal. The aim is to characterize each factor in terms of the fewest variables possible.

ROTATION NAME: covarimin

PARAMETERS: none

Factor loadings (ten variables with highest loadings on each factor):

F1		F2		F3		F4		F5	
variable	loading	variable	loading	variable	loading	variable	loading	variable	loading
$corrupt_{mig}$	0.855	$rulelaw_{qlm}$	0.761	$regulati_{grs}$	0.879	$voice_{hum}$	0.908	$stability_{bcri}$	0.847
$goveffect_{mig}$	0.798	$corrupt_{qlm}$	0.738	$goveffect_{grs}$	0.799	$voice_{frh}$	0.885	$stability_{gcs}$	0.830
$rulelaw_{mig}$	0.783	$corrupt_{gcs}$	0.574	$rulelaw_{grs}$	0.739	$voice_{rsf}$	0.738	$stability_{prs}$	0.665
$regulation_{mig}$	0.740	$corrupt_{eiu}$	0.566	$stability_{grs}$	0.625	$voice_{prs}$	0.633	$stability_{hum}$	0.621
$corrupt_{prs}$	0.515	$goveffect_{eiu}$	0.544	$corrupt_{grs}$	0.612	$voice_{eiu}$	0.628	$stability_{mig}$	0.603
$goveffect_{prs}$	0.400	$rulelaw_{her}$	0.541	$regulation_{prs}$	0.564	$rulelaw_{hum}$	0.619	$stability_{grs}$	0.421
$rulelaw_{her}$	0.371	$rulelaw_{gcs}$	0.541	$regulation_{gcs}$	0.414	$voice_{bcri}$	0.472	$stability_{eiu}$	0.417
$voice_{gcs}$	0.353	$corrupt_{bcri}$	0.495	$regulation_{bcri}$	0.403	$regulation_{her}$	0.461	$rulelaw_{bcri}$	0.297
$stability_{mig}$	0.337	$regulation_{gcs}$	0.492	$voice_{gcs}$	0.359	$regulation_{eiu}$	0.371	$rulelaw_{prs}$	0.291
$regulation_{her}$	0.286	$rulelaw_{bcri}$	0.460	$goveffect_{gcs}$	0.351	$goveffect_{prs}$	0.341	$rulelaw_{grs}$	0.230

Factor covariance matrix (Phi):

	F1	F2	F3	F4	F5
F1	1.000				
F2	0.606	1.000			
F3	0.559	0.536	1.000		
F4	0.466	0.372	0.414	1.000	
F5	0.512	0.372	0.478	0.380	1.000

Correlation of covarimin factor score predictions with varimax-rotated predictions:

(n=138)	rho
market infrastructure	0.982
civil liberties	0.961
downside governance risk	0.873
order	0.877
MIG	0.746

Growth regressions:

IV (2SLS) Results using covarimin factor scores - Rodrik et al.'s large sample

	(1)	(2)	(3)	(4)	(5)	(6)
	Second-stage results; dependent variable is log 2005 per capita income.					
GEO	0.0139**	0.0297***	0.0129*	0.0111	0.00809	0.00764
	(0.00680)	(0.00580)	(0.00688)	(0.00996)	(0.00787)	(0.0104)
INT	0.432**	0.434*	0.345*	0.313	-0.0162	-0.0215
	(0.201)	(0.251)	(0.204)	(0.236)	(0.358)	(0.365)
mrkt. infrastructure	0.849***		0.573***	0.519*	0.551***	0.537*
	(0.154)		(0.172)	(0.274)	(0.171)	(0.275)
civil liberties		0.673***	0.488***	0.480***	0.406***	0.404***
		(0.145)	(0.130)	(0.130)	(0.145)	(0.145)
downside governance risk				0.126		0.0335
				(0.501)		(0.510)
order					0.340	0.337
					(0.277)	(0.279)
Observations	138	138	138	138	138	138
R-squared	0.637	0.432	0.630	0.644	0.636	0.640
Sargan p-val.	0.000728	0.0408	0.155	0.0765	0.152	0.0694
Cragg-Donald Wald p-val	0.0000	0.0000	0.0000	0.3378	0.0057	0.2125

Constants not shown; standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1.

Table 20

Table 21 presents results from the promax rotation. The promax criterion rotates Λ_0 until it resembles as closely as possible the loading pattern in a target matrix (in this case, the orthogonal varimax loadings matrix) that has had each of its elements raised to a power (in this case, three).

ROTATION NAME: promax

PARAMETERS: power = 3

Factor loadings (ten variables with highest loadings on each factor):

F1		F2		F3		F4		F5	
variable	loading	variable	loading	variable	loading	variable	loading	variable	loading
$\text{rulelaw}_{\text{qlm}}$	1.052	$\text{regulation}_{\text{grs}}$	0.907	$\text{voice}_{\text{hum}}$	0.985	$\text{corrupt}_{\text{mig}}$	0.676	$\text{stability}_{\text{gcs}}$	0.805
$\text{corrupt}_{\text{qlm}}$	1.038	$\text{goveffect}_{\text{grs}}$	0.797	$\text{voice}_{\text{frh}}$	0.962	$\text{goveffect}_{\text{mig}}$	0.628	$\text{stability}_{\text{bcri}}$	0.794
$\text{corrupt}_{\text{gcs}}$	0.857	$\text{rulelaw}_{\text{grs}}$	0.738	$\text{voice}_{\text{rsf}}$	0.802	$\text{regulation}_{\text{mig}}$	0.596	$\text{stability}_{\text{prs}}$	0.648
$\text{rulelaw}_{\text{her}}$	0.835	$\text{stability}_{\text{grs}}$	0.644	$\text{voice}_{\text{prs}}$	0.677	$\text{rulelaw}_{\text{mig}}$	0.580	$\text{stability}_{\text{mig}}$	0.591
$\text{rulelaw}_{\text{gcs}}$	0.828	$\text{corrupt}_{\text{grs}}$	0.578	$\text{voice}_{\text{eiu}}$	0.666	$\text{stability}_{\text{mig}}$	0.304	$\text{stability}_{\text{hum}}$	0.571
$\text{corrupt}_{\text{eiu}}$	0.824	$\text{regulation}_{\text{prs}}$	0.537	$\text{rulelaw}_{\text{hum}}$	0.650	$\text{corrupt}_{\text{prs}}$	0.302	$\text{stability}_{\text{grs}}$	0.359
$\text{goveffect}_{\text{eiu}}$	0.800	$\text{regulation}_{\text{gcs}}$	0.333	$\text{regulation}_{\text{her}}$	0.496	$\text{goveffect}_{\text{prs}}$	0.224	$\text{stability}_{\text{eiu}}$	0.348
$\text{regulation}_{\text{gcs}}$	0.717	$\text{regulation}_{\text{bcri}}$	0.329	$\text{voice}_{\text{bcri}}$	0.493	$\text{voice}_{\text{gcs}}$	0.189	$\text{rulelaw}_{\text{prs}}$	0.230
$\text{corrupt}_{\text{bcri}}$	0.715	$\text{voice}_{\text{gcs}}$	0.287	$\text{regulation}_{\text{eiu}}$	0.386	$\text{regulation}_{\text{her}}$	0.174	$\text{rulelaw}_{\text{bcri}}$	0.208
$\text{goveffect}_{\text{gcs}}$	0.707	$\text{goveffect}_{\text{gcs}}$	0.267	$\text{goveffect}_{\text{prs}}$	0.352	$\text{rulelaw}_{\text{her}}$	0.154	$\text{regulation}_{\text{mig}}$	0.162

Factor covariance matrix (Phi):

	F1	F2	F3	F4	F5
F1	1.000				
F2	0.688	1.000			
F3	0.562	0.505	1.000		
F4	0.619	0.506	0.471	1.000	
F5	0.495	0.486	0.435	0.433	1.000

Correlation of promax factor score predictions with varimax-rotated predictions:

(n=138)	rho
market infrastructure	0.915
civil liberties	0.925
downside governance risk	0.864
order	0.908
MIG	0.862

Growth regressions:

IV (2SLS) Results using promax factor scores - Rodrik et al.'s large sample

	(1)	(2)	(3)	(4)	(5)	(6)
	Second-stage results; dependent variable is log 2005 per capita income.					
GEO	0.00971	0.0181***	0.00676	0.00735	0.00398	0.00463
	(0.00649)	(0.00685)	(0.00674)	(0.00716)	(0.00800)	(0.00861)
INT	0.0746	0.445*	0.154	0.0527	-0.0309	-0.0742
	(0.207)	(0.231)	(0.214)	(0.339)	(0.356)	(0.406)
mrkt. infrastructure	0.940***		0.624***	0.521	0.607***	0.545*
	(0.143)		(0.191)	(0.327)	(0.194)	(0.325)
civil liberties		0.846***	0.466**	0.274	0.450**	0.333
		(0.165)	(0.181)	(0.521)	(0.184)	(0.525)
downside governance risk				0.378		0.233
				(0.956)		(0.978)
order					0.173	0.155
					(0.265)	(0.281)
Observations	138	138	138	138	138	138
R-squared	0.673	0.517	0.658	0.630	0.655	0.640
Sargan p-val.	0.0130	0.0270	0.125	0.0886	0.0807	0.0406
Cragg-Donald Wald p-val	0.0000	0.0000	0.0000	0.5642	0.0094	0.4141

Constants not shown; standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1.

Table 21

The first general observation is that all three oblique rotations identify more or less the same set of concepts in the first five factors that the varimax rotation finds - although oblique rotations often extract the factors in a different sequence.[52] Thus, every oblique rotation finds a "market infrastructure" factor, a "civil liberties" factor, etc. This result is demonstrated by the high degree of similarity in the composition of top-ten variable lists across rotations. For example, in all three oblique rotations, the same three variables (rulelaw_{QLM}, corrupt_{QLM} and corrupt_{GCS}) emerge as the first-, second-, and third-most important variables for each rotation's version of "market infrastructure" (two of these are also in the top three of the varimax's market infrastructure). Likewise, the same three variables (voice_{HUM}, voice_{FRH} and voice_{RSF}) emerge in the same order as top three for each rotation's civil liberties factor (exactly as in the varimax version of civil liberties). Finally, the same seven variables (and in very similar order) show up at the top of each oblique rotation's "order" factor (also mirroring the pattern in varimax).

Another general result is that the five factors within any single oblique rotation are strongly intercorrelated with one another. While the oblique rotations tried here do not force factors to be correlated, they do relax the constraint that imposes orthogonality, resulting in oblique configurations whenever such configurations increase the value of the rotation's maximand (i.e., whenever an oblique configuration can produce a "simpler" loading pattern). The pairwise correlation of any two factors within a given oblique rotation is generally in the neighborhood of 0.5.

[52] In oblique rotations, the explanatory powers of the factors (i.e. - their associated eigenvalues) tend to be much closer in magnitude to one another than in orthogonal rotations. Therefore, f_1 does not play the dominant explanatory role that it does under orthogonal rotations; likewise, $f_2 - f_5$ do not play as insignificant a role in explaining Σ as they did under orthogonal rotations.

A third general observation about Tables 19-21 is that factor score predictions from oblique rotations are highly correlated with score predictions from the varimax rotation. The tables show the sample pairwise correlations between varimax and oblique predictions of corresponding factors for the Rodrik et al. (2004) sample of 138 countries. The minimum correlation for any pairing is 0.746, but the vast majority of correlations are 0.85 and above. Clearly, scoring countries based on oblique factors will yield very similar conclusions to scoring countries based on the orthogonal varimax factors. It is no surprise, then, that results of 2SLS growth regressions from Chapter 2 also change very little when I replace varimax factor scores with oblique-rotation factor scores. Both coefficient magnitudes and significances are preserved following the replacement. As with the 2SLS varimax results, market infrastructure and civil liberties enter highly significantly in nearly all specifications, and trade (INT) and geography (GEO) drop in magnitude and significance once three or more governance factors are controlled for. The second-stage results are shown in the tables.

In toto, the oblique rotation results present a bit of a puzzle. On the one hand, orthogonality evidently imposes a nontrivial constraint on the makeup of Λ for our data, since maximizing unconstrained objective functions similar to the varimax criterion produces highly correlated factors. This outcome seems to amount to a rejection of orthogonality. On the other hand, oblique rotations of Λ_0 do very little to alter either the varimax's matching of variables to factors or its ranking of countries in the factor scores. This outcome suggests that the practical impacts of imposing orthogonality on our data are small to inconsequential.[53]

[53] In the aggregate, the varimax and the oblique rotations rank countries in a highly correlated

Given these results, the choice between an oblique or an orthogonal orientation of the factors ought to hinge on the aims of the researcher. For some situations, orthogonality *per se* may be desirable, simply because the researcher believes the latent constructs are truly independent, or because a follow-on application requires a maximally distinct set of scores. In datasets like this one, however - where each variable is strongly correlated with nearly all others - orthogonal factors pose an added interpretive challenge. Essentially, the data are saying that market-infrastructure-like characteristics tend not to vary independently from civil-liberties-like characteristics. Therefore, in labeling orthogonal factors based on this kind of data, one must stress their residual nature ("civil liberties, after controlling for market infrastructure, order...etc.") so as to avoid equating the factors with the terms used to label them.

If the researcher's prior suggests the latent concepts under scrutiny are actually correlated, then an oblique rotations may be superior. With an oblique rotation, one no longer needs to qualify factor labels with "...controlling for the other factors"-type statements. But as the factors become highly correlated with one another, their distinguishing characteristics become harder to appreciate.

Implicitly in this dissertation, I have taken the stand that an orthogonal rotation is preferable for my purposes, but I also recognize the very reasonable arguments for using an oblique rotation. However, the invariance of many of my results to allowing

---

order. But there are individual countries for which the impacts of switching rotations is nontrivial. For example, some countries with poor human rights records have significantly better market infrastructure scores under varimax than under oblique rotations. The reason is similar to the explanation of deviations in quartimax and varimax scores in Chapter 1. The oblique versions of market infrastrucure are correlated with human rights variables, whereas the orthogonal varimax version of market infrastructure is not. Therefore, poor human rights records don't count against countries under varimax, but they do count against countries under oblique rotations.

factors to be correlated suggests that such arguments, stimulating though they may be, hold little consequence for ranking countries in governance.

2 Missing Data

2.1 Extent

There are a great many observations missing from the governance dataset. If one considers all the countries for which at least one governance variable is observed, then 32% of the cells in this 215×45 governance data matrix are missing. Such a high percentage of data missing means that any approach I take to deal with the missing data is likely to have a nontrivial impact on my estimation results.

The missing data pattern is such that the complete 45×1 data vector is observed for only 73 countries. Therefore, complete-case analysis - estimation based on only the complete observations - discards 142 partial observations (some missing as few as a single variable), amounting to a two-thirds reduction in the number of countries in my sample!

I estimate the factor model in this paper using complete-case analysis, so at a very minimum I need to contemplate the potential biases introduced by this decision.

2.2 Mechanism

To understand the potential impact of using complete-case analysis one must first assess the reason(s) that the data are missing, i.e. - identify the missing data mechanism. Complete-case analysis yields consistent (but inefficient) estimates under one type of missingness mechanism, but inconsistent estimates under all others.

Little and Rubin (2002) classify missing data mechanisms into three types, depending on how the property of missingness is related to values taken by the data.

The first type, called missing completely at random (MCAR), is the most innoccous. Data are said to be MCAR if the likelihood that a value is missing is independent of the value of that or any other variable in the data set (missing or observed).[54] If my data were MCAR, then the sample of 73 countries with no missing variables (the "complete-case sample") would in fact be a random sample of the population. MCAR therefore would mean factor model estimates based on the complete-case sample are unbiased (though not efficient).

The second type, called missing at random (MAR), requires more attention. Data are said to be MAR if the likelihood that a value is missing depends on observed but not unobserved values in the dataset. For instance, it may be that missingness for one corruption variable c_1 is negatively correlated with the overall level of corruption, but that after conditioning on another fully observed corruption measure c_2, missingness for c_1 is random.

The final type, called not missing at random (NMAR), is the most pernicious of all and requires the greatest amount of *a priori* information to treat properly. If data are NMAR, then the likelihood that a value is missing depends on both observed and unobserved values for that observation. Extending the corruption example given

[54]Formally, Little and Rubin (2002) characterize missingness mechanisms in terms of conditional distributions of the missingness pattern, given the data. When data are MCAR, the conditional distribution of the missingness pattern given the data is just the marginal density of the missingness pattern.

above, one would say that c_1 is NMAR if the likelihood that c_1 is missing for a given country depends on unobserved attributes of that country (such as the value of c_1 itself), even after conditioning on the value of c_2. If one knew the nature of the dependence on unobserved variables, then that information could be factored into the estimation process and perhaps produce consistent estimates. However in my case, without some kind of forensic evidence (e.g. - discussions with a data provider reveal an idiosyncratic sample selection procedure) it will generally be difficult to ascertain the nature of this dependence.

Which type of missingness mechanism is at work in the governance dataset? Consider that missingness in this case is a by-product of the aggregation of variables from a dozen independent data sources that did not coordinate their country-coverage decisions. What it literally means to say that the governance dataset has "missing" values is that different variables cover different (though overlapping) samples of countries. This situation differs from the more standard scenario in which multivariate data comes from a single survey, a single battery of test questions, or a set of instrument readings on a sample of observations. However, the theoretical issue remains the same: what is the relationship between the property of missingness and the values that the data take? In the context of expert opinion data, the answer to that question hinges on whether data providers decide whether to cover a country based on reasons that are correlated with the governance attributes being measured. If the answer to this question is 'no', then the data can be considered MCAR, and the complete-case sample will yield consistent estimates.

One might think that, since values in the governance dataset are missing deliberately (i.e. - because a data provider intentionally excluded a country) rather than by chance, missingness cannot be MCAR. But this is not so. Data can still be MCAR even though missingness has a systematic explanation. In our case, MCAR only requires that the rule generating missing values, whatever it may be, operates orthogonally to governance quality. So, for example, if a data provider excluded all countries in which the capital city was founded in an odd-numbered year, missingness would be systematic but (presumably) MCAR.

Is it possible that the governance data are MCAR? Consider that the majority of governance variables are scores published by profit-maximizing risk-consulting firms. In equilibrium one expects these firms to publish scores on a particular country only if demand for information on that country is sufficiently strong. Demand for information in turn depends largely on the interests of clients - chiefly investors and multinational corporations. Under this rationale, one should be willing to assume the data are MCAR if it can reasonably be expected that (across clients, and across countries) the basis for client interest in a country is independent of the quality of governance.

Prima facie, the complete-case sample does appear to be a diverse cross-section of countries, so I cannot reject MCAR out of hand. But on closer inspection the data suggest that client interest is in fact not independent of the quality of governance and thus that the data are not MCAR. Let $miss_i \in \{0, 1, 2, ..., 45\}$ be the number of values missing for country i. If the missingness mechanism is the same across data providers, then $miss_i$ measures the propensity for country i to be excluded under that mechanism. One then might ask: how is this propensity related to the quality of governance, as measured by the observed data? To answer that question, I calculate the pairwise correlations between $miss$ and each of the 45 variables. I find that $miss$ is negatively correlated with 35[55] out of 45 variables.[56] In other words, a country's propensity to be excluded by data providers is decreasing in its quality of governance; missingness conveys negative information about a country.[57]

The above exercise tells us something about the difference between more-complete and less-complete observations. Another way to shed light on the missingness mechanism is to directly compare the complete-case group of countries with the incomplete-case group of countries in terms of governance and non-governance variables. If systematic differences emerge, then these will constitute further evidence that the complete-case group is not a representative sample, forcing us to reject MCAR.

I run 45 independent t-tests on the equality of variable means across the two groups and find striking results. For 37 of the 45 governance variables, I can reject $H_0 : \mu_{complete} = \mu_{incomplete}$ in favor of $H_a : \mu_{complete} > \mu_{incomplete}$ at the 1% level.[58] For only one of the remaining eight variables can I reject equality in the oppositte direction, namely $H_{\tilde{a}} : \mu_{complete} < \mu_{incomplete}$ - and then, only at the 10% level. Evidently, complete-case countries on average have higher scores than incomplete-case countries,

[55] Pairwise correlations for 31 of the 35 are significant at the 5% level.

[56] For variable j, the sample upon which the pairwise correlation with $miss$ is estimated comprises the countries for which variable j is observed. These samples differ across variables.

[57] An alternative interpretation of the negative correlations would be that countries with fewer observed values appear to possess relatively poor governance because the data providers that do cover them are more pessimistic than other data providers. But given that all 45 variables in our dataset are highly positively correlated, it seems unlikely that providers who assess "poorly covered" countries are evaluating those countries in a way not representative of the way that all data providers would have assessed them.

[58] Interestingly, six of the eight variables for which we cannot reject $H_0 : \mu_{complete} = \mu_{incomplete}$ fall into the Political Stability/No Violence (PSNV) category of Kaufmann et al. Evidently, the complete-data sample of 73 countries is closer to being a random sample in the dimensions measured by PSNV variables.

suggesting again that missingness conveys negative information about governance.[59]

In the Table 22 below I present more t-test results - this time for covariates outside the area of governance. The results show still more systematic differences between complete-case and incomplete-case countries. Complete-case countries are richer, farther from the equator, more populous, and more linguistically Europeanized than incomplete-case countries.

Two-sample t-tests for difference in means - selected covariates

	real per capita income (2005)	population (2005)	distance from equator	eurfrac	tropics
mean, complete obs.	16,530.37	72,565.12	31.93	0.30	0.36
mean, incomplete obs.	9,801.38	9,516.94	20.80	0.17	0.76
std. dev., complete obs.	12,750.18	199,353.30	17.74	0.42	0.46
std. dev., incomplete obs.	13,274.33	17,767.12	14.38	0.35	0.40
n, complete obs.	73	73	73	73	61
n, incomplete obs.	113	113	128	110	80
t statistic (incomplete - complete)	-3.46	-2.70	-4.57	-2.11	5.30
upper one-sided p-value	1.00	1.00	1.00	0.98	0.00
lower one-sided p-value	0.00	0.00	0.00	0.02	1.00
two-sided p-value	0.00	0.01	0.00	0.04	0.00

Note: unequal variances assumed.

Table 22

The behavior of $miss$ and the t-test results strongly suggest that complete-case countries differ systematically from incomplete-case countries in level of governance, an observation inconsistent with the MCAR hypothesis. I therefore reject MCAR.[60]

[59] Two potential explanations for the observed pattern of missingness that are presumably independent of governance are population and geographic remoteness. 2005 population and $miss$ are correlated at $\rho = -0.2$, and interestingly, the countries with the most missing data (35 or more missing values) are almost exclusively sparsely populated island nations.

[60] The Little-Rubin test of MCAR is a formal tool for assessing whether data are MCAR. It is available in SPSS but not in any of the statistical software packages on computers in the Economics Department Graduate Student Computer Lab. For that reason, I have not been able to implement this test yet.

If my factor model estimates are based on countries with above-average governance, then clearly my parameter estimates must be biased to some degree. The difficult questions are: how much, and in what direction(s)? The evidence presented so far indicates only that complete-case countries have higher mean scores than incomplete-case countries. But the covariance structure is what matters for factor model estimation, and I have not ruled out that the two samples, though they differ in means, share the same covariance structure. If they did, then my reliance on the complete-case sample would not seem to be that problematic. There are formal tests of the homogeneity of covariance structures (Kim and Bentler, 2002) that could shed more light on this question.

Clearly the complete-case analysis approach has risks. My data do not appear to be MCAR. The magnitude and direction of biases introduced by the complete-case approach however are hard to predict because they depend on the representativeness of the covariance structure found in the complete-case sample, which I have not assessed. To gain further insight into the possible biases of the complete-case approach, two alternative missing data treatments - variable reduction, and multiple imputation - are tried in the next section.

2.3 Treatments

In this section, I undertake two alternative treatments of the missing data problem and compare the corresponding results with those from complete-case analysis. Both of the alternative approaches have the effect of expanding (beyond 73) the number of countries used to estimate the factor model parameters Λ, Φ, Ψ. The first approach drops variables from the model so as to maximally increase the complese-case sample size. The second approach (multiple imputation) rectangularizes the incomplete dataset by imputing missing values with draws from the conditional distribution of the missing values (X_{miss}) given the observed values (X_{obs}). Inferences from both approaches align quite closely with those of complete-case analysis, suggesting that the exclusion of countries with incomplete data has not distorted my interpretation of the factor structure or my analysis of governance and growth in any substantial way.

2.3.1 Expanding the sample by dropping variables.

One common technique for dealing with missing multivariate data is to consider dropping variables from an analysis when: i.) those variables are highly correlated with others in the dataset, and ii.) doing so would add nontrivially to the number of cases with complete data (Hair et al., 2005). Many of the variables in my governance dataset purport to measure very similar concepts and are in fact highly positively correlated. In addition, by dropping carefully chosen subsets of these variables, it is possible to increase the number of complete cases substantially. Therefore, as a simple robustness check on the complete-case analysis utilized in Chapters one and two, I re-estimated the original varimax-rotated five-factor model on a reduced set of governance variables for which a larger number of countries had complete observations. I chose which of the 45 variables to drop from my analysis based primarily on which variables' deletion could expand the complete-case sample size the most. I tried a number of configurations, and I present below results from the most extreme version, in which I drop 20 of the 45 governance variables.

Denote the original 45-variable governance dataset by X and denote the reduced dataset of $45 - 20 = 25$ retained variables by X^{small}. The reduction from X to X^{small} expands the complete-case sample from $n = 73$ to $n^{small} = 110$ countries.[61] I estimate an orthogonal five-factor model on X^{small} and rotate loadings to the varimax criterion. Four of the five factors that emerge from this model closely resemble factors from Chapter one's model estimated on X: factor one is a close analogue to market infrastructure; factor two resembles civil liberties; factor three resembles the MIG factor; and factor four resembles order. Factor five from the X^{small} model is a hodge-podge of *stability* and *rulelaw* variables.[62] Below, I present scatterplots comparing factor score predictions from the models based on X and X^{small}. The plots show that with respect to market infrastructure, civil liberties and order, the factor model based on X^{small} ranks countries in almost exactly the same way as the model based on X. This concordance provides a measure of confidence that the inferences based on complete-case analysis in Chapters one and two are robust to the inclusion of countries with incomplete data.

[61] Whenever one variable from provider A was dropped from X, all variables from A were dropped. This is because all variables from a given provider tended to omit the same subsample of countries. Thus, dropping only one variable from, say, Global Risk Service, would not expand the number of countries with complete cases. The 20 variables dropped from X to create X^{small} comprised all variables from the following data providers (see appendix for acronyms): GRS, QLM, EIU, GCS and RSF.

[62] Because all Global Risk Service variables were dropped from X^{small}, no downside governance risk factor emerges.

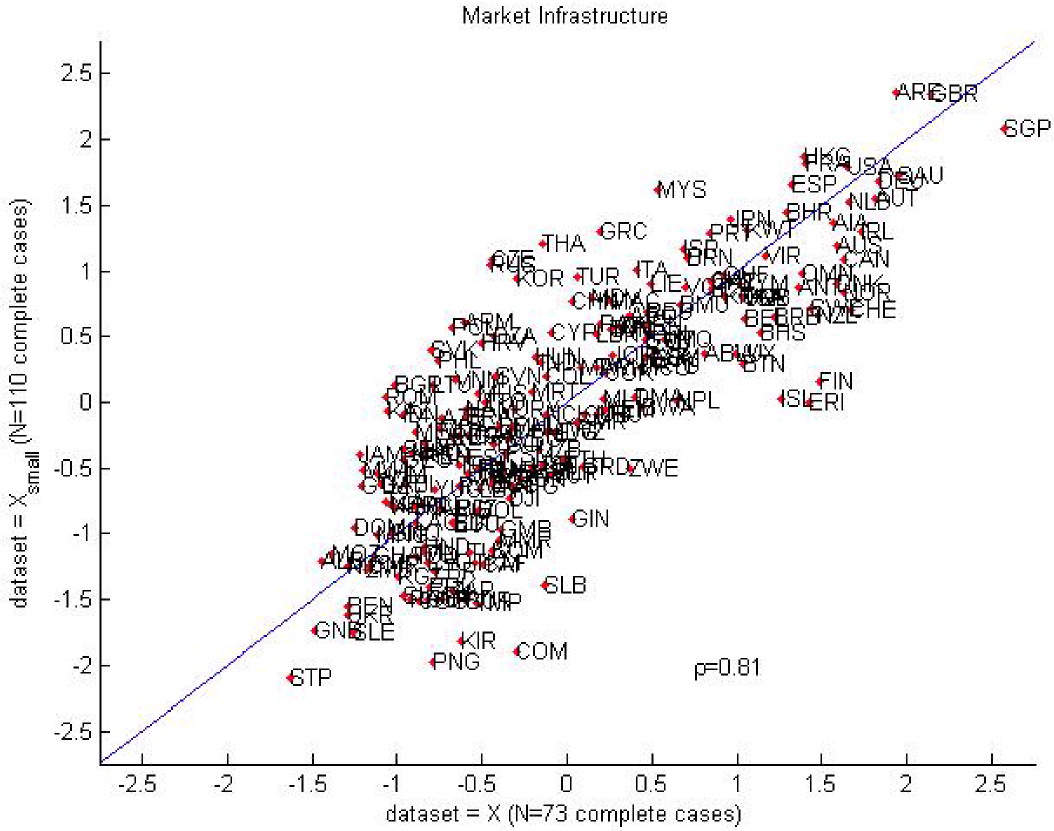
Market Infrastructure
dataset = X_{small} (N=110 complete cases)
dataset = X (N=73 complete cases)
ρ=0.81
SGP
MYS
THA
GRC
KOR
TUR
GIN
SLB
KIR
PNG
COM
STP
FIN
ERI
ZWE
BHS
BTN
NPL
ESP
HKG
CAN
CHE
AUT
NLD
IRL

Figure 4.

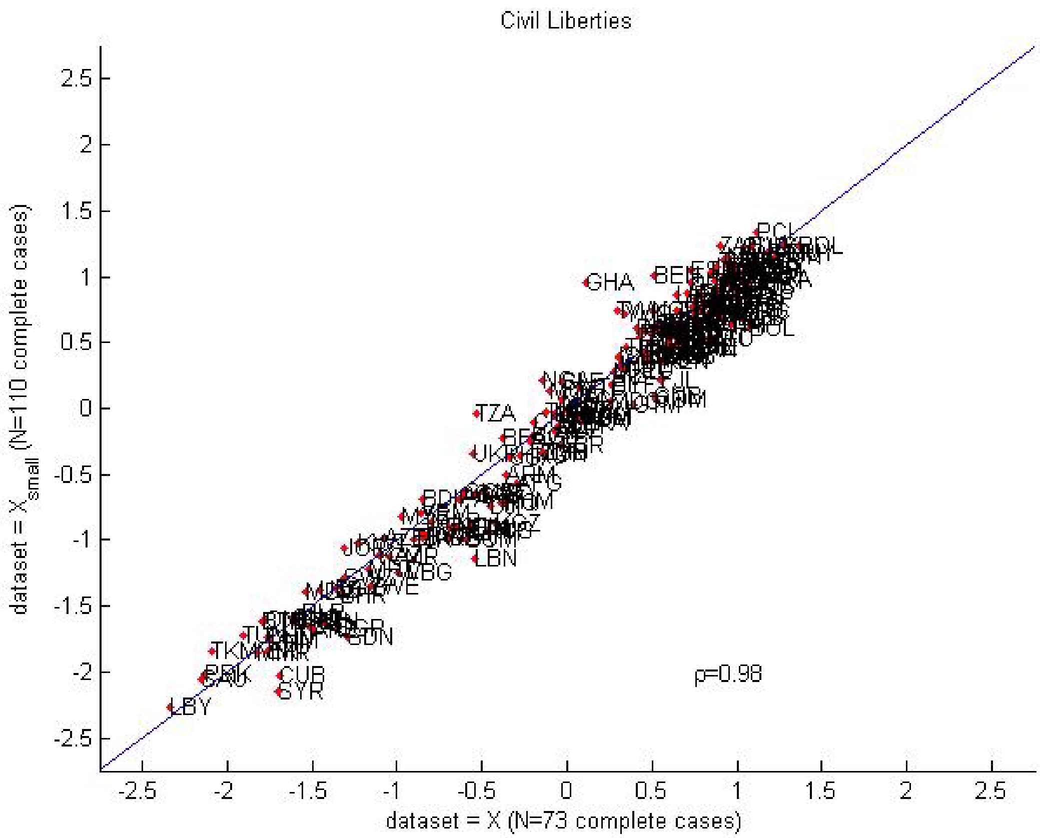
Civil Liberties
dataset = X_{small} (N=110 complete cases)
dataset = X (N=73 complete cases)
ρ=0.98

Figure 5.

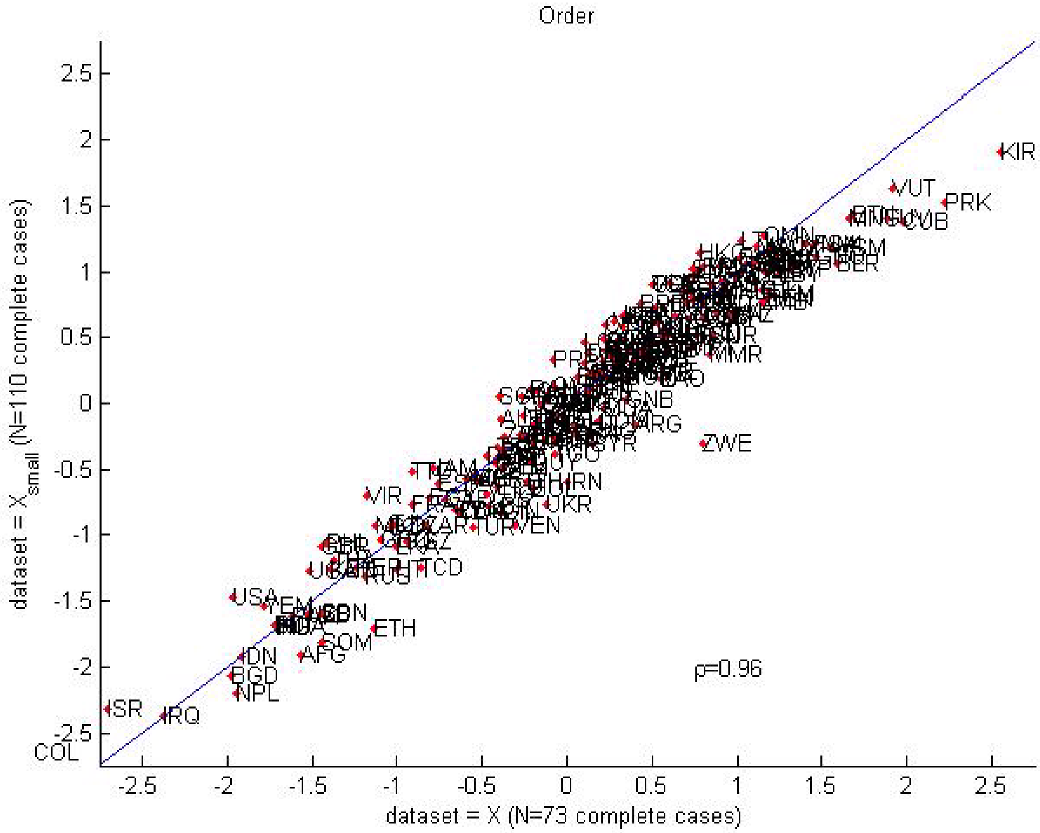

Figure 6.

I take the robustness check a step further by replicating the 2SLS growth regressions of Chapter two. The table below presents results like those of Table 17 but with the score predictions for market infrastructure, civil liberties and order taken from the X^{small} model. Not surprisingly given the close correlation between factor score predictions the two models, the results below closely resemble those of Table 17. From the second-stage results in Panel A, we see that market infrastructure and civil liberties again emerge as the dominant explanatory variables for per capita income; trade and geography are insignificant after controlling for three governance factors. Overall, coefficient magnitudes and significances are highly similar to the varimax results of Table 17. Note that the factor "downside governance risk" is not included as a regressor in Table 23 because dataset X^{small} contains no Global Risk Service variables (the variables that define downside governance risk under the full model). First-stage results from the X^{small} model are also very similar to those of the full model: distance from the equator (GEO), EURFRAC and English legal origin are most important for market infrastructure; GEO and % Muslim population are again most important for civil liberties; and Frankel-Romer tradeshare and % Muslim population best explain order.

IV (2SLS) results using limited model's factor score predictions - Rodrik et al.'s large sample

Panel A: second-stage results (Dependent variable is log 2005 per capita income.)

	(1)	(2)	(3)	(4)
GEO	0.0280***	0.0380***	0.0179**	0.00953
	(0.00818)	(0.00548)	(0.00871)	(0.0100)
INT	0.717***	0.550**	0.617***	0.0430
	(0.210)	(0.269)	(0.219)	(0.409)
mrkt. infrastructure[1]	0.484**		0.562***	0.706***
	(0.202)		(0.210)	(0.225)
civil liberties[1]		0.617***	0.642***	0.591***
		(0.159)	(0.129)	(0.131)
order[1]				0.535*
				(0.324)
Observations	138	138	138	138
R-squared	0.592	0.331	0.562	0.570
Sargan p-val	1.28E-07	0.0237	0.0132	0.0173
Cragg-Donald Wald p-val	0.0002	0	0.0001	0.0288

Panel B: first-stage results - regressing INT and governance on the instruments

	(1)	(2)	(3)	(4)
	INT	mrkt. infrastructure	civil liberties	order
GEO	0.00184	0.0373***	0.00930**	0.00379
	(0.00243)	(0.00458)	(0.00459)	(0.00468)
TRADESHARE	0.363***	-0.0187	0.0446	0.475***
	(0.0425)	(0.0802)	(0.0802)	(0.0818)
EURFRAC	-0.187*	0.434**	0.457**	0.0977
	(0.110)	(0.208)	(0.208)	(0.212)
$LEGOR_E$	0.146*	0.379**	0.221	-0.323*
	(0.0879)	(0.166)	(0.166)	(0.169)
$LEGOR_{GE}$	-0.0578	0.579*	0.302	0.0789
	(0.181)	(0.341)	(0.341)	(0.348)
$LEGOR_{SC}$	-0.240	-0.553	0.248	-0.111
	(0.218)	(0.411)	(0.411)	(0.419)
CATHO80	-0.00452	-0.243	0.417	-0.514*
	(0.148)	(0.278)	(0.278)	(0.284)
MUSLIM80	-0.264**	0.427*	-0.923***	-0.698***
	(0.128)	(0.242)	(0.242)	(0.247)
Observations	138	138	138	138
R-squared	0.400	0.453	0.377	0.259
F-stat.	10.77	13.38	9.739	5.645

Constants not shown; standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

[1]Factor score predictions from a reduced-dimension factor model (m=5, p=25, N=110, varimax rotation).

Table 23

2.3.2 Multiple imputation.

Multiple imputation (MI) is the second tactic I employ to check robustness of the complete-case analysis. Little and Rubin (2002) and Schafer (1997) provide exhaustive derivations and background on the theory and application of multiple imputation. In short, MI entails filling in the values missing from multivariate data with draws from their conditional distribution, given the observed values. The completed dataset (observed+imputed) is rectangular and suitable for standard (complete-data) estimation techniques. Of course, point estimates using imputed data would vary from one iteration of imputation to the next, since the imputed values are random draws. Such variation simply represents the uncertainty due to the missing data. To increase the precision of point estimates and to properly account for the estimate uncertainty caused by the missing data, $M > 1$ rounds of imputation are performed, yielding M completed datasets. The model of interest is estimated separately on each completed dataset and the results are then combined to yield a single, averaged, point estimate. Standard errors for the point estimate are constructed in a way that captures both the within- and across-imputation variability (Schafer, 1997).

MI is a valid approach only if the data are, at worst, MAR (Little and Rubin, 2002). While my analysis of the governance variables suggests they may not meet the threshhold of MAR, what makes MI nevertheless a potentially valid approach is the breadth of predictor variables I include in my imputation model. The role of a predictor variable in the imputation model is simply to enrich the information set upon which the conditional distribution of X_{miss} is constructed. So, while it may be that the property of missingness is still correlated the missing values themselves if one conditions only on X_{obs}, missingness may become independent of the missing values once one conditions on the wider information set $[X_{obs}, Z]$.

My predictor variables are strongly associated with governance and include distance from equator, legal origin, linguistic and ethnic fractionalization, % speaking European languages, GDP per capita, population, and trade openness. I leverage the fact that the predictors are observed completely on a wider sample of countries than is X_{obs} ($n = 154$ vs. $n = 73$). In essence, the strong association of Z with governance, plus the wide availability of Z mean that the relationships from the true joint distribution of X that would have been apparent had X_{miss} been observed can be better approximated by drawing imputations of X_{miss} from $P\left(X_{miss}|Z, X_{obs}\right)$ than from $P\left(X_{miss}|X_{obs}\right)$.

The beneficial effect of MI is to allow data from more countries to be incorporated into the estimation of the factor model of governance that is presented in Chapter one. The purpose of incorporating more countries, of course, is to examine the robustness of complete-case inferences presented in Chapters one and two. There are three stages of inference upon which the effects of MI will have an impact: i.) estimates of the factor structure itself (i.e., Λ, Φ, Ψ, and the interpretation of governance factors); ii.) predictions of the factor scores; and, iii.) 2SLS growth regression results using factor scores as regressors. One could analyze the effects of MI on each stage of results, but the results are so sequentially dependent that examining only the final stage (the 2SLS regressions) seems reasonable, and this is what I do. Clearly, if the application of MI alters the factor structure implied by Λ, Φ, Ψ substantively, then the effects should ripple through to the factor score predictions, and thereby to the growth regressions. However, if the growth regressions are basically unchanged, then it is probably safe to assume that the effects of MI on estimates of the factor structure and the factor scores has been minimal.[63]

The practical importance of this decision is that it dictates the stage of my analysis at which the M different point estimates and standard errors are to be aggregated. For instance, one could aggregate at the factor model stage, averaging the $\widehat{\Lambda}^{(t)}, \widehat{\Phi}^{(t)}, \widehat{\Psi}^{(t)}$ over t for $t = \{1, 2, ..., M\}$. Or one could could aggregate at the factor score prediction stage, averaging $\widehat{F}^{(t)} = \widehat{E}\left[f | X^{(t)}_{completed}, \widehat{\Lambda}^{(t)}, \widehat{\Phi}^{(t)}, \widehat{\Psi}^{(t)}\right]$ over t. Instead, my decision to make the final stage of analysis (the growth regressions) the focus of this robustness check implies that I should hold off combining parameter estimates and standard errors until the growth regressions stage. In other words, I carry out M factor model estimations on M completed datasets; I produce the corresponding M separate $n \times m$ matrices of factor score predictions (i.e. - each $\widehat{F}^{(t)}$ is based on $X^{(t)}_{completed}, \widehat{\Lambda}^{(t)}, \widehat{\Phi}^{(t)}, \widehat{\Psi}^{(t)}$ for $t = \{1, 2, ..., M\}$) ; and I carry out the 2SLS growth regressions M different times, using a different $\widehat{F}^{(t)}$ in the regressor matrix each time. Finally, I aggregate the M point estimates of the 2SLS second-stage regression coefficents and their standard errors according to Rubin's rules for combining completed-data estimates[64] (Rubin,

[63] Analysis of multiply imputed data entails aggregating results from M different estimates, each of which was based on a different completed dataset. In my case, analysis has three distinct stages (factor model, factor prediction, growth regression). One advantage to aggregating results at the final stage (the growth-regression stage) is that estimates of the 2SLS coefficient standard errors will explicitly account for the uncertainty resulting from between-imputation variation.

[64] The formula presented for $\overline{VC}\left(\boldsymbol{\beta}_{2sls}\right)$ reflects a small-sample (small M) adjustment as presented in Schafer (1997), pg. 114.

2002 p. 86):

$$\text{MI point estimates:} \qquad \underset{k\times 1}{\overline{\boldsymbol{\beta}}_{2sls}} = \frac{1}{M}\sum_{t=1}^{M}\widehat{\boldsymbol{\beta}}_{2sls}^{(t)} \tag{13}$$

$$\text{MI standard errors:} \qquad \underset{k\times k}{\overline{VC}\left(\widehat{\boldsymbol{\beta}}_{2sls}\right)} = (1+r_1)\,\overline{U} \tag{14}$$

$$\textit{where,}$$

$$r_1 \equiv \left(1+M^{-1}\right) tr\left(B\overline{U}^{-1}\right)/k \tag{15}$$

$$\underset{k\times k}{\overline{U}} \equiv \frac{1}{M}\sum_{t=1}^{M}\widehat{VC}\left(\boldsymbol{\beta}_{2sls}\right)^{(t)} \tag{16}$$

$$\underset{k\times k}{B} \equiv \frac{1}{M-1}\sum_{t=1}^{M}\left(\widehat{\boldsymbol{\beta}}_{2sls}^{(t)}-\overline{\boldsymbol{\beta}}_{2sls}\right)\left(\widehat{\boldsymbol{\beta}}_{2sls}^{(t)}-\overline{\boldsymbol{\beta}}_{2sls}\right)' \tag{17}$$

As can be seen in the expressions above, the MI point estimate vector is simply the average of the M completed-data point estimates. The MI standard errors (square roots of the diagonal of $\overline{VC}\left(\boldsymbol{\beta}_{2sls}\right)$) are more complicated but are in essence a combination of within- and between-imputation variation in the point estimates. Matrix B captures the between-imputation variation, while matrix $\overline{U}$ averages the within-imputation variation.

The imputation model I use is a multivariate normal regression model. For incomplete observation i, missing values are filled in with draws from the conditional distribution of i's missing variables, given i's scores in the observed variables and in the predictors. Under the assumption that all the data, $[X, Z]$ are jointly normal, this conditional distribution is also normal and is derived in a straightforward way from the estimated full joint distribution of $[X, Z]$. See Anderson (2003) p. 35 for a derivation of the conditional distribution of A given B when A and B are jointly normal. I set the number of imputations, M, to 25, which should be more than adequate (Schafer, 1997) to obtain reliable estimates.

The parameter estimates for the full joint distribution of $[X, Z]$ are arrived at iteratively, through a Markov chain Monte Carlo (MCMC) technique known as data augmentation (DA). The roles of MCMC and DA techniques in MI are covered extensively in Schafer (1997) and Rubin and Little (2002). DA consists of two steps, carried out each iteration s - an imputation step and a posterior step. In the imputation step, missing values in observation X_i are filled in with draws from $P\left(X_{miss,i}|X_{obs,i}, Z_i, \mu^{(s)}, \Sigma^{(s)}\right)$, the conditional distribution of i's missing variables, given the data observed for i and given the current iteration's parameter estimates, $\theta^{(s+1)} = \left[\mu^{(s+1)}, \Sigma^{(s+1)}\right]$. Note that $X_{miss,i}$ $(X_{obs,i})$ signifies the subset of variables in X which are missing (observed) for observation i. In the posterior step, the next iteration's parameter estimates, $\theta^{(s+1)} = \left[\mu^{(s+1)}, \Sigma^{(s+1)}\right]$, are drawn from their conditional distribution, given the newly completed dataset, which is $P\left(\theta|X_{obs}, X_{imputed}^{(s+1)}, Z\right)$.This process is repeated for S iterations, and a subset of the completed datasets (say, $X_{completed}^{(100)}, X_{completed}^{(200)}, X_{completed}^{(300)}, \ldots,$ etc.) is saved until M completed datasets have been compiled. For more detail see Schafer (1997), Chs. 3-6; Rubin and Little (2002), Ch. 10; and STATA v. 11 documentation for command *mi*.

I present my MI results in Table 24 below. The format of the table mirrors that of Table 17 in Chapter 2, but the results below come from multiple imputation, not complete-case analysis. They were derived as follows. A a total of $M = 25$ completed datasets were produced using MI. The method of impuation was a multivariate normal regression model using DA. An orthogonal factor model with five factors was then estimated on each completed dataset. Loadings from each model were rotated according to the varimax rotation. These steps resulted in 25 sets of factor model parameter estimates - one set corresponding to each completed dataset. For the tth data/parameter combination, a 215×5 matrix of factor score predictions was then generated using the Thomson predictor: $\widehat{F}^{(t)} = \widehat{E}\left[f|X^{(t)}_{completed}, \widehat{\Lambda}^{(t)}, \widehat{\Phi}^{(t)}, \widehat{\Psi}^{(t)}\right]$, for $t = \{1, 2, ..., 25\}$. Finally, the 2SLS regression specifications from Table 17 were each estimated 25 times - once with $\widehat{F}^{(1)}$as right-hand side variables, once with $\widehat{F}^{(2)}$, once with $\widehat{F}^{(3)}, ..., \widehat{F}^{(25)}$. This procedure produced 25 conditionally independent estimates of each scalar coefficient in each specification. The estimates for each coefficient were averaged and reported in the table below. Standard errors in parentheses below each point estimate were derived from matrix $\overline{VC}\left(\boldsymbol{\beta}_{2sls}\right)$, as described above.

IV (2SLS) Results with multiply imputed governance scores - Rodrik et al.'s large sample

	(1)	(2)	(3)	(4)	(5)	(6)
	Second-stage results; dependent variable is log 2005 per capita income.					
GEO	0.0197***	0.0368***	0.0153**	0.0023	0.0151**	0.0004
	(0.007)	(0.0056)	(0.0072)	(0.0093)	(0.0072)	(0.0101)
INT	0.452**	0.5691**	0.393*	0.2461	0.4203	0.0612
	(0.2096)	(0.2615)	(0.2133)	(0.2011)	(0.327)	(0.3439)
mrkt. infrastructure	0.676***		0.5997***	0.7568***	0.6026***	0.7702***
	(0.1544)		(0.158)	(0.1588)	(0.159)	(0.1666)
civil liberties		0.5123***	0.46***	0.5003***	0.4627***	0.4989***
		(0.1425)	(0.1141)	(0.0998)	(0.1146)	(0.1038)
downside governance risk				0.5592*		0.6584*
				(0.2966)		(0.3443)
order					-0.0228	0.1336
					(0.2049)	(0.2006)
Observations	138	138	138	138	138	138

1. Constants not shown; standard errors in parentheses; ***p<0.01, **p<0.05, *p<0.1.

2. Multiple imputation regression estimates based on 25 imputations of governance variables; point estimates ($\hat{\beta}^{mi}$) and standard errors ($\hat{\sigma}^{mi}_{\beta}$) constucted according) to Rubin's rules for combining completed data estimates (see Rubin and Little (2002), Schafer (1997)). Reference distribution for significance levels is standard normal; z-statistic constructed with $\hat{\beta}^{mi}/\hat{\sigma}^{mi}_{\beta}$.

3. Instruments = GEO, TRADESHARE, EURFRAC, $LEGOR_E$, $LEGOR_{GE}$, $LEGOR_{SC}$, CATHO80, MUSLIM80

Table 24

Coefficient magnitudes and patterns of significance are generally unchanged from the corresponding complete-case results in Table 17 of Chapter two. Market infrastructure and civil liberties again show the most strength in explaining cross-country income variation, while GEO and INT lose significance once three or more governance variables are included on the right-hand side. The coefficient on downside governance risk is larger and more significant than in Table 17, suggesting a more influential role for this variable than implied by complete-case analysis. The R^2, and p-values for Sargan and Cragg-Donald Wald statistics are not reported becuase it is unclear how these values should be combined from the individual imputation results.

Significance levels were arrived at by constructing a z-statistic for each 2SLS scalar coefficient β_k using its MI point estimate and standard error estimate ($z_k = \widehat{\beta}_k^{mi} / \widehat{\sigma}_{\beta_k}^{mi}$) and comparing this value to the standard normal distribution. Normality is an asymptotic approximation. In my case, this approximation may overstate the actual significance levels since the true distribution of z_k under the null may have fatter tails than a normal distribution. The reason to suspect fat tails stems from noise in the estimate of B in (17). When M is small, the sampling variation of B (i.e., its variation across different sets of M imputations) could be be very high, causing the sampling distribution of $\widehat{\sigma}_{\beta_k}^{mi}$ to be more diffuse than expected. In these circumstances, extreme values of z_k would be observed under the null more frequently than the standard normal suggests, meaning nominal significance levels would overstate actual levels. The potential severity of this distortion in my data depends on M (larger is better), the variability in $\widehat{\boldsymbol{\beta}}_{2sls}^{(t)}$ about its mean $\overline{\boldsymbol{\beta}}_{2sls}$ across imputations (less is better), and the relative magnitudes of $\overline{U}$ and B (B "small" relative to $\overline{U}$ is better). If the potential distortion from the normality assumption appears large, one could alternatively try to approximate the sampling distribution of z_k directly by using a bootstrap approach.

2.4 Conclusion on Missing Data

The two exercises above shed light on the robustness of complete-case analysis presented in Chapters one and two. Expanding the sample of complete cases by dropping variables, and using multiple imputation to incorporate data from incomplete observations both produce very similar pictures of the governance landscape. Both missing data techniques find the same factors that complete-case estimation found, and both techniques rank countries in those factors in a highly correlated order. Not surprisingly, 2SLS regressions of per capita income on the factor scores resulting from either missing data technique generate inferences that are almost identical to complete-case analysis.

The concordance of results from complete-case analysis and MI is reassuring. MI allows the use of all available data, greatly increasing the amount of information entering the factor model estimation. [65] But the concordance is not a definitive confirmation of my complete-case results. It could be that even controlling for the wide assortment of predictor variables in the imputation model (i.e., the Z), missingness is still dependent on the missing values themselves (i.e. - the data are not MAR). If the data are not MAR even after controlling for Z, then MI suffers from the same problem that afflicts complete-case analysis because imputed values will be based on an inconsistent estimate of the joint distribution of X. Estimates from the two methods

[65] In addition to the 73 complete cases, MI incorporates all observed data from the 142 countries that are incomplete in at least one of the 45 governance variables. This has the effect of doubling the number of country-variable observations to 6585 from 3285. This should lead to more efficient estimation compared to complete-case analysis. Of course, the imputation of missing values with random draws increases the level of uncertainty surrounding point estimates, but this increase is accounted for explicitly in the construction of MI standard errors.

will align, and both will be inconsistent.

Still it should be pointed out that nothing in the robustness checks performed above contradicts the hypothesis that complete-case inferences are basically sound. And one would need evidence of a particularly pernicious missingness mechanism to presume that, even after conditioning on Z and X_{obs} (the latter often coming from data providers in a similar line of work and with similar skills and incentives as the providers of the variables to be imputed), missing data are still not MAR.

3 Conclusion

The exercises performed in this chapter demonstrate that interpretation of the data's factor structure, the prediction of country factor scores, and growth-governance inferences are quite robust to alternative rotations and radically different treatments of the missing data problem.

Chapter 4

1 Introduction

In this paper I provide evidence on the robustness of factor score predictions constructed in Chapter 1 and applied in Chapter 2. I developed factor score predictions in Chapter 1 by applying ML factor model estimation to 45 governance variables, the distributions of which take on a variety of shapes and degrees of discretization. To the extent that the individual variables depart from continuity and normality, they violate the explicit assumptions behind by the ML estimator. This paper uses Monte Carlo simulations to study the effects of discretization on the quality of factor score predictions.

Social science researchers frequently encounter polytomous data that measure concepts plausibly thought of as varying continuously. One example is Likert scale data, where respondents rate their concurrence with a statement on a scale of, say, 1 to 5 (where 1 signifies "strongly disagree", and 5 signifies "strongly agree"). Other examples include categorical survey questions on age (e.g. - 20-30, 40-50, over 65), income ($< \$10,000, \$40,000$-$\$50,000, > \$100,000$) or frequency of engaging in an activity (e.g. - 1-2 times a month, weekly, daily, etc.). Survival data from medical trials may exist only in coarse increments (patient was alive at 3, 6, 9, 12, or >12 months post-treatment).

The 45 governance variables on which Chapter 1's factor model is estimated also fit this template. For example, one measure of "independence of the judiciary" covers

191 countries but takes on only three distinct values. It seems reasonable to interpret this polytomous variable as an approximation of a continuum rather than a depiction of the true distribution of "independence of the judiciary". Even when the number of categories in governance variables is higher than three, there is little reason to believe that the concepts being assessed (e.g. - "corruption", "unfair competitive practices", "property rights") truly take on only a small number of discrete levels. The unique historical, cultural and political forces at work in each country seem destined to produce governance variation of a finer level than can be described by a handful of categories.

In this paper, I assess the extent to which the discrete nature of polytomous data confounds prediction of common factors generating the underlying continuous data. I want to know, for example, the correlation between predictions of "market infrastructure" based on three-category polytomous variables and predictions of "market infrastructure" based on the continuous versions of the same variables. To estimate these correlations, I conduct Monte Carlo experiments in which normally distributed continuous random variables, generated by a standard factor-model data-generating process (DGP), undergo a transformation to polytomous variables. The transformed variables are then factor analyzed using ML, and the parameter estimates so derived are used to construct predictions of the factor scores that gave rise to the original (untransformed) data. These score predictions are then compared to the (infeasible) predictions based on the original (untransformed) data.

I expect the two sets of predictions to differ for two reasons. First, a polytomous variable distinguishes true performance coarsely and unevenly; countries with insufficiently large differences in true performance are assigned the same score (unless they happen to fall on opposite sides of a cutoff). Second, depending on the placement of the category cutoffs, the distribution of the polytomous variable can be highly non-normal even if the underlying continuous variable is normal. For example, if the upper cutoff for the lowest score category is near the mean of the continuous variable's distribution, then the polytomous variable may exhibit large positive skew.

The psychometric and statistical literatures have extensively studied factor model estimation in the presence of both polytomous data (e.g. - Bartholomew, 1980; Mislevy, 1986; Dolan, 1994) and censored data (e.g. - Muthen, 1989; Waller and Muthen, 1991; Kamakura and Wedel, 2001). A common mode of inquiry is to examine the relative performance of different factor model estimators (e.g. - ML, weighted least-squares (WLS) and generalized least-squares (GLS)) in terms of: i.) variability and bias in parameter estimates (factor loadings, uniqueness variances), and ii.) deviations in the actual distribution of the likelihood-ratio fit statistic, which follows a χ^2 distribution under conforming data. Less attention, however, has been paid to the impact of censoring and discretization on factor score prediction (which involves complicated functions of the parameter estimates) - the focus of this paper. I study only the normal-theory ML factor model estimator because that is the estimator used in Chapter 1, and it is specifically the robustness of Chapter 1's predictions that I wish to assess.

Shi and Lee (1997) use Monte Carlo experiments to examine the performance of a Bayesian factor score predictor with censored, truncated and polytomous data. Their simulated datasets consist of a mix of continuous and polytomous variables. Accuracy of the Bayesian factor predictor was found to increase in the number of categories (for polytomous variables), and in the proportion of all variables that were continuous. The pattern of censoring (symmetric vs. asymmetric thresholds) was found to have little impact.

This paper departs from Shi and Lee in a number of ways. First, I study a more common approach to factor model prediction: Thomson's (1951) regression predictor using ML parameter estimates. Second, I vary independently the pattern of censoring (location and symmetry of upper and lower cutoffs), and the coarseness of discretization for polytomous variables. These two design features were linked in the experimental setup of Shi and Lee, so the independent effects of censoring and discretization were not directly assessed. Third, I examine factor score prediction under a wider variety of DGPs. Shi and Lee used the same simple oblique two-factor DGP for all experiments. In this paper, however, I examine prediction performance under one-, two- and three-factor models with simple and complex loading patterns.

2 Interpreting Polytomous Governance Data

The specific motivation of this paper is the use of factor models to analyze expert opinion assessments of country governance. In Chapter 1, I estimate an orthogonal multi-factor model on a dataset of 45 such variables using ML and uncover four strongly defined factors of governance - market infrastructure, civil liberties, downside governance risk, and order. I then estimate each country's score in each of the four factors using the regression predictor (Thomson, 1951).

Many of the 45 variables used in that estimation are polytomous in nature.[66] Many variables also exhibit distributions with distinctly non-normal shapes - including bi-modal, negative skew, positive skew and U-shaped. I present empirical histograms of these variables in the Appendix. The concern taken up by this paper is the extent to which a country's predicted factor scores based on this data may deviate from its true factor score.

Before taking up that question though, I need to motivate my basic assumptions about the nature of the data. How do I know governance actually varies continuously across countries? And if it does, what explains the polytomous, non-normal distributions of observed governance variables?

I cannot prove that the governance dimensions assessed by the data vary continuously, but I can try to show why this assumption is the most plausible. First, some governance variables are literally probability assessments, and the natural range of these is the continuous interval from 0 to 1. Second, other variables assess attributes such as the quality of the bureaucracy, freedom of expression, or the impartiality of the legal system. Given the complexity of cultural, economic and political influences shaping these outcomes, it seems implausible that no detail is lost by representing their cross-country variation with a small number of categories. For the same reason, Likert scale variables in the dataset (e.g. - agree/disagree on a scale of 1-7 with: "The threat of terrorism ...imposes significant costs on business") can probably be regarded as simplifications of the true variation in the attribute they measure.

[66] The coarseness of discretization varies widely across the 45 variables. Six variables take on nine or fewer values (3,5,5,7,8,9), while the remaining 39 each take on at least ten. Eighteen variables take on 30 or more unique values.

Why are governance data polytomous if the concepts they assess are continuous? The explanation I propose is limited powers of discernment. I suggest that country experts have constrained abilities to discern a country's actual level of performance. Experts are able to group countries into ordered categories, but that's all. One could think of experts as only observing the integer portion of a country's true score. I could also draw an analogy to a scale that registers increments no finer than a kilogram and which has an upper limit of, say M kg. The scale is incapable of precisely ranking objects by mass, although it can order them approximately so. And for objects weighing M kg or more, the scale provides only a lower bound on the object's mass. This is the type of mechanism I have in mind.

Depending on the coarseness of the discretization, and on the placement of the cutoffs for the lower-most and upper-most score categories, the resulting cross-country distribution of an expert's assessments could take a wide variety of shapes. For example, an expert who can distinguish equally well between countries at the bottom, the middle and the top of the governance distribution would transform a symmetrically distributed unimodal continuous variable into a symmetrically distributed, unimodal polytomous variable. On the other hand, an expert who is capable of making finer distinctions between countries at the top of the governance distribution than at the bottom would transform a symmetrically distributed continuous variable into a polytomous variable with positive skew (long right tail). Finally, an expert incapable of making fine distinctions at either extreme could generate a polytomous variable with a U-shaped distribution.

While I think I've identified a plausible explanation for the observed data, there is no way to prove that it is the correct explanation. Other reasons may explain why one observes skewed, polytomous distributions even though the underlying governance concepts are normal and continuous. For instance, non-normality could have arisen because disturbances in the underlying factor model follow a skewed or bimodal distribution. And polytomous data may reflect the preferences of clients rather than the cognitive limitations of country experts.

But it's important to point out which assumptions matter, and which do not matter for my results. The assumptions that do matter are that polytomous variables reflect continuous variables; that the continuous variables are normally distributed; and that the covariance structure of the continuous variables exhibit a factor structure. If these assumptions are not met by the DGP behind the governance data, then the results from this paper's simulation exercises do not offer valid insights into the estimates of Chapter 1.

On the other hand, assumptions about how the continuous variables are transformed into polytomous variables do not matter for inference because nothing in the ML estimation procedure presumes knowledge of this process. The crucial thing is that the proposed explanation generate data with similar observable properties as the governance variables entering the factor model in Chapter 1. My explanation does that.

3 The Model

I now turn to the formal model of expert assessment data. I review the standard factor model briefly, reproducing the exposition first presented in Chapter 1 for convenience. I then introduce some new notation to describe the polytomous nature of the data.

Let $X = [X_1, X_2, ...X_k, ...X_p]'$ represent a vector of normally distributed random variables. The factor analysis model decomposes this p-dimensional vector in the following way:

$$\underset{p\times 1}{X} = \underset{p\times m}{\Lambda}\underset{m\times 1}{f} + \underset{p\times 1}{u} \tag{18}$$

Matrix Λ is called the loadings matrix; in an orthogonal factor model its (i, j)-th entry is the covariance between the i-th variable in X and the j-th common factor. When λ_{ij} is large and positive, variable X_i is said to load heavily on factor f_j. Vector f is the vector of $m < p$ (unobserved) common factors, modeled as a random vector with density $f \sim N(\mathbf{0}, \Phi)$. Random disturbance vector u is iid with density $N(\mathbf{0}, \Psi)$, Ψ diagonal.

For clarity, I assume all elements of X have been centered and scaled such that $E(X_k) = 0$ and $sd(X_k) = 1, \forall k \in \{1, 2, ..., p\}$. It is assumed that all m factors are independent of all p disturbances (otherwise, disturbance u_k would not represent variable X_k's idiosyncratic variation).

Given equation (1) and the stochastic assumptions stated above, one can write Σ, the VC matrix of X, as follows:

$$\Sigma = \Lambda\Phi\Lambda' + \Psi \tag{19}$$

Equation (2) is the fundamental hypothesis of the factor model. The factor model posits that Σ can be decomposed into the sum of a symmetric, positive definite matrix of rank $m < p$ (i.e. - $\Lambda\Phi\Lambda'$), and a diagonal, positive definite matrix of rank p (i.e. - Ψ). Matrix $\Lambda\Phi\Lambda'$ represents the variation in X that is due to factors common to all elements of X, while matrix Ψ captures the variation in each $X_k \in \{X_1, ..., X_p\}$ that is idiosyncratic.

To predict f, one constructs its conditional mean, given data vector X. Such a predictor is called the regression predictor and is due to Thomson (1951). Anderson (2003) and Lawley and Maxwell (1971) provide full derivations of the regression predictor. Here I only reproduce the final result, noting that it follows directly from the joint distribution of f and X implied by (1), (2) and accompanying stochastic assumptions.

The $m \times 1$ population conditional mean vector of f, given observed data X can be written:

$$E[f|X] = \Phi\Lambda'\Sigma^{-1}X \tag{20}$$

Since in practice the parameters Λ, Φ, Ψ are unknown, their ML estimates are used. Designate the estimated conditional mean and of f as:

$$\widehat{E}[f|X] = \widehat{\Phi}\widehat{\Lambda}'\widehat{\Sigma}^{-1}X \tag{21}$$

Unfortunately, X is not observed. Rather, a polytomous version of X that I designate X^* is observed.A realization of scalar variable $X_k^* \in \left\{X_1^*, ..., X_p^*\right\}$ takes on one of T discrete values, $\{1, 2, ..., T\}$, according to the following rule:

$$X_k^* = t, \text{ if } c_{t-1} < X_k \leq c_t \text{ for } t \in \{1, 2, ..., T\} \tag{22}$$

where the c are cutoffs for the score categories and where I define $c_0 = -\infty$, and $c_T = \infty$. To simplify things, I will assume that $c_1, ..., c_{T-1}$ are all equally spaced and that all p variables in X^* have the same cutoff points $c_0, ..., c_T$.[67]

4 Monte Carlo Design

In this section I describe the design of my Monte Carlo experiments. The basic procedure involves estimating identically specified factor models on two parallel samples - a continuous-data sample and a polytomous version of that sample - and comparing the factor predictions from each. My experimental design draws on numerous papers in the psychometric literature (e.g. - Hakstian et al., 1982; Muthen, 1989; Dolan, 1994; Shi and Lee, 1997).

The ideal design for any Monte Carlo study depends on the nature of the question being investigated and on the desired generality of the results. Fundamental elements to consider in any factor model simulation include the number of factors, the correlation of factors, the complexity of loading patterns, magnitude of communalities (the proportion of each X_k's variation attributable to the common factors), the number of variables, and the sample size. Further design elements that seem important for

[67] The latter assumption is contradicted by our data. Both the number and (evidently) the location of cutoffs differ across variables. However we can extrapolate from our results based on homogeneous cutoffs to datasets with mixed cutoffs.

my specific question include the number of discrete values that polytomous variables may take on, and the proportion of the underlying continuous distribution censored by the discretization process.

Varying all of these elements independently in one experiment would yield a comprehensive but also unwieldy body of results. I pursue a narrower design tailored to assess factor prediction accuracy under conditions resembling those of Chapter 1. My study is not designed to uncover in full generality the impact of polytomous data on the ML-based factor score predictor. Rather, my goal is to be able to say something informative about the sensitivity of factor score predictions from estimations like those of Chapter 1. My tactic is to simplify and extrapolate; I experiment on smaller-dimensional models and use stylized loading patterns and homogeneous patterns of discretization to gain a clear sense of how these model attributes affect factor prediction.

I now describe the Monte Carlo procedure in detail. The procedure begins by specifying the DGP. The DGP comprises the factor model parameters (Λ, Φ, Ψ) as well as the number of polytomous categories (T) and the pattern of upper and lower category cutoffs $((a, b))$.

I look exclusively at orthogonal-factor models and therefore restrict $\Phi = I$ in all simulations. I examine one-,two-,three- and five-factor models. For two- and three-factor models, I experiment with three different loading patterns. The three patterns differ in complexity, i.e. - in the number of factors on which each variable has a nonzero loading, or in the frequency of negative loadings. In the first loading pattern (the "simple" pattern), each variable in X has a nonzero loading on exactly one factor, and every factor has exactly seven variables that load on it. In the second loading pattern (the "complex" pattern), each variable in X has nonzero loadings on all factors but loads most heavily on one dominant factor. Every factor has exactly seven variables for which it is the dominant factor. In the third loading pattern (the "complex and negative" patterns), each variable in X loads heavily on a dominant factor but also exhibits small positive, zero or negative loadings on the remaining factors.

In my five-factor model, I consider a single loading pattern specially constructed to mimic characteristics of the varimax-rotated estimate of Λ from the actual governance data. The characteristics I attempt to match with my constructed Λ are: 45 variables and five factors; factor communalities for $f_1, .., f_5$ of $0.41, 0.24, 0.15, 0.12, 0.08$, respectively; and 83% of total variation explained by all factors. All variants of Λ, Ψ used in my simulations are displayed in the appendix.

With the DGP specified, I draw n observations on f and u from $N(0, I_m)$ and $N(0, \Psi)$, respectively and use them to construct n observations on X in accordance with (1). I fix n at 75 in all experiments to mimic the size of the complete-case sample in Chapter 1 (where $n = 73$). Continuous data, X, are then transformed into T-category polytomous data, X^*, using the transformation rule in (5). I let T take values in $\{2, 3, 7, 15, 25, 50, 250, 500\}$.

The exact placement of the category cutoffs, c, in (5) is determined by the parameters (a, b), which denote percentiles of the marginal densities of the X_k. For example, the values $T = 3$, and $(a, b) = (0.1, 0.9)$ establish the following cutoffs for polytomous variable X_k^*: $[c_0, c_1, c_2, c_3] = [-\infty, z_{0.1}, z_{0.9}, \infty]$ where z_α is the value corresponding to the α-th percentile of X_k's marginal density (the standard normal density, in all my simulations). Thus, I would have $X_k^* = $ "1" for $X_k \leq -1.28$; $X_k^* = $ "2" for $-1.28 < X_k \leq 1.28$; and finally, $X_k^* = $ "3" for $X_k > 1.28$. I allow (a, b) to take on values in $\{(0.1, 1), (0.5, 1), (0.75, 1), (0.4, 0.6), (0.1, 0.9)\}$. Each pair (a, b) implies a certain manifestation of an expert's limited powers of discernment. The pair $(0.75, 1)$, for example, induces the extreme left-censoring of X_k one would expect to see if expert k perceived all countries in the lower 75% of the true governance distribution as having equivalent governance. Likewise, the pair $(0.1, 0.9)$ induces mild right- and left-censoring of X_k consistent with an expert incapable of discriminating between countries within the bottom 10% of the true governance distribution or between countries within the top 10% of that distribution.

Let the full $n \times p$ datasets of continuous and polytomous observations be denoted $\mathbf{X}$ and $\mathbf{X}^*$ I estimate identically specified factor models on $\mathbf{X}$ and $\mathbf{X}^*$ using maximum likelihood.[68] (For multi-factor models, I rotate to the varimax solution.) Denote the parameter estimates based on $\mathbf{X}$ as $\widehat{\theta} = \left(\widehat{\Lambda}, \widehat{\Psi}\right)$ and the parameter estimates based $\mathbf{X}^*$ as $\widehat{\theta}^* = \left(\widehat{\Lambda}^*, \widehat{\Psi}^*\right)$. I then compare factor score predictions based on $\widehat{\theta}$ with those based on $\widehat{\theta}^*$.For a particular pair of observations X and X^*, the predictions being compared are (I_m substituted for Φ):

$$\underset{(m \times 1)}{\widehat{E}[f|X]} = \widehat{\Lambda}' \widehat{\Sigma}^{-1} X \tag{23}$$

$$\underset{(m \times 1)}{\widehat{E}[f|X^*]} = \widehat{\Lambda}^{*\prime} \widehat{\Sigma}^{*-1} X^* \tag{24}$$

I look at two related measures of correspondence between (6) and (7): mean-

[68] The presumed number of factors in the estimated model is always set equal to the actual number of factors in the DGP.

squared difference and correlation. The mean-squared difference in predictors for the jth factor, for a particular sample is:

$$msd_j \equiv \frac{1}{n}\sum_{i=1}^{n}\left(\widehat{E}\left[f_j|X^*_{(i)}\right] - \widehat{E}\left[f_j|X_{(i)}\right]\right)^2, \;\; j \in \{1, ..., m\} \tag{25}$$

where $X_{(i)}$ and $X^*_{(i)}$ are the ith observations on X and X^*, respectively. The correlation between predictors for the jth factor, for a particular sample, is:

$$\rho_j \equiv \frac{1}{n}\sum_{i=1}^{n}\left(\widehat{E}\left[f_j|X^*_{(i)}\right] \cdot \widehat{E}\left[f_j|X_{(i)}\right]\right), \;\; j \in \{1, ..., m\} \tag{26}$$

A DGP is defined by specific values for Λ, Φ, Ψ, T,and (a, b). For each DGP, I repeat the above steps 100 times, using 100 independently drawn samples. I save the msd_j and the ρ_j resulting from each sample. I then average the values of the msd_j and ρ_j across the 100 samples and display these average values ($\overline{msd}_j = \frac{1}{100}\sum_{i=1}^{100} msd_{j(i)}; \overline{\rho}_j = \frac{1}{100}\sum_{i=1}^{100}\rho_{j(i)}$ where i indexes the sample) in Tables 1-8.

The values of $\overline{mse}$ and $\overline{\rho}$ measure the average loss in precision attributable to discretization. Literally, they measure the correspondence between factor predictions from a model that uses conforming data and a model that uses coarse approximations of that data.

It might seem more logical to compare factor predictors to the true factor scores (i.e. - compare $\widehat{E}\left[f_j|X^*\right]$ and $\widehat{E}\left[f_j|X\right]$ to f) rather than to each other. For two reasons, I don't report those results here. First, the goal of my simulations is to quantify the loss in predictive accuracy due to a specific deviation from the ML estimator assumptions. The predictions based on data conforming to those assumptions (i.e. - $\widehat{E}[f|X]$) therefore constitute a natural benchmark. Second, I have examined the pairwise correlations between predictors and true factor scores and have found, not surprisingly, that $\overline{\rho}\left(\widehat{E}[f_j|X], f_j\right) > \overline{\rho}\left(\widehat{E}[f_j|X^*], f_j\right), \forall j$ always. That is, in no case does the polytomous-data factor predictor outperform the continuous-data predictor. Therefore, a low correlation between the two predictors will in general reflect the polytomous-data predictor's low correlation with the true factor scores.

5 Simulation Results

I present my simulation results in Tables 25-32 below. Discussion of these results follows the tables.

5.1 Tables

Simulation Results - Simple One-Factor Models

	$E[(\hat{f}^* - \hat{f})^2]$	$\rho(\hat{f}^*, \hat{f})$		$E[(\hat{f}^* - \hat{f})^2]$	$\rho(\hat{f}^*, \hat{f})$
(a,b)	f_1	f_1	(a,b)	f_1	f_1
	(categories = 2)			*(categories = 3)*	
(0.1,1)	0.55	0.67	(0.1,1)	0.08	0.95
(0.5,1)	0.15	0.91	(0.5,1)	0.10	0.94
(0.75,1)	0.24	0.85	(0.75,1)	0.29	0.83
(0.4,0.6)	-	-	(0.4,0.6)	0.11	0.94
(0.1,0.9)	-	-	(0.1,0.9)	0.16	0.90
	(categories = 7)			*(categories = 15)*	
(0.1,1)	0.02	0.99	(0.1,1)	0.02	0.99
(0.5,1)	0.12	0.93	(0.5,1)	0.15	0.91
(0.75,1)	0.35	0.79	(0.75,1)	0.40	0.76
(0.4,0.6)	0.11	0.94	(0.4,0.6)	0.11	0.93
(0.1,0.9)	0.02	0.99	(0.1,0.9)	0.02	0.99
	(categories = 25)			*(categories = 50)*	
(0.1,1)	0.02	0.99	(0.1,1)	0.02	0.99
(0.5,1)	0.15	0.91	(0.5,1)	0.17	0.90
(0.75,1)	0.40	0.76	(0.75,1)	0.40	0.76
(0.4,0.6)	0.12	0.93	(0.4,0.6)	0.11	0.93
(0.1,0.9)	0.02	0.99	(0.1,0.9)	0.02	0.99
	(categories = 250)			*(categories = 500)*	
(0.1,1)	0.02	0.99	(0.1,1)	0.02	0.99
(0.5,1)	0.17	0.90	(0.5,1)	0.18	0.89
(0.75,1)	0.42	0.75	(0.75,1)	0.41	0.75
(0.4,0.6)	0.12	0.93	(0.4,0.6)	0.12	0.93
(0.1,0.9)	0.02	0.99	(0.1,0.9)	0.02	0.99

Design parameters: p=7, n=75, lambda=0.7, psi=0.51, 100 samples per DGP

Table 25

Simulation Results - Simple Two-Factor Models

	$E[(\hat{f}^* - \hat{f})^2]$		$\rho(\hat{f}^*, \hat{f})$			$E[(\hat{f}^* - \hat{f})^2]$		$\rho(\hat{f}^*, \hat{f})$	
(a,b)	f_1	f_2	f_1	f_2	(a,b)	f_1	f_2	f_1	f_2
	(categories = 2)					*(categories = 3)*			
(0.1,1)	0.60	0.58	0.64	0.65	(0.1,1)	0.08	0.08	0.95	0.95
(0.5,1)	0.16	0.16	0.90	0.90	(0.5,1)	0.11	0.11	0.94	0.94
(0.75,1)	0.27	0.28	0.84	0.83	(0.75,1)	0.27	0.28	0.84	0.84
(0.4,0.6)	-	-	-	-	(0.4,0.6)	0.11	0.11	0.94	0.94
(0.1,0.9)	-	-	-	-	(0.1,0.9)	0.17	0.17	0.90	0.90
	(categories = 7)					*(categories = 15)*			
(0.1,1)	0.03	0.02	0.99	0.99	(0.1,1)	0.02	0.02	0.99	0.99
(0.5,1)	0.14	0.13	0.92	0.92	(0.5,1)	0.15	0.15	0.91	0.91
(0.75,1)	0.35	0.35	0.79	0.79	(0.75,1)	0.40	0.38	0.77	0.78
(0.4,0.6)	0.12	0.12	0.93	0.93	(0.4,0.6)	0.12	0.11	0.93	0.93
(0.1,0.9)	0.02	0.02	0.99	0.99	(0.1,0.9)	0.02	0.02	0.99	0.99
	(categories = 25)					*(categories = 50)*			
(0.1,1)	0.03	0.02	0.99	0.99	(0.1,1)	0.02	0.03	0.99	0.98
(0.5,1)	0.17	0.17	0.90	0.90	(0.5,1)	0.17	0.17	0.90	0.90
(0.75,1)	0.40	0.41	0.76	0.76	(0.75,1)	0.43	0.40	0.75	0.76
(0.4,0.6)	0.12	0.12	0.93	0.93	(0.4,0.6)	0.12	0.12	0.93	0.93
(0.1,0.9)	0.02	0.02	0.99	0.99	(0.1,0.9)	0.02	0.02	0.99	0.99
	(categories = 250)					*(categories = 500)*			
(0.1,1)	0.02	0.02	0.99	0.99	(0.1,1)	0.02	0.02	0.99	0.99
(0.5,1)	0.17	0.18	0.90	0.90	(0.5,1)	0.17	0.17	0.90	0.90
(0.75,1)	0.43	0.43	0.75	0.75	(0.75,1)	0.43	0.42	0.75	0.75
(0.4,0.6)	0.12	0.12	0.93	0.93	(0.4,0.6)	0.12	0.12	0.93	0.93
(0.1,0.9)	0.02	0.02	0.99	0.99	(0.1,0.9)	0.02	0.02	0.99	0.99

p=14, n=75, lambda=0.7 or 0, psi=0.51, 100 samples per DGP

Table 26

Simulation Results - Simple Three-Factor Models

(a,b)	$E[(\hat{f}^* - \hat{f})^2]$ f_1	f_2	f_3	$\rho(\hat{f}^*, \hat{f})$ f_1	f_2	f_3	(a,b)	$E[(\hat{f}^* - \hat{f})^2]$ f_1	f_2	f_3	$\rho(\hat{f}^*, \hat{f})$ f_1	f_2	f_3
			(categories = 2)							*(categories = 3)*			
(0.1,1)	0.64	0.62	0.68	0.62	0.63	0.60	(0.1,1)	0.09	0.09	0.09	0.95	0.95	0.95
(0.5,1)	0.17	0.17	0.16	0.90	0.90	0.90	(0.5,1)	0.12	0.11	0.11	0.93	0.94	0.94
(0.75,1)	0.28	0.28	0.29	0.83	0.83	0.83	(0.75,1)	0.29	0.27	0.28	0.83	0.84	0.84
(0.4,0.6)	-	-	-	-	-	-	(0.4,0.6)	0.11	0.11	0.11	0.94	0.93	0.93
(0.1,0.9)	-	-	-	-	-	-	(0.1,0.9)	0.18	0.19	0.19	0.89	0.89	0.89
			(categories = 7)							*(categories = 15)*			
(0.1,1)	0.09	0.09	0.09	0.98	0.98	0.98	(0.1,1)	0.02	0.02	0.02	0.99	0.99	0.99
(0.5,1)	0.12	0.11	0.11	0.92	0.92	0.92	(0.5,1)	0.15	0.16	0.15	0.91	0.91	0.91
(0.75,1)	0.29	0.27	0.28	0.78	0.79	0.79	(0.75,1)	0.41	0.40	0.41	0.76	0.76	0.75
(0.4,0.6)	0.11	0.11	0.11	0.93	0.93	0.93	(0.4,0.6)	0.12	0.12	0.12	0.93	0.93	0.93
(0.1,0.9)	0.18	0.19	0.19	0.99	0.99	0.99	(0.1,0.9)	0.02	0.02	0.02	0.99	0.99	0.99
			(categories = 25)							*(categories = 50)*			
(0.1,1)	0.03	0.02	0.03	0.98	0.99	0.98	(0.1,1)	0.03	0.03	0.02	0.98	0.99	0.99
(0.5,1)	0.16	0.17	0.16	0.90	0.90	0.90	(0.5,1)	0.17	0.18	0.17	0.90	0.90	0.90
(0.75,1)	0.41	0.45	0.42	0.76	0.74	0.75	(0.75,1)	0.43	0.44	0.43	0.75	0.74	0.74
(0.4,0.6)	0.12	0.12	0.13	0.93	0.93	0.93	(0.4,0.6)	0.12	0.13	0.12	0.93	0.92	0.93
(0.1,0.9)	0.02	0.02	0.02	0.99	0.99	0.99	(0.1,0.9)	0.02	0.02	0.02	0.99	0.99	0.99
			(categories = 250)							*(categories = 500)*			
(0.1,1)	0.02	0.02	0.02	0.99	0.99	0.99	(0.1,1)	0.02	0.02	0.02	0.99	0.99	0.99
(0.5,1)	0.18	0.18	0.18	0.89	0.90	0.90	(0.5,1)	0.18	0.18	0.18	0.89	0.90	0.89
(0.75,1)	0.45	0.45	0.45	0.74	0.74	0.74	(0.75,1)	0.44	0.46	0.43	0.74	0.73	0.75
(0.4,0.6)	0.13	0.12	0.12	0.93	0.93	0.93	(0.4,0.6)	0.13	0.12	0.13	0.92	0.93	0.92
(0.1,0.9)	0.02	0.02	0.03	0.99	0.99	0.99	(0.1,0.9)	0.02	0.03	0.02	0.99	0.99	0.99

p=21, n=75, lambda=0.7 or 0, psi=0.51, 100 samples per DGP

Table 27

Simulation Results - Complex Two-Factor Models

	$E[(\hat{f}^* - \hat{f})^2]$		$\rho(\hat{f}^*, \hat{f})$			$E[(\hat{f}^* - \hat{f})^2]$		$\rho(\hat{f}^*, \hat{f})$	
(a,b)	f_1	f_2	f_1	f_2	(a,b)	f_1	f_2	f_1	f_2
	(categories = 2)					*(categories = 3)*			
(0.1,1)	0.79	0.83	0.54	0.52	(0.1,1)	0.24	0.22	0.85	0.86
(0.5,1)	0.37	0.36	0.78	0.78	(0.5,1)	0.23	0.24	0.86	0.85
(0.75,1)	0.52	0.52	0.69	0.69	(0.75,1)	0.50	0.51	0.70	0.69
(0.4,0.6)	-	-	-	-	(0.4,0.6)	0.26	0.26	0.84	0.84
(0.1,0.9)	-	-	-	-	(0.1,0.9)	0.43	0.44	0.74	0.74
	(categories = 7)					*(categories = 15)*			
(0.1,1)	0.07	0.08	0.95	0.95	(0.1,1)	0.07	0.06	0.96	0.96
(0.5,1)	0.25	0.26	0.85	0.84	(0.5,1)	0.30	0.28	0.82	0.83
(0.75,1)	0.58	0.55	0.66	0.67	(0.75,1)	0.64	0.62	0.62	0.64
(0.4,0.6)	0.28	0.29	0.83	0.83	(0.4,0.6)	0.28	0.27	0.83	0.84
(0.1,0.9)	0.04	0.05	0.97	0.97	(0.1,0.9)	0.06	0.05	0.96	0.97
	(categories = 25)					*(categories = 50)*			
(0.1,1)	0.08	0.07	0.95	0.96	(0.1,1)	0.08	0.08	0.95	0.95
(0.5,1)	0.29	0.29	0.83	0.83	(0.5,1)	0.30	0.30	0.82	0.82
(0.75,1)	0.62	0.65	0.63	0.62	(0.75,1)	0.65	0.64	0.62	0.62
(0.4,0.6)	0.29	0.29	0.82	0.82	(0.4,0.6)	0.28	0.31	0.83	0.81
(0.1,0.9)	0.05	0.06	0.97	0.96	(0.1,0.9)	0.05	0.06	0.97	0.96
	(categories = 250)					*(categories = 500)*			
(0.1,1)	0.05	0.04	0.97	0.97	(0.1,1)	0.06	0.06	0.96	0.96
(0.5,1)	0.29	0.29	0.83	0.82	(0.5,1)	0.30	0.30	0.82	0.82
(0.75,1)	0.64	0.65	0.63	0.62	(0.75,1)	0.65	0.64	0.62	0.62
(0.4,0.6)	0.30	0.30	0.82	0.82	(0.4,0.6)	0.30	0.32	0.82	0.81
(0.1,0.9)	0.05	0.05	0.97	0.97	(0.1,0.9)	0.07	0.06	0.96	0.97

p=14, n=75, lambda=0.7 or 0.4, psi=0.4, 100 samples per DGP

Table 28

Simulation Results - Complex Three-Factor Models

	$E[(\hat{f}^{*}-\hat{f})^2]$			$\rho(\hat{f}^{*},\hat{f})$				$E[(\hat{f}^{*}-\hat{f})^2]$			$\rho(\hat{f}^{*},\hat{f})$		
(a,b)	f_1	f_2	f_3	f_1	f_2	f_3	(a,b)	f_1	f_2	f_3	f_1	f_2	f_3
			(categories = 2)							*(categories = 3)*			
(0.1,1)	0.91	0.91	0.91	0.48	0.48	0.48	(0.1,1)	0.25	0.26	0.23	0.85	0.84	0.86
(0.5,1)	0.42	0.43	0.43	0.75	0.75	0.75	(0.5,1)	0.26	0.26	0.25	0.84	0.85	0.85
(0.75,1)	0.57	0.60	0.57	0.66	0.65	0.66	(0.75,1)	0.53	0.53	0.52	0.69	0.69	0.69
(0.4,0.6)	-	-	-	-	-	-	(0.4,0.6)	0.27	0.28	0.27	0.84	0.83	0.84
(0.1,0.9)	-	-	-	-	-	-	(0.1,0.9)	0.52	0.47	0.45	0.69	0.72	0.73
			(categories = 7)							*(categories = 15)*			
(0.1,1)	0.07	0.07	0.06	0.96	0.96	0.96	(0.1,1)	0.06	0.06	0.06	0.97	0.97	0.96
(0.5,1)	0.25	0.26	0.24	0.85	0.85	0.86	(0.5,1)	0.29	0.31	0.29	0.83	0.82	0.83
(0.75,1)	0.64	0.62	0.59	0.63	0.64	0.66	(0.75,1)	0.68	0.68	0.65	0.61	0.61	0.63
(0.4,0.6)	0.29	0.29	0.28	0.83	0.82	0.83	(0.4,0.6)	0.32	0.29	0.33	0.81	0.83	0.80
(0.1,0.9)	0.05	0.05	0.05	0.97	0.97	0.97	(0.1,0.9)	0.06	0.05	0.05	0.97	0.97	0.97
			(categories = 25)							*(categories = 50)*			
(0.1,1)	0.08	0.06	0.08	0.95	0.96	0.96	(0.1,1)	0.07	0.09	0.09	0.96	0.95	0.95
(0.5,1)	0.32	0.30	0.29	0.81	0.82	0.83	(0.5,1)	0.31	0.31	0.31	0.82	0.82	0.82
(0.75,1)	0.67	0.65	0.71	0.62	0.63	0.59	(0.75,1)	0.73	0.72	0.70	0.58	0.59	0.60
(0.4,0.6)	0.32	0.32	0.30	0.81	0.81	0.82	(0.4,0.6)	0.32	0.31	0.34	0.81	0.81	0.80
(0.1,0.9)	0.05	0.05	0.06	0.97	0.97	0.97	(0.1,0.9)	0.06	0.06	0.06	0.96	0.96	0.96
			(categories = 250)							*(categories = 500)*			
(0.1,1)	0.07	0.05	0.06	0.96	0.97	0.97	(0.1,1)	0.05	0.07	0.07	0.97	0.96	0.96
(0.5,1)	0.32	0.30	0.32	0.82	0.82	0.81	(0.5,1)	0.33	0.33	0.34	0.81	0.81	0.80
(0.75,1)	0.71	0.69	0.71	0.59	0.60	0.59	(0.75,1)	0.70	0.70	0.70	0.60	0.60	0.59
(0.4,0.6)	0.32	0.31	0.33	0.81	0.82	0.80	(0.4,0.6)	0.32	0.31	0.31	0.81	0.82	0.82
(0.1,0.9)	0.06	0.06	0.07	0.96	0.96	0.96	(0.1,0.9)	0.06	0.06	0.07	0.97	0.97	0.96

p=21, n=75, lambda=0.7,0.4 or 0.3, psi=0.51, 100 samples per DGP

Table 29

Simulation Results - Complex Five-Factor Models

	$E[(\hat{f}^{*}-\hat{f})^2]$					$\rho(\hat{f}^{*},\hat{f})$						$E[(\hat{f}^{*}-\hat{f})^2]$					$\rho(\hat{f}^{*},\hat{f})$				
(a,b)	f_1	f_2	f_3	f_4	f_5	f_1	f_2	f_3	f_4	f_5	(a,b)	f_1	f_2	f_3	f_4	f_5	f_1	f_2	f_3	f_4	f_5
					(categories = 2)											*(categories = 3)*					
(0.1,1)	0.87	0.80	0.68	0.92	0.89	0.54	0.59	0.64	0.48	0.49	(0.1,1)	0.28	0.34	0.31	0.35	0.54	0.85	0.82	0.83	0.80	0.68
(0.5,1)	0.47	0.53	0.51	0.58	0.93	0.75	0.72	0.73	0.67	0.46	(0.5,1)	0.28	0.33	0.32	0.34	0.48	0.85	0.83	0.83	0.81	0.71
(0.75,1)	0.60	0.66	0.63	0.75	1.07	0.69	0.66	0.66	0.58	0.39	(0.75,1)	0.54	0.60	0.61	0.61	0.88	0.72	0.69	0.67	0.66	0.49
(0.4,0.6)	-	-	-	-	-	-	-	-	-	-	(0.4,0.6)	0.34	0.38	0.38	0.42	0.48	0.82	0.80	0.80	0.76	0.71
(0.1,0.9)	-	-	-	-	-	-	-	-	-	-	(0.1,0.9)	0.50	0.56	0.58	0.66	0.91	0.74	0.71	0.69	0.63	0.47
					(categories = 7)											*(categories = 15)*					
(0.1,1)	0.05	0.07	0.08	0.14	0.46	0.97	0.96	0.96	0.92	0.74	(0.1,1)	0.06	0.07	0.09	0.17	0.54	0.97	0.97	0.95	0.91	0.70
(0.5,1)	0.29	0.31	0.33	0.34	0.35	0.85	0.84	0.82	0.81	0.79	(0.5,1)	0.34	0.35	0.34	0.40	0.41	0.82	0.82	0.82	0.78	0.76
(0.75,1)	0.64	0.68	0.70	0.76	0.93	0.67	0.65	0.63	0.58	0.47	(0.75,1)	0.68	0.69	0.77	0.81	0.99	0.65	0.64	0.60	0.56	0.43
(0.4,0.6)	0.38	0.42	0.41	0.39	0.60	0.80	0.79	0.78	0.78	0.65	(0.4,0.6)	0.40	0.43	0.44	0.42	0.64	0.79	0.78	0.77	0.77	0.63
(0.1,0.9)	0.06	0.07	0.07	0.07	0.08	0.97	0.96	0.96	0.96	0.95	(0.1,0.9)	0.07	0.08	0.08	0.07	0.06	0.96	0.96	0.96	0.96	0.96
					(categories = 25)											*(categories = 50)*					
(0.1,1)	0.07	0.08	0.10	0.14	0.92	0.96	0.96	0.95	0.92	0.48	(0.1,1)	0.07	0.07	0.12	0.14	0.71	0.96	0.96	0.94	0.92	0.60
(0.5,1)	0.36	0.36	0.39	0.38	0.47	0.82	0.81	0.79	0.79	0.72	(0.5,1)	0.36	0.37	0.39	0.41	0.45	0.81	0.81	0.79	0.77	0.74
(0.75,1)	0.68	0.73	0.75	0.84	0.94	0.65	0.63	0.61	0.54	0.46	(0.75,1)	0.72	0.75	0.77	0.87	1.04	0.63	0.62	0.59	0.52	0.41
(0.4,0.6)	0.40	0.45	0.45	0.40	0.68	0.79	0.77	0.76	0.77	0.61	(0.4,0.6)	0.39	0.44	0.44	0.45	0.56	0.80	0.77	0.77	0.74	0.67
(0.1,0.9)	0.08	0.09	0.09	0.07	0.07	0.96	0.95	0.95	0.96	0.96	(0.1,0.9)	0.09	0.09	0.09	0.10	0.08	0.96	0.95	0.95	0.95	0.95
					(categories = 250)											*(categories = 500)*					
(0.1,1)	0.06	0.08	0.07	0.09	0.44	0.97	0.96	0.96	0.95	0.75	(0.1,1)	0.07	0.08	0.09	0.11	0.48	0.97	0.96	0.95	0.94	0.73
(0.5,1)	0.37	0.37	0.39	0.41	0.44	0.81	0.81	0.79	0.77	0.74	(0.5,1)	0.37	0.39	0.39	0.40	0.38	0.81	0.80	0.79	0.78	0.77
(0.75,1)	0.72	0.73	0.77	0.87	1.06	0.63	0.63	0.60	0.52	0.40	(0.75,1)	0.71	0.75	0.79	0.83	0.98	0.63	0.62	0.58	0.54	0.44
(0.4,0.6)	0.39	0.45	0.44	0.42	0.53	0.80	0.77	0.76	0.76	0.68	(0.4,0.6)	0.41	0.46	0.43	0.45	0.65	0.79	0.76	0.77	0.75	0.62
(0.1,0.9)	0.08	0.09	0.09	0.08	0.08	0.96	0.95	0.95	0.96	0.95	(0.1,0.9)	0.09	0.10	0.10	0.10	0.08	0.95	0.95	0.95	0.95	0.95

p=45, n=75, lambda, psi - see Appendix, 100 samples per DGP

Table 30

Simulation Results - Complex, Negative Two-Factor Models

(a,b)	$E[(\hat{f}^* - \hat{f})^2]$ f_1	f_2	$\rho(\hat{f}^*, \hat{f})$ f_1	f_2	(a,b)	$E[(\hat{f}^* - \hat{f})^2]$ f_1	f_2	$\rho(\hat{f}^*, \hat{f})$ f_1	f_2
	(categories = 2)					*(categories = 3)*			
(0.1,1)	0.79	0.78	0.53	0.54	(0.1,1)	0.13	0.18	0.92	0.90
(0.5,1)	0.31	0.31	0.81	0.81	(0.5,1)	0.16	0.17	0.91	0.90
(0.75,1)	0.38	0.44	0.77	0.74	(0.75,1)	0.39	0.41	0.77	0.76
(0.4,0.6)	-	-	-	-	(0.4,0.6)	0.16	0.16	0.91	0.90
(0.1,0.9)	-	-	-	-	(0.1,0.9)	0.31	0.33	0.81	0.80
	(categories = 7)					*(categories = 15)*			
(0.1,1)	0.05	0.04	0.97	0.97	(0.1,1)	0.03	0.05	0.98	0.97
(0.5,1)	0.18	0.20	0.89	0.88	(0.5,1)	0.20	0.21	0.88	0.88
(0.75,1)	0.49	0.53	0.71	0.69	(0.75,1)	0.52	0.54	0.69	0.69
(0.4,0.6)	0.17	0.18	0.90	0.89	(0.4,0.6)	0.19	0.24	0.89	0.86
(0.1,0.9)	0.03	0.05	0.98	0.97	(0.1,0.9)	0.03	0.04	0.98	0.98
	(categories = 25)					*(categories = 50)*			
(0.1,1)	0.04	0.04	0.98	0.98	(0.1,1)	0.03	0.04	0.98	0.98
(0.5,1)	0.22	0.25	0.87	0.86	(0.5,1)	0.23	0.27	0.86	0.84
(0.75,1)	0.57	0.57	0.66	0.67	(0.75,1)	0.56	0.62	0.67	0.64
(0.4,0.6)	0.18	0.19	0.89	0.88	(0.4,0.6)	0.19	0.23	0.88	0.87
(0.1,0.9)	0.03	0.04	0.98	0.98	(0.1,0.9)	0.03	0.03	0.98	0.98
	(categories = 250)					*(categories = 500)*			
(0.1,1)	0.03	0.03	0.98	0.98	(0.1,1)	0.03	0.03	0.98	0.98
(0.5,1)	0.22	0.28	0.87	0.84	(0.5,1)	0.24	0.25	0.86	0.86
(0.75,1)	0.62	0.63	0.64	0.63	(0.75,1)	0.58	0.57	0.66	0.67
(0.4,0.6)	0.17	0.20	0.90	0.88	(0.4,0.6)	0.20	0.24	0.88	0.86
(0.1,0.9)	0.03	0.04	0.98	0.98	(0.1,0.9)	0.04	0.04	0.98	0.98

p=14, n=75, lambda=0.7, 0.4, 0 or -0.3; psi = various; 100 samples per DGP

Table 31

Simulation Results - Complex, Negative Three-Factor Models

(a,b)	$E[(\hat{f}^* - \hat{f})^2]$ f_1	f_2	f_3	$\rho(\hat{f}^*, \hat{f})$ f_1	f_2	f_3	(a,b)	$E[(\hat{f}^* - \hat{f})^2]$ f_1	f_2	f_3	$\rho(\hat{f}^*, \hat{f})$ f_1	f_2	f_3
			(categories = 2)							*(categories = 3)*			
(0.1,1)	0.96	0.72	1.04	0.43	0.60	0.41	(0.1,1)	0.44	0.12	0.58	0.74	0.93	0.65
(0.5,1)	0.63	0.24	0.80	0.62	0.87	0.53	(0.5,1)	0.44	0.16	0.72	0.73	0.91	0.57
(0.75,1)	0.54	0.39	0.90	0.67	0.78	0.47	(0.75,1)	0.59	0.41	0.92	0.65	0.77	0.47
(0.4,0.6)	-	-	-	-	-	-	(0.4,0.6)	0.50	0.15	0.78	0.70	0.91	0.54
(0.1,0.9)	-	-	-	-	-	-	(0.1,0.9)	0.64	0.26	0.86	0.63	0.85	0.49
			(categories = 7)							*(categories = 15)*			
(0.1,1)	0.28	0.04	0.45	0.83	0.98	0.73	(0.1,1)	0.27	0.03	0.35	0.84	0.98	0.79
(0.5,1)	0.37	0.23	0.89	0.78	0.87	0.48	(0.5,1)	0.41	0.24	0.91	0.76	0.87	0.47
(0.75,1)	0.67	0.49	1.05	0.60	0.73	0.41	(0.75,1)	0.63	0.55	1.02	0.63	0.69	0.43
(0.4,0.6)	0.54	0.17	0.78	0.68	0.90	0.54	(0.4,0.6)	0.62	0.18	0.82	0.62	0.90	0.52
(0.1,0.9)	0.29	0.03	0.34	0.82	0.98	0.80	(0.1,0.9)	0.23	0.03	0.46	0.86	0.98	0.73
			(categories = 25)							*(categories = 50)*			
(0.1,1)	0.37	0.04	0.50	0.78	0.98	0.71	(0.1,1)	0.31	0.04	0.46	0.80	0.98	0.73
(0.5,1)	0.42	0.26	0.95	0.74	0.86	0.46	(0.5,1)	0.47	0.28	0.90	0.72	0.84	0.49
(0.75,1)	0.68	0.55	1.05	0.60	0.69	0.41	(0.75,1)	0.57	0.58	1.04	0.67	0.68	0.41
(0.4,0.6)	0.49	0.19	0.85	0.71	0.90	0.50	(0.4,0.6)	0.56	0.18	0.86	0.66	0.90	0.50
(0.1,0.9)	0.20	0.03	0.36	0.88	0.98	0.79	(0.1,0.9)	0.24	0.04	0.38	0.86	0.98	0.78
			(categories = 250)							*(categories = 500)*			
(0.1,1)	0.25	0.04	0.38	0.85	0.98	0.78	(0.1,1)	0.31	0.04	0.52	0.81	0.98	0.70
(0.5,1)	0.54	0.28	0.91	0.68	0.85	0.48	(0.5,1)	0.58	0.27	0.91	0.65	0.85	0.48
(0.75,1)	0.65	0.56	1.04	0.62	0.69	0.42	(0.75,1)	0.75	0.56	1.05	0.55	0.69	0.41
(0.4,0.6)	0.56	0.18	0.73	0.65	0.90	0.58	(0.4,0.6)	0.43	0.18	0.76	0.75	0.90	0.55
(0.1,0.9)	0.23	0.04	0.40	0.86	0.98	0.77	(0.1,0.9)	0.34	0.05	0.46	0.79	0.98	0.73

p=21, n=75, lambda=0.7,0.4,,0.3,0,-0.3, or -0.4; psi=various; 100 samples per DGP

Table 32

5.2 Discussion

Tables 25-32 present my simulation results. Each table presents results from simulations using a fixed loading matrix, Λ. All loadings matrices are shown in the appendix. For example, Table 26's "Simple Two-Factor Models" refer to DGPs using the loadings matrix Λ_{simple} shown in the "Two-factor models" section of the Appendix. For indicated values of (a, b) and T, and for each factor f_j of the given model, the tables display two statistics: $\overline{msd}_j$ and $\overline{\rho}_j$. Note that in all tables, cells corresponding to $T = 2$, $(a, b) \in [(0.4, 0.6), (0.1, 0.9)]$ are blank because imposing two finite-valued cutoffs (e.g. - at the 40th and 60th percentiles) implies at least three score categories.

Tables25-32 hold a massive amount of figures, but a few patterns stand out. First, $\overline{\rho}_j$ is increasing in T only for very small values of T. After the number categories for polytomous variables in X^* reaches 7, the polytomous nature of the data *per se* no longer constrains the accuracy of factor score prediction. What does constrain $\overline{\rho}_j$, even at high values of T is censoring induced by the value of (a, b). Looking across the eight tables, one can see that for any given Λ and any given value of T, the lowest value of $\overline{\rho}_j$ almost always occurs at $(a, b) = (0.75, 1)$, where the discretization assigns the same score (a "1") to all countries in the bottom 75% of the true distribution of governance. Such an extreme form of left-censoring so greatly distorts the actual variation in governance that it is no surprise to find the factor predictors weakly correlated (it is more surprising that their correlation isn't even weaker!). Even though the value $(0.4, 0.6)$ actually censors a greater portion of X's density than does $(0.75, 1)$, the value of $\overline{\rho}_j$ increases when (a, b) changes to $(0.4, 0.6)$. This observation suggests that the censoring's asymmetry, not just it's extent, matters for $\overline{\rho}_j$.

Remarkably, even with T as low as 2 or 3, the polytomous data's predictor, $E[f|X^*]$, performs relatively well (a high $\overline{\rho}_j$) under the simple and the complex models and even performs decently for some scenarios in the complex and negative models.

Complexity of the loading pattern reduces the relative accuracy of the polytomous-data predictor.[69] Regardless of T,m or (a, b), moving from a simple model to a complex or complex and negative model with the same number of factors (e.g., from Table 26 to Table 28 or Table 29) causes $\overline{\rho}_j$ to fall. Thus, one should expect polytomous data to cause bigger problems for factor prediction whenever invividual variables tend to load nontrivially on multiple common factors.

Tables 7 and 8 show some surprising results. I had expected that the complex, negative loading patterns in these models to induce the poorest relative performance of the polytomous-data predictor, but the results contradict this hypothesis. Adding negative loadings to an already complex loading pattern can either raise or lower $\overline{\rho}_j$, depending on the particular characteristics of the DGP.

I focus finally on Table 30, which presents results for a model equal in size (same number of factors) and similar in loading pattern to the estimated model Λ used to generate score predictions in Chapter 1. Of all the models we've experimented with in this paper, I believe the five-factor complex model in Table 6 is the one most closely resembling the DGP that gave rise to my governance data. In designing the

[69] The correlation of the continuous-data predictor, $E[f|X]$, with f also falls with more complex loading patterns, but not nearly as much as does the correlation of $E[f|X]$ with f. Thus the drop in $\overline{\rho}_j$ reflects a fall in the relative accuracy of $E[f|X^*]$ compared with $E[f|X]$.

five-factor model, I have tried to match the observable attributes of Chapter 1's estimated model (same number of factors, same number of variables, same proportion of variation explained by common factors ($\approx 83\%$), and same communality attributable to each factor). Of course, since Chapter 1's estimated model was itself derived from polytomous data, that model may contain biased parameter estimates. But it is currently the best benchmark I have.

I take encouragement from Table 30. It shows that, under a factor model with characteristics similar to the one estimated from the governance variables in Chapter 1, the factor score predictions one will produce using polytomous data will be very highly correlated with the predictions one would have generated using the underlying continuous variables. Except under the most extreme censoring (when $(a, b) = (0.75, 1)$), Table 6 shows that $\overline{\rho}_j \geq 0.75$ for the first four factors - even with T as low as 3! [70] In short, the ML factor prediction strategy used in Chapter 1, though formally not suited for polytomous data, is nevertheless quite robust to even the coarsest discretization of information from continuous variables.

Overall, my results suggest that governance inference based on polytomous variables will not differ substantially from (and in many circumstances, will coincide closely with) governance inference based on a richer, continuous dataset of variables that measure the same concepts.

Of course, the meaning of "differ substantially" is contextual. Conclusions from one study relying on factor score predictions may not be seriously altered if $E[f|X]$

[70] The values of $\overline{\rho}_j$ are noticeably lower for f_5 compared to $f_1 - f_4$. The fifth factor, not surprisingly, is the factor with the smallest explanatory power in our simulations ($\approx 8.4\%$ of explained variation is attributable to f_5). The equivalent of f_5 in our data is the MIG factor, which we do not use in our regression analysis in Chapter 2.

is replaced with $E[f|X^*]$, while conclusions from another study may be.

How might the conclusions from my growth regressions in Chapter 2 be affected by such a substitution? In that situation, $E[f|X^*]$ was an explanatory variable for per capita income. Assuming that $E[f|X^*] = E[f|X] + e$ - where e is white noise, and thus $E[f|X^*]$ a noisy but unbiased proxy for $E[f|X]$ - I face measurement error in a regressor. Under OLS, the estimate of the coefficient on governance is attenuated to zero, meaing the use of $E[f|X^*]$ instead of $E[f|X]$ causes us to understate the impact of governance on growth. However, the regressions in Chapter 2 are estimated using instrumental variables, and so the effect of measurement error in an endogenous regressor is less clear.

6 Bimodal Discrete Distributions

In the results of Tables 25-32, continuous variables were transformed into polytomous variables by dividing up the continous variable's support using $(T-1)$ equally spaced cutoff values. For example, I generate a seven-category polytomous variable from a continuous standard normal variable by placing cutoffs at $[z_{0.1}, z_{0.26}, z_{0.42}, z_{0.58}, z_{0.74}, z_{0.9}]$. The highest (lowest) categorical score can constitute a mode of the resulting discrete distribution whenever $z_{\max}$ $(z_{\min})$ is located sufficiently close to the mean.

Equally spaced bins however cannot generate a discrete distribution with multiple interior modes - i.e. with two or more modes at categorical scores other than the highest or lowest. A discrete distribution with two interior modes could arise in a perceptions-based governance data if, for example, the data provider tends to lump countries into one of two middling categories. For instance, on a discrete scale of 1 to 5, we may observe modes at scores "2" and "4" if the analyst views most countries as either "somewhat below average" or "somewhat above average".

I carry out a second set of Monte Carlo exercises to assess the impact of such a distortion on factor score prediction. These simulations are much less extensive than those presented in Tables 25-32 but they nevertheless provide a sense of the robustness of factor score prediction based on polytomous data with two interior modes.

Results are presented in Tables 34-37 below. Each table presents the results for all models of a particular size (a particular # of factors). I focus on a much narrower variety of discretization schemes than in Tables 25-32, looking only at polytomous data with $T = 5, 7$ and 15, and considering only a single placement of cutoff values per value of T. Otherwise, the simulation procedures and parameter values are all identical to the procedures and parameters used in Tables 25-32. The placement of cutoff values and the resulting location of interior modes is presented in Table 33 below.

Cutoff Locations (z_α) for Polytomous Data with Interior Modes

number of categories	cutoff locations	modal score categories
5	$[z_{0.05}, z_{0.45}, z_{0.55}, z_{0.95}]$	"2", "4"
7	$[z_{0.02}, z_{0.40}, z_{0.45}, z_{0.55}, z_{0.60}, z_{0.98}]$	"2", "6"
15	$[z_{0.01}, z_{0.04}, z_{0.10}, z_{0.25}, z_{0.40}, z_{0.46}, z_{0.49}, z_{0.51}, z_{0.54}, z_{0.60}, z_{0.75}, z_{0.90}, z_{0.96}, z_{0.99}]$	"4" and "5"; "11" and "12"

$F(z_\alpha) = \alpha$ under the standard normal distribution. So, e.g. - when the first cutoff location is $z_{0.05}$, all continuous values less than -1.649 receive discrete score "1".

Table 33

Thus, cutoffs for the five-category data generate modes at scores "2" and "4"; cutoffs for the seven-category data create modes at "2" and "6"; and cutoffs for the fifteen-category data produce modes at score clusters 4 and 5, and 11 and 12.

Table 34 presents results for one-factor models.

One-Factor Models

categories	$E[(\hat{f}^* - \hat{f})^2]$ f_1	$\rho(\hat{f}^*, \hat{f})$ f_1
5	0.06	0.97
7	0.08	0.95
15	0.03	0.98

**n=75, p=7, 100 samples per DGP*

Table 34

Next, Table 35 presents results from simple, complex, and complex and negative two-factor models.

Two-Factor Models

	$E[(\hat{f}^* - \hat{f})^2]$		$\rho(\hat{f}^*, \hat{f})$	
categories	f_1	f_2	f_1	f_2
		Simple two-factor models*		
5	0.06	0.06	0.96	0.97
7	0.08	0.09	0.95	0.95
15	0.04	0.04	0.98	0.98
		Complex two-factor models*		
5	0.17	0.17	0.90	0.90
7	0.24	0.25	0.86	0.85
15	0.10	0.10	0.94	0.94
		Complex, negative two-factor models*		
5	0.09	0.09	0.95	0.95
7	0.13	0.15	0.92	0.91
15	0.07	0.06	0.96	0.97

**n=75, p=14, 100 samples per DGP;*

Table 35

Table 36 presents results from the three-factor models.

Three-Factor Models

	$E[(\hat{f}^* - \hat{f})^2]$			$\rho(\hat{f}^*, \hat{f})$		
categories	f_1	f_2	f_3	f_1	f_2	f_3
Simple three-factor models*						
5	0.06	0.06	0.06	0.96	0.96	0.96
7	0.09	0.09	0.09	0.95	0.95	0.95
15	0.04	0.04	0.04	0.98	0.98	0.98
Complex three-factor models*						
5	0.18	0.17	0.19	0.89	0.90	0.88
7	0.24	0.24	0.24	0.86	0.86	0.86
15	0.11	0.12	0.12	0.94	0.93	0.93
Complex, negative three-factor models*						
5	0.54	0.08	0.65	0.67	0.96	0.62
7	0.43	0.13	0.64	0.74	0.93	0.62
15	0.35	0.06	0.47	0.79	0.97	0.73

**n=75, p=21, 100 samples per DGP*

Table 36

And finally, Table 37 presents results from a complex five-factor model similar to the model estimated from the actual governance data.

Five-Factor Models*

	$E[(\hat{f}^* - \hat{f})^2]$					$\rho(\hat{f}^*, \hat{f})$				
categories	f_1	f_2	f_3	f_4	f_5	f_1	f_2	f_3	f_4	f_5
5	0.22	0.27	0.25	0.25	0.31	0.89	0.86	0.87	0.86	0.81
7	0.31	0.35	0.33	0.36	0.41	0.84	0.82	0.82	0.80	0.75
15	0.14	0.16	0.16	0.18	0.20	0.93	0.92	0.91	0.90	0.88

**n=75, p=45, 100 samples per DGP*

Table 37

The results of these Monte Carlo exercises using bimodal polytomous data are largely consistent with the results of Tables 25-32. Factor score predictions based on bimodal polytomous variables are highly correlated with score predictions based on the underlying continous variables. Indeed, often $\rho \geq 0.90$. As before, this correlation increases with the number of score categories in the polytomous variable. The correlation is generally decreasing in loading complexity, as is especially apparent in the three-factor models (see Table 36).

In summary, the results of Tables 34-37 produce further evidence that even radical discretizations of continuous normal data with a factor structure will not substantively distort the prediction of factor scores.

7 Conclusion

This paper has conducted a targeted set of Monte Carlo exercises to assess the impact on factor score prediction of using polytomous versions of continuous variables. I have found that the correlation ($\overline{\rho}_j$) between the polytomous-data-based prediction of a factor (i.e., $E[f_j|X^*]$), and the continous-data-based prediction (i.e., $E[f_j|X]$) is increasing in i.) the number of categories in the polytomous variables (but only when the initial number of categories is very small), and ii.) the share of communality attributable to the factor, f_j, being predicted. However, $\overline{\rho}_j$ is decreasing in i.) loading pattern complexity, and ii.) the extent of censoring induced by the polytomous variable (especially asymmetrical censoring). Over the range of model sizes with which I experimented, the number of factors *per se* has no discernible effect on $\overline{\rho}_j$, but this result is probably connected with the fact that I kept the ratio of variables to factors constant in my one-,two- and three-factor models.

On the whole, my simulation results demonstrate that in many circumstances, the standard ML factor model estimator can be employed with polytomous data without sacrificing a great deal of predictive accuracy for factor scores. However, this does not imply that in estimating factor models, one should always simply treat polytomous data as if it were continuous - especially if the number of categories is very small, or the degree of censoring imposed by the categorization is known to be extreme. Other models, estimators and methods exist (e.g. - the Tobit factor model; GLS and WLS estimators; the use of polychoric correlations in place of Pearson product-moment correlations), which try to explicitly account for the polytomous and/or non-normal distribution of the variables and which may therefore yield more accurate factor score predictions.

Conclusion

Measuring governance is frought with both practical and conceptual complications, even though most observers can agree on the broad outlines of good and bad practices. To get anywhere, one needs first to impose some basic assumptions on what is to be measured and on the relationship between those objects of measurement and the observable data. This paper has offered a cohesive and transparent approach that addresses both imperatives and also generates intuitive new findings on growth and governance.

By design, my new governance measures explain the intercorrelation of expert opinions, but substantively they assess fundamental dimensions of a nation's political, legal and civic order: are citizens equal before the law.? is the bureaucracy efficient and accountable? are contracting and property rights upheld? does the state respect and defend human rights? are there pervasive threats to peace and societal cohesion? That the scope and dynamism of markets might hinge on the anwers to these questions seems uncontroversial. But establishing such assertions empirically requires reliable governance measures.

In Chapter 1, I applied factor model techniques to construct such measures. Chapter 2 applied these assessments to explain observed cross-country variation in per capita income. Finally, Chapter 3 produced evidence suggesting that potential distortion in the factor score predictions due to my reliance on polytomous data is unlikely to be substantial.

While my results suggest a major (perhaps dominant) role for governance in economic performance, clearly much more work needs to be done in this area. From information aggregation techniques, to the econometrics of governance and growth, to the theoretical modeling of governance reform - scholars and policymakers have much to learn about how to measure governance, how to improve it, and what economic impacts to expect from those improvements. I hope that my investigation has shed some new light on these topics.

Appendices

1 Appendix to Chapter 1

1.1 Factor Model Parameter Estimates

Here I provide i.) the full $\widehat{\Lambda}, \widehat{\Psi}$ matrices from three rotations of the $m = 5$ orthogonal factor model estimated on the 45 governance variables, ii.) a key to variable names, and iii.) summary statistics for all the variables in the governance dataset, for balanced and unbalanced panels

UNROTATED Λ,Ψ						
Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Uniqueness
$corrupt_{GRS}$	0.902	0.078	0.274	-0.155	0.079	0.075
$corrupt_{EIU}$	0.948	-0.134	-0.054	-0.004	-0.072	0.076
$corrupt_{GCS}$	0.892	-0.266	0.023	-0.011	0.007	0.132
$corrupt_{MIG}$	0.834	-0.022	0.105	0.439	0.176	0.068
$corrupt_{PRS}$	0.881	-0.071	-0.209	0.160	0.123	0.134
$corrupt_{QLM}$	0.941	-0.229	-0.193	-0.024	-0.060	0.021
$corrupt_{BCRI}$	0.959	-0.049	-0.007	-0.059	-0.063	0.071
$stability_{GRS}$	0.775	0.207	0.445	-0.109	-0.062	0.144
$stability_{EIU}$	0.835	0.174	0.128	0.053	-0.173	0.223
$stability_{GCS}$	0.443	0.186	0.285	0.251	-0.501	0.374
$stability_{HUM}$	0.779	0.065	0.176	0.257	-0.304	0.199
$stability_{IJT}$	0.846	-0.020	0.102	0.053	0.006	0.271
$stability_{MIG}$	0.585	0.238	0.274	0.351	-0.239	0.346
$stability_{PRS}$	0.437	-0.171	0.459	0.219	-0.273	0.447
$stability_{BCRI}$	0.732	0.081	0.428	0.157	-0.421	0.071
$voice_{EIU}$	0.910	0.307	-0.173	-0.030	0.008	0.047
$voice_{FRH}$	0.728	0.606	-0.232	0.013	0.021	0.048
$voice_{GCS}$	0.851	-0.057	0.036	-0.014	0.211	0.226
$voice_{HUM}$	0.663	0.609	-0.279	0.029	-0.066	0.106
$voice_{PRS}$	0.771	0.411	-0.025	-0.158	-0.080	0.205
$voice_{RSF}$	0.691	0.514	-0.133	0.029	-0.088	0.232
$voice_{BCRI}$	0.931	0.208	-0.033	-0.127	0.036	0.071
$regulation_{GRS}$	0.593	0.137	0.359	-0.216	0.402	0.292
$regulation_{EIU}$	0.833	0.132	-0.049	-0.057	0.043	0.280
$regulation_{GCS}$	0.696	-0.333	0.163	-0.150	0.109	0.344
$regulation_{HER}$	0.821	0.260	-0.050	0.045	0.081	0.247
$regulation_{MIG}$	0.780	0.090	0.102	0.400	0.120	0.198
$regulation_{PRS}$	0.807	0.174	0.150	-0.115	0.216	0.237
$regulation_{BCRI}$	0.955	0.083	0.072	-0.124	0.034	0.059
$rulelaw_{GRS}$	0.848	0.098	0.416	-0.156	0.104	0.063
$rulelaw_{EIU}$	0.954	-0.005	0.002	0.015	-0.006	0.090
$rulelaw_{GCS}$	0.870	-0.318	0.096	-0.033	0.084	0.124
$rulelaw_{HER}$	0.942	-0.149	-0.146	0.050	0.080	0.061
$rulelaw_{HUM}$	0.784	0.270	-0.221	-0.082	-0.130	0.239
$rulelaw_{MIG}$	0.833	-0.151	0.049	0.375	0.159	0.115
$rulelaw_{PRS}$	0.759	-0.234	0.158	0.093	-0.052	0.332
$rulelaw_{QLM}$	0.939	-0.245	-0.161	-0.070	-0.075	0.022
$rulelaw_{BCRI}$	0.917	-0.106	0.088	-0.024	-0.093	0.130
$goveffect_{GRS}$	0.813	0.023	0.446	-0.200	0.141	0.079
$goveffect_{EGV}$	0.520	-0.329	-0.032	-0.036	0.203	0.579
$goveffect_{EIU}$	0.913	-0.125	-0.097	-0.033	0.032	0.138
$goveffect_{GCS}$	0.893	-0.233	0.157	-0.030	0.059	0.118
$goveffect_{MIG}$	0.848	0.056	0.065	0.383	0.194	0.089
$goveffect_{PRS}$	0.820	0.043	-0.200	0.074	0.144	0.260
$goveffect_{BCRI}$	0.962	0.016	0.007	-0.100	-0.017	0.064

QUARTIMAX Λ,Ψ						
Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Uniqueness
$\text{corrupt}_{\text{GRS}}$	0.910	-0.009	0.060	0.294	-0.083	0.075
$\text{corrupt}_{\text{EIU}}$	0.946	-0.065	0.004	-0.144	-0.068	0.076
$\text{corrupt}_{\text{GCS}}$	0.898	-0.223	-0.038	-0.093	-0.038	0.132
$\text{corrupt}_{\text{MIG}}$	0.843	-0.046	0.087	-0.030	0.459	0.068
$\text{corrupt}_{\text{PRS}}$	0.879	0.037	-0.157	-0.202	0.164	0.134
$\text{corrupt}_{\text{QLM}}$	0.937	-0.100	-0.106	-0.267	-0.096	0.021
$\text{corrupt}_{\text{BCRI}}$	0.956	-0.005	0.019	-0.057	-0.102	0.071
$\text{stability}_{\text{GRS}}$	0.781	0.050	0.296	0.384	-0.089	0.144
$\text{stability}_{\text{EIU}}$	0.830	0.152	0.252	0.030	-0.031	0.223
$\text{stability}_{\text{GCS}}$	0.434	0.112	0.649	-0.064	-0.003	0.374
$\text{stability}_{\text{HUM}}$	0.774	0.041	0.428	-0.105	0.077	0.199
$\text{stability}_{\text{IJT}}$	0.849	-0.029	0.073	0.028	0.041	0.271
$\text{stability}_{\text{MIG}}$	0.582	0.150	0.497	0.028	0.213	0.346
$\text{stability}_{\text{PRS}}$	0.447	-0.299	0.505	0.058	0.069	0.447
$\text{stability}_{\text{BCRI}}$	0.732	-0.035	0.621	0.067	-0.046	0.071
$\text{voice}_{\text{EIU}}$	0.897	0.379	-0.053	-0.033	-0.027	0.047
$\text{voice}_{\text{FRH}}$	0.707	0.670	-0.029	0.023	0.036	0.048
$\text{voice}_{\text{GCS}}$	0.858	-0.053	-0.144	0.086	0.077	0.226
$\text{voice}_{\text{HUM}}$	0.639	0.695	0.014	-0.055	0.005	0.106
$\text{voice}_{\text{PRS}}$	0.758	0.419	0.066	0.120	-0.161	0.205
$\text{voice}_{\text{RSF}}$	0.673	0.553	0.095	0.008	-0.002	0.232
$\text{voice}_{\text{BCRI}}$	0.926	0.234	-0.047	0.083	-0.094	0.071
$\text{regulation}_{\text{GRS}}$	0.612	-0.021	-0.142	0.558	0.039	0.292
$\text{regulation}_{\text{EIU}}$	0.829	0.166	-0.051	0.026	-0.036	0.280
$\text{regulation}_{\text{GCS}}$	0.710	-0.353	-0.096	0.090	-0.099	0.344
$\text{regulation}_{\text{HER}}$	0.815	0.281	-0.019	0.051	0.081	0.247
$\text{regulation}_{\text{MIG}}$	0.785	0.061	0.132	-0.004	0.406	0.198
$\text{regulation}_{\text{PRS}}$	0.813	0.115	-0.079	0.288	0.017	0.237
$\text{regulation}_{\text{BCRI}}$	0.955	0.079	-0.008	0.114	-0.094	0.059
$\text{rulelaw}_{\text{GRS}}$	0.861	-0.048	0.124	0.418	-0.059	0.063
$\text{rulelaw}_{\text{EIU}}$	0.953	0.028	0.016	-0.036	-0.005	0.090
$\text{rulelaw}_{\text{GCS}}$	0.882	-0.306	-0.070	-0.013	-0.017	0.124
$\text{rulelaw}_{\text{HER}}$	0.941	-0.054	-0.142	-0.165	0.045	0.061
$\text{rulelaw}_{\text{HUM}}$	0.766	0.370	-0.004	-0.123	-0.145	0.239
$\text{rulelaw}_{\text{MIG}}$	0.842	-0.141	0.023	-0.100	0.382	0.115
$\text{rulelaw}_{\text{PRS}}$	0.767	-0.246	0.123	-0.044	0.037	0.332
$\text{rulelaw}_{\text{QLM}}$	0.936	-0.125	-0.097	-0.239	-0.142	0.022
$\text{rulelaw}_{\text{BCRI}}$	0.919	-0.093	0.096	-0.033	-0.082	0.130
$\text{goveffect}_{\text{GRS}}$	0.830	-0.133	0.084	0.448	-0.081	0.079
$\text{goveffect}_{\text{EGV}}$	0.531	-0.287	-0.233	-0.043	0.035	0.579
$\text{goveffect}_{\text{EIU}}$	0.913	-0.048	-0.106	-0.113	-0.044	0.138
$\text{goveffect}_{\text{GCS}}$	0.904	-0.249	-0.001	0.048	-0.016	0.118
$\text{goveffect}_{\text{MIG}}$	0.854	0.041	0.045	-0.005	0.422	0.089
$\text{goveffect}_{\text{PRS}}$	0.816	0.134	-0.178	-0.114	0.110	0.260
$\text{goveffect}_{\text{BCRI}}$	0.960	0.046	-0.011	0.011	-0.108	0.064

VARIMAX Λ,Ψ						
Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Uniqueness
$\text{corrupt}_{\text{GRS}}$	0.550	0.382	0.615	0.277	0.147	0.075
$\text{corrupt}_{\text{EIU}}$	0.792	0.371	0.240	0.270	0.168	0.076
$\text{corrupt}_{\text{GCS}}$	0.808	0.206	0.292	0.224	0.192	0.132
$\text{corrupt}_{\text{MIG}}$	0.533	0.289	0.239	0.285	0.653	0.068
$\text{corrupt}_{\text{PRS}}$	0.722	0.425	0.129	0.084	0.375	0.134
$\text{corrupt}_{\text{QLM}}$	0.887	0.350	0.140	0.175	0.140	0.021
$\text{corrupt}_{\text{BCRI}}$	0.742	0.427	0.318	0.277	0.137	0.071
$\text{stability}_{\text{GRS}}$	0.330	0.360	0.630	0.458	0.105	0.144
$\text{stability}_{\text{EIU}}$	0.468	0.486	0.307	0.446	0.166	0.223
$\text{stability}_{\text{GCS}}$	0.118	0.256	0.058	0.732	0.089	0.374
$\text{stability}_{\text{HUM}}$	0.460	0.351	0.160	0.613	0.256	0.199
$\text{stability}_{\text{IJT}}$	0.595	0.338	0.339	0.289	0.250	0.271
$\text{stability}_{\text{MIG}}$	0.172	0.342	0.173	0.603	0.337	0.346
$\text{stability}_{\text{PRS}}$	0.276	-0.108	0.233	0.615	0.179	0.447
$\text{stability}_{\text{BCRI}}$	0.361	0.255	0.316	0.785	0.127	0.071
$\text{voice}_{\text{EIU}}$	0.532	0.740	0.255	0.159	0.183	0.047
$\text{voice}_{\text{FRH}}$	0.228	0.906	0.183	0.103	0.188	0.048
$\text{voice}_{\text{GCS}}$	0.640	0.329	0.405	0.080	0.292	0.226
$\text{voice}_{\text{HUM}}$	0.193	0.901	0.084	0.132	0.138	0.106
$\text{voice}_{\text{PRS}}$	0.344	0.711	0.346	0.228	0.018	0.205
$\text{voice}_{\text{RSF}}$	0.231	0.783	0.172	0.225	0.144	0.232
$\text{voice}_{\text{BCRI}}$	0.578	0.622	0.400	0.177	0.132	0.071
$\text{regulation}_{\text{GRS}}$	0.259	0.229	0.740	-0.020	0.199	0.292
$\text{regulation}_{\text{EIU}}$	0.548	0.517	0.316	0.155	0.166	0.280
$\text{regulation}_{\text{GCS}}$	0.679	0.008	0.415	0.118	0.094	0.344
$\text{regulation}_{\text{HER}}$	0.447	0.601	0.301	0.165	0.271	0.247
$\text{regulation}_{\text{MIG}}$	0.433	0.358	0.229	0.306	0.583	0.198
$\text{regulation}_{\text{PRS}}$	0.447	0.449	0.550	0.104	0.219	0.237
$\text{regulation}_{\text{BCRI}}$	0.641	0.494	0.460	0.232	0.144	0.059
$\text{rulelaw}_{\text{GRS}}$	0.458	0.314	0.709	0.317	0.160	0.063
$\text{rulelaw}_{\text{EIU}}$	0.698	0.445	0.320	0.265	0.229	0.090
$\text{rulelaw}_{\text{GCS}}$	0.803	0.122	0.369	0.188	0.213	0.124
$\text{rulelaw}_{\text{HER}}$	0.811	0.379	0.212	0.124	0.279	0.061
$\text{rulelaw}_{\text{HUM}}$	0.487	0.685	0.137	0.185	0.036	0.239
$\text{rulelaw}_{\text{MIG}}$	0.635	0.217	0.201	0.241	0.581	0.115
$\text{rulelaw}_{\text{PRS}}$	0.643	0.110	0.274	0.337	0.231	0.332
$\text{rulelaw}_{\text{QLM}}$	0.892	0.329	0.173	0.186	0.096	0.022
$\text{rulelaw}_{\text{BCRI}}$	0.716	0.324	0.332	0.345	0.148	0.130
$\text{goveffect}_{\text{GRS}}$	0.473	0.227	0.742	0.276	0.135	0.079
$\text{goveffect}_{\text{EGV}}$	0.585	-0.009	0.209	-0.062	0.178	0.579
$\text{goveffect}_{\text{EIU}}$	0.774	0.374	0.256	0.152	0.185	0.138
$\text{goveffect}_{\text{GCS}}$	0.751	0.175	0.420	0.250	0.217	0.118
$\text{goveffect}_{\text{MIG}}$	0.513	0.375	0.259	0.242	0.617	0.089
$\text{goveffect}_{\text{PRS}}$	0.615	0.486	0.178	0.036	0.305	0.260
$\text{goveffect}_{\text{BCRI}}$	0.705	0.473	0.375	0.242	0.132	0.064

1.2 Data Documentation

1.2.1 Variable Names

Variables in my analysis are named according to the convention category$_{\text{PROVIDER}}$. The names I use for categories reflect the governance categories to which variables were assigned by Kaufmann et al.:

Kaufmann et al. Governance Categories

Variable prefix in this paper	**Governance category in Kaufmann et al. (2006a)**
corrupt	Control of Corruption
rulelaw	Rule of Law
regulation	Regulatory Quality
voice	Voice and Accountability
goveffect	Government Effectiveness
stability	Political Stability/No Violence

1.2.2 Data Sources

The table below cross-references provider mnemonics used in the variable names with the actual providers and publications from which governance data was taken.

Data Sources

Mnemonic	Data Provider	Publication
BCRI	Global Insight	*Business Conditions and Risk Indicators*
GRS	Global Insight	*Global Risk Service*
GCS	World Economic Forum	*Global Competitiveness Survey*
EIU	Economist Intelligence Unit	*Country Risk Service, Country Forecast*
EGV	Professor Darrell M. West - Brown University, Brookings Institution	*Global E-Governance Index*
QLM	Business Environment Risk Intelligence	*Business Risk Service, Lender Risk Rating, and Quantitative Risk Measure in Foreign Lending*
HUM	U.S. State Department, Amnesty International, Cingranelli-Richards Human Rights Dataset, Professor Marc Gibney (U. of NC)	*U.S. State Dept.: Country Report on Human Rights Practices; Amnesty: Annual Reports; Gibney: Political Terror Scale*
RSF	Reporters Without Borders	*Press Freedom Index*
FRH	Freedom House	*Freedom in the World*
HER	Heritage Foundation/Wall Street Journal	*Index of Economic Freedom*
MIG	Merchant International Group	*Grey Area Dynamics*
IJT	Ijet	*Security Risk Ratings*

1.2.3 Variable Descriptions

The table below provides descriptions for all 45 governance variables used in this paper. The descriptions are taken more or less exactly as they appear in *Kaufmann et al. Governance Matters V: Appendices* (2006).

Governance Variables

Variable	**Description (Kaufmann et al., 2006b)**
corrupt_{BCRI}	Corruption: An assessment of the intrusiveness of the country's beaucracy. The amount of red tape likely to countered is assessed, as is the likelihood of encountering corrupt officials and other groups.
corrupt_{EIU}	Corruption among public officials
corrupt_{GCS}	Public trust in financial honesty of politicians; Diversion of public funds due to corruption is common; Frequency of bribery in the economy; Frequent for firms to make extra payments connected to: public utilities, tax payments, loan applications, awarding of public contracts, influencing laws, policies regulations, decrees, getting favourable judicial decisions; Extent to which firms' illegal payments to influence government policies impose costs on other firms; Extent to which influence of powerful firms with political ties impose costs on other firms
corrupt_{GRS}	A one-point increase on a scale from "0" to "10" in corruption during any 12-month period
corrupt_{MIG}	Corruption
corrupt_{PRS}	Corruption: Measures corruption within the political system, which distorts the economic and financial environment, reduces the efficiency of government and business by enabling people to assume positions of power through patronage rather than ability, and introduces an inherently instability in the political system.
corrupt_{QLM}	Political Risk Index - Internal causes of Political Risk: Mentality, including xenophobia, nationalism, corruption, nepotism, willingness to compromise. Indirect diversion of funds

continued on next page

Governance Variables - continued from previous page

Variable	**Description (Kaufmann et al., 2006b)**
goveffect$_{BCRI}$	Bureaucracy: An assessment of the quality of the country's bureaucracy. The better the bureaucracy the quicker decisions are made and the more easily foreign investors can go about their business. Policy consistency and forward planning: How confidence businesses can be of the continuity of economic policy stance - whether a change in government will entail major policy disruption, and whether the current government has pursued a coherent strategy. This factor also looks at the extent to which policy-making is far-sighted, or conversely aimed at short-term economic advantage.
goveffect$_{EGV}$	Global E-governance index
goveffect$_{EIU}$	Quality of bureaucracy / institutional effectiveness; Excessive bureaucracy / red tape
goveffect$_{GCS}$	Competence of public sector personnel; Quality of general infrastructure; Quality of public schools; Time spent by senior management dealing with government officials; Public Service vulnerability to political pressure; Wasteful government expenditure; Strength and expertise of the civil service to avoid drastic interruptions in government services in times of political instability; Government economic policies are independent of pressure from special interest groups.
goveffect$_{GRS}$	An increase in government personnel turnover rate at senior levels that reduces the GDP growth rate by 2% during any 12-month period; A decline in government personnel quality at any level that reduces the GDP growth rate by 1% during any 12-month period; A deterioration of government capacity to cope with national problems as a result of institutional rigidity or gridlock that reduces the GDP growth rate by 1% during any 12-month period.

continued on next page

Governance Variables - continued from previous page

Variable	Description (Kaufmann et al., 2006b)
goveffect$_{MIG}$	Bureaucracy
goveffect$_{PRS}$	Measures institutional strength and quality of the civil service, assess how much strength and expertise bureaucrats have and how able they are to manage political alternations without drastic interruptions in government services, or policy changes. Good performers have somewhat autonomous bureaucracies, free from political pressures, and an established mechanism for recruitment and training.
regulation$_{BCRI}$	Tax Effectiveness How efficient the countryŠs tax collection system is. The rules may be clear and transparent, but whether they are enforced consistently. This factor looks at the relative effectiveness too of corporate and personal, indirect and direct taxation. Legislation An assessment of whether the necessary business laws are in place, and whether there any outstanding gaps. This includes the extent to which the country's legislation is compatible with, and respected by, other countries' legal systems.
regulation$_{EIU}$	Unfair competitive practices; Price controls; Discriminatory tariffs; Excessive protections; Discriminatory taxes
regulation$_{GCS}$	Administrative regulations are burdensome; Tax system is distortionary; Import barriers as obstacle to growth; Competition in local market is limited; Anti-monopoly policy is lax and ineffective; Environmental regulations hurt competitiveness; Complexity of Tax System
regulation$_{GRS}$	Exports: A 2Imports: A 2Other Business: An increase in other regulatory burdens, with respect to the level at the time of the assessment, that reduces total aggregate investment in real LCU terms by 10Ownership of Business by Non-Residents: A 1-point increase on a scale from "0" to "10" in legal restrictions on ownership of business by non-residents during any 12-month period. Ownership of Equities by Non-Residents: A 1-point increase on a scale from "0" to "10" in legal restrictions on ownership of equities by non-residents during any 12-month period.

continued on next page

Governance Variables - continued from previous page

Variable	**Description (Kaufmann et al., 2006b)**
regulation_{HER}	Regulation; Government intervention; Wages/Prices; Trade; Foreign investment; Banking
regulation_{MIG}	Unfair trade; Unfair competition
regulation_{PRS}	Includes the risk to operations (scored from 0 to 4, increasing in risk); taxation (scored from 0 to 3), repatriation (scored from 0 to 3); repatriation (scored from 0 to 3) and labor costs (scored from 0 to 2). They all look at the governmentŠs attitude towards investment.
rulelaw_{BCRI}	Judicial Independence An assessment of how far the state and other outside actors can influence and distort the legal system. This will determine the level of legal impartiality investors can expect. Crime How much of a threat businesses face from crime such as kidnapping, extortion, street violence, burglary and so on. These problems can cause major inconvenience for foreign investors and require them to take expensive security precautions.
rulelaw_{EIU}	Violent crime; Organized crime; Fairness of judicial process; Enforceability of contracts; Speediness of judicial process; Confiscation/expropriation; Intellectual property rights protection; Private property protection
rulelaw_{GCS}	Common crime imposes costs on business; Organized crime imposes costs on business; Money laundering through banks is pervasive; Quality of Police; The judiciary is independent from political influences of government, citizens, or firms; Legal framework to challenge the legality of government actions is inefficient; Intellectual Property protection is weak; protection of financial assets is weak; Percentage of firms which are unofficial or unregistered / Tax evasion

continued on next page

Governance Variables - continued from previous page

Variable	Description (Kaufmann et al., 2006b)
rulelaw$_{GRS}$	Losses and Costs of Crime: A 1-point increase on a scale from "0" to "10" in crime during any 12-month period. Kidnapping of Foreigners: An increase in scope, intensity, or frequency of kidnapping of foreigners that reduces the GDP growth rate by 1% during any 12-month period. Enforceability of Government Contracts: A 1 point decline on a scale from "0" to "10" in the enforceability of contracts during any 12-month period. Enforceability of Private Contracts: A 1-point decline on a scale from "0" to "10" in the legal enforceability of contracts during any 12-month period.
rulelaw$_{HER}$	Black market; Property rights
rulelaw$_{HUM}$	Independence of judiciary
rulelaw$_{MIG}$	Legal safeguards; Organized crime
rulelaw$_{PRS}$	Law and Order. The Law sub-component is an assessment of the strength and impartiality of the legal system, while the Order sub-component is an assessment of popular observance of the law.
rulelaw$_{QLM}$	Enforceability of contracts; Direct financial fraud, money laundering and organized crime
stability$_{BCRI}$	Civil unrest How widespread political unrest is, and how great a threat it poses to investors. Demonstrations in themselves may not be cause for concern, but they will cause major disruption if they escalate into severe violence. At the extreme, this factor would amount to civil war. Terrorism Whether the country suffers from a sustained terrorist threat, and from how many sources. The degree of localisation of the threat is assessed, and whether the active groups are likely to target or affect businesses.

continued on next page

Governance Variables - continued from previous page

Variable	**Description (Kaufmann et al., 2006b)**
stability$_{EIU}$	Armed conflict; violent demonstrations; social unrest; international tensions/terrorist threat
stability$_{GCS}$	The threat of terrorism in the country imposes significant costs on business.
stability$_{GRS}$	Military Coup Risk: A military coup dŠetat (or a series of such events) that reduces the GDP growth rate by 2% during any 12-month period. Major Insurgency/Rebellion: An increase in scope or intensity of one or more insurgencies/rebellions that reduces the GDP growth rate by 3% during any 12-month period. Political Terrorism: An increase in scope or intensity of terrorism that reduces the GDP growth rate by 1% during any 12-month period. Political Assassination: A political assassination (or a series of such events) that reduces the GDP growth rate by 1% during any 12-month period. Civil War: An increase in scope or intensity of one or more civil wars that reduces the GDP growth rate by 4% during any 12-month period. Major Urban Riot: An increase in scope, intensity, or frequency of rioting that reduces the GDP growth rate by 1% during any 12-month period.

continued on next page

Governance Variables - continued from previous page

Variable	**Description (Kaufmann et al., 2006b)**
stability$_{HUM}$	Frequency of political killings; Frequency of disappearances; Frequency of tortures; Political terror scale
stability$_{IJT}$	Security risk rating
stability$_{MIG}$	Extremism
stability$_{PRS}$	Government Stability. Measures the governmentŠs ability to carry out its declared programs, and its ability to stay in office. This will depend on issues as: the type of governance, the cohesion of the government and governing party or parties, the closeness of the next election, the government command of the legislature, and approval of government policies. Internal Conflict. Assess political violence and its influence on governance. Highest scores go to countries with no armed opposition, and where the government does not indulge in arbitrary violence, direct or indirect. Lowest ratings go to civil war torn countries. Intermediate ratings are awarded on the basis of the threats to the government and busines. External conflict: The external conflict measure is an assessment both of the risk to the incumbent government and to inward investment. It ranges from trade restrictions and embargoes, whether imposed by a single country, a group of countries, or the international community as a whole, through geopolit

continued on next page

Governance Variables - continued from previous page

Variable	Description (Kaufmann et al., 2006b)
voice_{BCRI}	Institutional permanence An assessment of how mature and well-established the political system is. It is also an assessment of how far political opposition operates within the system or attempts to undermine it from outside. Representativeness How well the population and organised interests can make their voices heard in the political system. Provided representation is handled fairly and effectively, it will ensure greater stability and better designed policies.
voice_{EIU}	Orderly transfers, Vested Interests, Accountability of Public Officials, human Rights, Freedom of Association
voice_{FRH}	Political Rights (includes many subindices); Civil Liberties (includes many subindices); Freedom of the Press (includes many subindices)
voice_{GCS}	Newspapers can publish stories of their choosing without fear of censorship or retaliation; When deciding upon policies and contracts, Government officials favor well-connected firms; Effectiveness of national Parliament/Congress as a law making and oversight institution
voice_{HUM}	Restrictions on domestic and foreign travel; Freedom of political participation; Imprisonments because of ethnicity, race, or political, religious beliefs; Government censorship
voice_{PRS}	Military in Politics The military are not elected by anyone, so their participation in government, either direct or indirect, reduces accountability and therefore represents a risk. The threat of military intervention might lead as well to an anticipated potentially inefficient change in policy or even in government. It also works as an indication that the government is unable to function effectively and that the country has an uneasy environment for foreign business. Democratic Accountability. Quantifies how responsive government is to its people, on the basis that the less response there is the more likely is that the government will fall, peacefully or violently. It includes not only if free and fair elections are in place, but also how likely is the government to remain in power or remain popular.

continued on next page

Governance Variables - continued from previous page

Variable	Description (Kaufmann et al., 2006b)
voice_{RSF}	Press Freedom Index
For a detailed description of all data providers, see Kaufmann, Kraay and Mastruzzi "Governance Matters V: Appendices", World Bank, September 2006	

1.2.4 Summary Statistics

For reference, the table below presents summary statistics for each governance variable - once for the balanced dataset sample of 73 countries, and once using all available observations for each variable.

Summary Statistics, all governance variables

	All available obs.			*Balanced panel*		
Variable	**Mean**	**(Std. Dev.)**	**N**	**Mean**	**(Std. Dev.)**	**N**
cgdp95	7,364.22	(7,473.33)	186	10,053.94	(7,650.32)	73
$\text{corrupt}_{\text{GRS}}$	0.622	(0.27)	121	0.697	(0.259)	73
$\text{corrupt}_{\text{EIU}}$	0.336	(0.345)	119	0.421	(0.351)	73
$\text{corrupt}_{\text{GCS}}$	0.541	(0.181)	117	0.575	(0.175)	73
$\text{corrupt}_{\text{MIG}}$	0.291	(0.161)	153	0.331	(0.167)	73
$\text{corrupt}_{\text{PRS}}$	0.415	(0.204)	139	0.46	(0.218)	73
$\text{corrupt}_{\text{QLM}}$	0.377	(0.291)	115	0.443	(0.299)	73
$\text{corrupt}_{\text{BCRI}}$	0.55	(0.262)	198	0.615	(0.272)	73
$\text{stability}_{\text{GRS}}$	0.833	(0.147)	120	0.869	(0.133)	73
$\text{stability}_{\text{EIU}}$	0.565	(0.263)	118	0.62	(0.253)	73
$\text{stability}_{\text{GCS}}$	0.664	(0.134)	117	0.656	(0.149)	73
$\text{stability}_{\text{HUM}}$	0.647	(0.253)	190	0.614	(0.253)	73
$\text{stability}_{\text{IJT}}$	0.551	(0.257)	180	0.545	(0.206)	73
$\text{stability}_{\text{MIG}}$	0.368	(0.16)	152	0.36	(0.161)	73
$\text{stability}_{\text{PRS}}$	0.741	(0.104)	138	0.746	(0.097)	73

continued on next page

Summary statistics - continued from previous page

	All available obs.			*Balanced panel*		
Variable	**Mean**	**(Std. Dev.)**	**N**	**Mean**	**(Std. Dev.)**	**N**
$\text{stability}_{\text{BCRI}}$	0.696	(0.217)	198	0.686	(0.204)	73
$\text{voice}_{\text{EIU}}$	0.454	(0.284)	119	0.544	(0.272)	73
$\text{voice}_{\text{FRH}}$	0.595	(0.298)	194	0.676	(0.268)	73
$\text{voice}_{\text{GCS}}$	0.476	(0.138)	114	0.498	(0.138)	73
$\text{voice}_{\text{HUM}}$	0.628	(0.354)	182	0.659	(0.351)	73
$\text{voice}_{\text{PRS}}$	0.671	(0.256)	134	0.756	(0.212)	73
$\text{voice}_{\text{RSF}}$	0.747	(0.22)	158	0.801	(0.188)	73
$\text{voice}_{\text{BCRI}}$	0.583	(0.248)	192	0.698	(0.223)	73
$\text{regulation}_{\text{GRS}}$	0.885	(0.092)	121	0.904	(0.086)	73
$\text{regulation}_{\text{EIU}}$	0.543	(0.24)	119	0.616	(0.212)	73
$\text{regulation}_{\text{GCS}}$	0.486	(0.107)	115	0.505	(0.098)	73
$\text{regulation}_{\text{HER}}$	0.503	(0.178)	152	0.559	(0.16)	73
$\text{regulation}_{\text{MIG}}$	0.4	(0.138)	150	0.434	(0.147)	73
$\text{regulation}_{\text{PRS}}$	0.73	(0.215)	135	0.794	(0.201)	73
$\text{regulation}_{\text{BCRI}}$	0.599	(0.254)	197	0.707	(0.225)	73
$\text{rulelaw}_{\text{GRS}}$	0.785	(0.169)	121	0.831	(0.154)	73
$\text{rulelaw}_{\text{EIU}}$	0.512	(0.263)	119	0.585	(0.257)	73
$\text{rulelaw}_{\text{GCS}}$	0.527	(0.174)	117	0.560	(0.169)	73
$\text{rulelaw}_{\text{HER}}$	0.441	(0.278)	156	0.534	(0.285)	73
$\text{rulelaw}_{\text{HUM}}$	0.534	(0.384)	191	0.582	(0.391)	73
$\text{rulelaw}_{\text{MIG}}$	0.338	(0.136)	154	0.368	(0.146)	73
$\text{rulelaw}_{\text{PRS}}$	0.637	(0.21)	139	0.683	(0.195)	73
$\text{rulelaw}_{\text{QLM}}$	0.444	(0.301)	115	0.509	(0.302)	73
$\text{rulelaw}_{\text{BCRI}}$	0.608	(0.232)	201	0.666	(0.22)	73
$\text{goveffect}_{\text{GRS}}$	0.692	(0.207)	121	0.755	(0.184)	73
$\text{goveffect}_{\text{EGV}}$	0.256	(0.064)	194	0.278	(0.063)	73
$\text{goveffect}_{\text{EIU}}$	0.376	(0.306)	118	0.455	(0.309)	73
$\text{goveffect}_{\text{GCS}}$	0.509	(0.143)	114	0.542	(0.137)	73
$\text{goveffect}_{\text{MIG}}$	0.309	(0.144)	149	0.347	(0.161)	73
$\text{goveffect}_{\text{PRS}}$	0.536	(0.28)	135	0.653	(0.234)	73
$\text{goveffect}_{\text{BCRI}}$	0.582	(0.232)	191	0.681	(0.213)	73

2 Appendix to Chapter 2

2.1 Weak Instruments?

Some of my first-stage regressions exhibit either low R^2's, low F statistics, or both (see, Columns (4) and (5) of Panel B, Tables 15-18). In addition, the same instrument sometimes enters highly significant for multiple endogenous regressors (see Panel B, Tables 15-18). These are indications of potential weak-instrument problems. Low correlation between instruments and endogenous regressors, high correlation between instruments, or multiple endogenous regressors with overly similar dependencies on the excluded instruments all can cause the concentration parameter from the first-stage regressions to be small i.e. - can cause it's smallest eigenvalue to be small (Dollar, Kraay 2003; Stock and Yogo, 2005) .[71] When the concentration parameter is small, the 2SLS estimator of $\boldsymbol{\alpha}, \beta, \gamma$ in (1) no longer follows a normal distribution in finite samples, meaning standard errors, and therefore confidence intervals for $\boldsymbol{\alpha}, \beta, \gamma$, will be miscalculated. I may falsely conclude variables are significant when they are not. In addition, the 2SLS estimator is biased in the direction of the OLS estimator when instruments are weak (Bound, Jaeger and Baker, 1995; Stock and Yogo, 2005).

One approach to diagnosing weak instrument problems is to use a different estimator. Chao and Swanson (2005) demonstrate that the LIML IV estimator is more robust to instrument weakness than the 2SLS estimator, in the sense that consistency of LIML is preserved under weaker restrictions on the growth rate of the concentration

[71] Letting Z be the $n \times K$ matrix of observations on the instruments, letting Γ be the matrix of coefficients on the instruments from the first-stage regressions, and letting Σ be the VC matrix of disturbances from the first-stage regressions, the concentration parameter is written: $\Sigma^{-\frac{1}{2}}\Pi' Z' Z\Pi$ $\Sigma^{-\frac{1}{2}}$.

parameter (it is allowed to grow more slowly relative to the number of instruments) as $n \longrightarrow \infty$. Thus one can say in a crude sense that LIML estimates are consistent under a weaker set of assumptions on instrument strength than are 2SLS estimates. I therefore re-estimate all specifications in Tables 15-18 using the LIML estimator and find qualitatively similar results.

IV (LIML) Results - balanced dataset

Panel A: second-stage results (Dependent variable is log 2005 per capita income.)

	(1)	(2)	(3)	(4)	(5)	(6)
GEO	0.0186***	0.0303***	0.0165***	0.432	-0.0221	-0.0361
	(0.00690)	(0.00607)	(0.00592)	(5.448)	(0.0376)	(0.0363)
INT	0.149	0.405	0.261	9.361	-1.104	-1.325
	(0.250)	(0.315)	(0.216)	(119.0)	(1.378)	(1.141)
mrkt. infrastructure	0.607***		0.521***	-3.489	0.820*	0.983**
	(0.179)		(0.148)	(53.00)	(0.432)	(0.395)
civil liberties		0.437***	0.315***	-8.427	0.0457	0.447
		(0.168)	(0.116)	(114.3)	(0.389)	(0.328)
d-side. gov. risk				-20.03		0.900
				(262.0)		(0.751)
order					1.997	1.748
					(1.813)	(1.169)
Observations	64	64	64	64	64	64
R-squared	0.662	0.447	0.720			
Basmann p-val	0.00289	0.00256	0.0160	0.113	0.482	

Panel B: first-stage results - regressing INT and governance on the instruments

	(1)	(2)	(3)	(4)	(5)
	INT	mrkt. infrastructure	civil liberties	d-side. gov. risk	order
GEO	-0.00417	0.0279***	0.0143***	0.01000	0.00736
	(0.00330)	(0.00603)	(0.00520)	(0.00734)	(0.00762)
TRADESHARE	0.460***	0.342***	0.0566	0.123	0.320**
	(0.0650)	(0.119)	(0.102)	(0.144)	(0.150)
EURFRAC	-0.171	1.015***	0.0740	-0.0271	-0.000132
	(0.164)	(0.300)	(0.259)	(0.366)	(0.379)
$LEGOR_E$	-0.0408	0.836***	0.223	-0.353	-0.393
	(0.123)	(0.225)	(0.194)	(0.273)	(0.284)
$LEGOR_{GE}$	-0.0777	0.415	-0.0743	-0.292	-0.0286
	(0.219)	(0.401)	(0.346)	(0.488)	(0.506)
$LEGOR_{SC}$	-0.271	0.988**	0.157	-1.675***	0.392
	(0.264)	(0.482)	(0.416)	(0.587)	(0.609)
CATHO80	-0.210	-0.119	0.911***	-1.210**	-0.134
	(0.214)	(0.391)	(0.337)	(0.475)	(0.493)
MUSLIM80	-0.326*	0.426	-1.280***	-0.274	-0.303
	(0.179)	(0.326)	(0.281)	(0.397)	(0.412)
R-squared	0.536	0.599	0.629	0.261	0.218
F-stat.	7.937	10.28	11.67	2.431	1.912

Constants suppressed; standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A1

IV (LIML) Results - settler mortality sample

Panel A: second-stage results (Dependent variable is log 2005 per capita income.)

	(1)	(2)	(3)	(4)	(5)	(6)
GEO	-0.0764	0.0376***	0.00630	0.0155	0.0348	0.0326
	(0.0565)	(0.0125)	(0.0138)	(0.0355)	(0.0283)	(0.0299)
INT	-1.245	1.131**	0.519	0.264	2.031	1.757
	(1.414)	(0.451)	(0.351)	(0.899)	(1.241)	(1.564)
mrkt. infrastructure	3.472**		0.944***	1.834	0.374	0.642
	(1.419)		(0.308)	(1.421)	(0.590)	(1.134)
civil liberties		1.396***	1.072***	0.957*	1.165***	1.128***
		(0.295)	(0.204)	(0.497)	(0.302)	(0.336)
d-side. gov. risk				-1.887		-0.395
				(2.610)		(1.485)
order					-0.997	-0.851
					(0.750)	(0.904)
Observations	79	79	79	79	79	79
R-squared			0.439			
Basmann p-val	0.000820	0.0346	0.163	0.523	0.789	

Panel B: first-stage results - regressing INT and governance on the instruments

	(1)	(2)	(3)	(4)	(5)
	INT	mrkt. infrastructure	civil liberties	d-side. gov. risk	order
GEO	-0.00176	0.0160*	0.00327	0.0150	0.0273***
	(0.00558)	(0.00824)	(0.00853)	(0.0115)	(0.00989)
TRADESHARE	0.405***	0.0145	0.0171	-0.0246	0.621***
	(0.0670)	(0.0981)	(0.102)	(0.137)	(0.118)
SETMORT	-0.129**	-0.273***	0.0170	-0.0310	0.0893
	(0.0498)	(0.0730)	(0.0755)	(0.102)	(0.0876)
EURFRAC	-0.0663	0.377	0.975***	0.361	-0.190
	(0.160)	(0.235)	(0.243)	(0.328)	(0.282)
CATHO80	-0.201	-0.904***	0.0566	-0.683*	-0.254
	(0.187)	(0.274)	(0.283)	(0.382)	(0.329)
MUSLIM80	-0.236	-0.251	-0.517*	-0.0311	-0.848**
	(0.200)	(0.299)	(0.310)	(0.418)	(0.359)
R-squared	0.405	0.463	0.416	0.127	0.317
F-stat.	8.171	10.06	8.294	1.701	5.409

Constants suppressed; standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A2

IV (LIML) Results - Rodrik et al.'s large sample

Panel A: second-stage results (Dependent variable is log 2005 per capita income.)

	(1)	(2)	(3)	(4)	(5)	(6)
GEO	0.0157	0.0352***	0.0164**	0.0325	0.00839	0.0114
	(0.0100)	(0.00587)	(0.00713)	(0.0258)	(0.00899)	(0.0155)
INT	0.550**	0.520*	0.446**	0.548	-0.156	-0.117
	(0.274)	(0.306)	(0.222)	(0.386)	(0.459)	(0.509)
mrkt. infrastructure	0.877***		0.616***	0.664**	0.787***	0.788***
	(0.272)		(0.173)	(0.283)	(0.205)	(0.218)
civil liberties		0.729***	0.592***	0.409	0.540***	0.512***
		(0.177)	(0.128)	(0.324)	(0.134)	(0.183)
d-side. gov. risk				-0.919		-0.155
				(1.312)		(0.633)
order					0.555	0.535
					(0.365)	(0.394)
Observations	138	138	138	138	138	138
R-squared	0.560	0.281	0.586		0.542	0.478
Basmann p-val	3.21e-05	0.0165	0.173	0.164	0.273	0.149

Panel B: first-stage results - regressing INT and governance on the instruments

	(1)	(2)	(3)	(4)	(5)
	INT	mrkt. infrastructure	civil liberties	d-side. gov. risk	order
GEO	.0018	0.0287***	0.0121***	0.0199***	0.00186
	(0.00241)	(0.00429)	(0.00423)	(0.00534)	(0.00506)
TRADESHARE	0.363***	0.130*	0.0380	0.0415	0.436***
	(0.0424)	(0.0751)	(0.0741)	(0.0934)	(0.0885)
EURFRAC	-0.187*	0.660***	0.483**	0.0198	0.0150
	(0.110)	(0.195)	(0.192)	(0.243)	(0.230)
$LEGOR_E$	0.1458	0.616***	0.161	-0.00886	-0.470**
	(0.0873)	(0.155)	(0.153)	(0.193)	(0.183)
$LEGOR_{GE}$	-0.0578	0.893***	0.0691	0.134	-0.115
	(0.181)	(0.320)	(0.315)	(0.398)	(0.377)
$LEGOR_{SC}$	-0.240	1.080***	0.168	-1.036**	-0.126
	(0.218)	(0.385)	(0.380)	(0.479)	(0.454)
CATHO80	-0.0045	0.0501	0.431*	-0.458	-0.540*
	(0.147)	(0.261)	(0.257)	(0.324)	(0.307)
MUSLIM80	-0.2642**	0.693***	-0.956***	-0.0697	-0.705***
	(0.128)	(0.227)	(0.224)	(0.282)	(0.267)
R-squared	0.40	0.506	0.434	0.142	0.218
F-stat.	10.77	16.49	12.34	2.669	4.504

Constants suppressed; standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A3

IV (LIML) Results - full sample

Panel A: second-stage results (Dependent variable is log 2005 per capita income.)

	(1)	(2)	(3)	(4)	(5)	(6)
GEO	0.0188**	0.0349***	0.0168**	0.0434	0.00759	0.0680
	(0.00780)	(0.00593)	(0.00671)	(0.0414)	(0.0215)	(0.141)
INT	0.720**	0.568*	0.468**	0.731	-2.110	2.626
	(0.297)	(0.331)	(0.233)	(0.653)	(4.364)	(8.445)
mrkt. infrastructure	0.793***		0.651***	0.789*	0.900	0.717
	(0.214)		(0.167)	(0.426)	(0.550)	(0.695)
civil liberties		0.695***	0.549***	0.410	0.473	0.368
		(0.172)	(0.125)	(0.340)	(0.299)	(0.593)
d-side. gov. risk				-1.653		-2.794
				(2.378)		(7.149)
order					1.951	-1.313
					(3.275)	(5.667)
Observations	155	155	155	155	155	155
R-squared	0.548	0.230	0.542			
Basmann p-val	4.64e-05	0.00619	0.141	0.231	0.160	0.126

Panel B: first-stage results - regressing INT and governance on the instruments

	(1)	(2)	(3)	(4)	(5)
	INT	mrkt. infrastructure	civil liberties	d-side. gov. risk	order
GEO	0.005*	0.0283***	0.0107**	0.0215***	0.00173
	(0.00289)	(0.00407)	(0.00422)	(0.00554)	(0.00505)
TRADESHARE	0.373***	0.146**	0.0219	0.0695	0.500***
	(0.0500)	(0.0701)	(0.0725)	(0.0952)	(0.0868)
EURFRAC	-0.159	0.632***	0.515***	0.0819	-0.0211
	(0.132)	(0.185)	(0.192)	(0.252)	(0.230)
$LEGOR_E$	0.054	0.655***	0.110	0.0482	-0.310*
	(0.0999)	(0.141)	(0.146)	(0.191)	(0.174)
$LEGOR_{GE}$	-0.161	0.900***	0.0774	0.167	-0.0713
	(0.222)	(0.311)	(0.322)	(0.423)	(0.386)
$LEGOR_{SC}$	-0.401	1.080***	0.223	-0.971*	-0.110
	(0.266)	(0.372)	(0.385)	(0.506)	(0.461)
CATHO80	-0.054	0.0563	0.427	-0.313	-0.517*
	(0.178)	(0.250)	(0.258)	(0.339)	(0.309)
MUSLIM80	-0.345**	0.621***	-0.944***	-0.0832	-0.709***
	(0.142)	(0.199)	(0.206)	(0.270)	(0.246)
R-squared	0.307	0.495	0.437	0.122	0.217
F-stat.	8.10	17.87	14.14	2.545	5.052

Constants suppressed; standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A4

Another approach for assessing weak instrument problems is that of Stock and

Yogo (e.g. - 2005). These authors develop two alternative formal definitions of weak instruments and then tabulate cutoff values for the Cragg-Donald statistic (1993), thereby enabling formal testing of the null hypothesis that instruments are weak. The first criterion deems instruments weak if the bias of the IV estimator relative to the bias of the OLS estimator could potentially reach some threshold (e.g. - 20%). The second criterion deems instruments weak if the actual size of a Wald test on the IV parameters exceeds the test's nominal size (α) by some specified magnitude. STATA implements the Stock and Yogo tests based on the relative bias definition. I show those results in the tables below. Note that I can only implement Stock and Yogo's approach for specifications with three or fewer endogenous regressors because those authors have not tabulated cutoff values for specifications involving more than three endogenous regressors.

Stock and Yogo (2005) Tests of Weak Instruments, Selected Specifications

	(IV estimator is 2SLS; dependent variable is log 2005 per capita income throughout.)			
	balanced panel	settler mortality sample	Rodrik et al.'s large sample	fullest sample
GEO	0.0152***	0.00674	0.0160**	0.0166***
	(0.00546)	(0.0128)	(0.00674)	(0.00642)
INT	0.284	0.504	0.451**	0.477**
	(0.192)	(0.320)	(0.210)	(0.216)
market infrastructure	0.567***	0.946***	0.643***	0.672***
	(0.122)	(0.273)	(0.157)	(0.154)
civil liberties	0.311***	1.008***	0.553***	0.511***
	(0.104)	(0.188)	(0.120)	(0.118)
constant	7.389***	6.337***	6.458***	6.307***
	(0.868)	(1.538)	(0.962)	(0.980)
Observations	64	79	138	155
R-squared	0.727	0.471	0.600	0.556
Cragg-Donald F-statistic	5.388	3.805	5.767	7.18
10% maximal IV relative bias	8.5	6.61	8.5	8.5
20% maximal IV relative bias	5.56	4.99	5.56	5.56
30% maximal IV relative bias	4.44	4.3	4.44	4.44

*Constants suppressed; standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1*

Instruments for Cols. (1), (3), and (4): GEO, eurfrac, legal origin dummies, catho80, muslim80

Instruments for Col. (2): GEO, eurfrac, SETMORT, catho80, muslim80

Table A5

For the specification shown above, I can reject 20% maximal IV relative bias in my two largest samples (the Rodrik et al. large sample, and the fullest possible sample). I can reject only 30% maximal IV relative bias in the balanced panel sample, and in the settler mortality sample (where the estimation uses a different set of instruments) I cannot reject even 30% maximal IV relative bias. I can take some heart in the fact that my two largest samples enable the strongest degree of rejection of weak instruments. However, the fact that I cannot reject smaller than 20% maximal IV relative bias suggests I should base inference about structural parameters on weak-instrument-robust confidence intervals.

Additionally, the specifications shown above include the two governance factors that my instruments explain best. Performing the same tests with either or both of my downside governance risk and order factors (which have lower R^2 and F statistics in first-stage regressions) leads to an inability to reject even 30% maximal IV relative bias. Overall, then, I can say that weak instrument problems appear to be driven principally by downside governance risk and order. Barring new instruments that can better explain the variation in these two variables, a prudent course of action may be to omit them from my growth regressions.

Correlation of fitted regressors

(N=138)	GEO	$\hat{INT}$	$\hat{f}_1$	$\hat{f}_2$	$\hat{f}_3$	$\hat{f}_4$
GEO	1					
$\hat{INT}$	-0.0112	1				
$\hat{f}_1$	0.794	0.080	1			
$\hat{f}_2$	0.374	0.073	0.338	1		
$\hat{f}_3$	0.782	0.081	0.641	0.046	1	
$\hat{f}_4$	0.196	0.754	0.062	0.182	0.141	1

Note: Rodrik et al. large sample; excluded inst's.=eurfrac, legal origin, catho80, muslim80

Table A6

3 Appendix to Chapter 4

3.1 Comparison of $\widehat{E}[f|X^*]$ to f

Tables 25-32 and 34-37 present comparisons of $\widehat{E}[f|X^*]$ with $\widehat{E}[f|X]$ (rather than with f), reflecting my interest in assessing the distortion in factor score prediction due exclusively to the polytomous nature of governance data. In those results, I demonstrate that over a wide variety of factor model DGPs, $\widehat{E}[f|X^*]$ and $\widehat{E}[f|X]$ are highly correlated. However, I did not produce evidence on the absolute accuracy of $\widehat{E}[f|X^*]$, i.e. how close the polytomous-data predictor $\widehat{E}[f|X^*]$ comes to a country's true factor score, f. Tables A7-A14 below present these results. The results were compiled during the same simulation runs as the results in Tables 25-32. Tables A7-A14 are structured exactly like Tables 25-32, but with the basis of comparison for $\widehat{E}[f|X^*]$ switched from $\widehat{E}[f|X]$ to f.

There is little changed from Tables 25-32 except mean-squared error values are generally a bit higher, and correlations generally a bit lower. The influences of loadings complexity, number of factors, number of score categories and degree of censoring on prediction accuracy appear unchanged from those observed in Tables 25-32.

Simulation Results - Simple One-Factor Models

(a,b)	$E[(\hat{f}^{*}-f)^2]$ f_1	$\rho(\hat{f}^{*},f)$ f_1	(a,b)	$E[(\hat{f}^{*}-f)^2]$ f_1	$\rho(\hat{f}^{*},f)$ f_1
	(categories = 2)			*(categories = 3)*	
(0.1,1)	0.68	0.63	(0.1,1)	0.23	0.89
(0.5,1)	0.30	0.85	(0.5,1)	0.25	0.87
(0.75,1)	0.38	0.80	(0.75,1)	0.44	0.77
(0.4,0.6)	-	-	(0.4,0.6)	0.26	0.87
(0.1,0.9)	-	-	(0.1,0.9)	0.31	0.84
	(categories = 7)			*(categories = 15)*	
(0.1,1)	0.18	0.92	(0.1,1)	0.17	0.92
(0.5,1)	0.28	0.86	(0.5,1)	0.30	0.85
(0.75,1)	0.50	0.73	(0.75,1)	0.54	0.71
(0.4,0.6)	0.25	0.87	(0.4,0.6)	0.26	0.87
(0.1,0.9)	0.17	0.92	(0.1,0.9)	0.17	0.92
	(categories = 25)			*(categories = 50)*	
(0.1,1)	0.18	0.92	(0.1,1)	0.18	0.91
(0.5,1)	0.30	0.85	(0.5,1)	0.31	0.84
(0.75,1)	0.55	0.70	(0.75,1)	0.54	0.71
(0.4,0.6)	0.26	0.87	(0.4,0.6)	0.26	0.87
(0.1,0.9)	0.16	0.92	(0.1,0.9)	0.18	0.91
	(categories = 250)			*(categories = 500)*	
(0.1,1)	0.18	0.92	(0.1,1)	0.18	0.91
(0.5,1)	0.30	0.85	(0.5,1)	0.31	0.84
(0.75,1)	0.55	0.70	(0.75,1)	0.54	0.71
(0.4,0.6)	0.26	0.87	(0.4,0.6)	0.26	0.87
(0.1,0.9)	0.16	0.92	(0.1,0.9)	0.18	0.91

Design parameters: p=7, n=75, lambda=0.7, psi=0.51, 100 samples per DGP

Table A7

Simulation Results - Simple Two-Factor Models

	$E[(\hat{f}^* - f)^2]$		$\rho(\hat{f}^*, f)$			$E[(\hat{f}^* - f)^2]$		$\rho(\hat{f}^*, f)$	
(a,b)	f_1	f_2	f_1	f_2	(a,b)	f_1	f_2	f_1	f_2
	(categories = 2)					*(categories = 3)*			
(0.1,1)	0.75	0.72	0.60	0.61	(0.1,1)	0.23	0.23	0.88	0.88
(0.5,1)	0.31	0.30	0.84	0.84	(0.5,1)	0.26	0.25	0.86	0.87
(0.75,1)	0.41	0.40	0.78	0.78	(0.75,1)	0.42	0.43	0.78	0.77
(0.4,0.6)	-	-	-	-	(0.4,0.6)	0.25	0.26	0.87	0.87
(0.1,0.9)	-	-	-	-	(0.1,0.9)	0.32	0.30	0.83	0.84
	(categories = 7)					*(categories = 15)*			
(0.1,1)	0.18	0.18	0.92	0.91	(0.1,1)	0.18	0.17	0.92	0.92
(0.5,1)	0.29	0.28	0.85	0.86	(0.5,1)	0.30	0.31	0.85	0.84
(0.75,1)	0.48	0.50	0.74	0.73	(0.75,1)	0.53	0.52	0.72	0.72
(0.4,0.6)	0.26	0.27	0.87	0.86	(0.4,0.6)	0.27	0.26	0.86	0.86
(0.1,0.9)	0.18	0.18	0.92	0.91	(0.1,0.9)	0.18	0.18	0.92	0.92
	(categories = 25)					*(categories = 50)*			
(0.1,1)	0.18	0.17	0.91	0.91	(0.1,1)	0.17	0.18	0.91	0.91
(0.5,1)	0.31	0.32	0.84	0.84	(0.5,1)	0.31	0.31	0.84	0.84
(0.75,1)	0.54	0.56	0.71	0.70	(0.75,1)	0.57	0.56	0.69	0.71
(0.4,0.6)	0.27	0.27	0.86	0.86	(0.4,0.6)	0.26	0.27	0.87	0.86
(0.1,0.9)	0.17	0.18	0.92	0.92	(0.1,0.9)	0.18	0.18	0.91	0.91
	(categories = 250)					*(categories = 500)*			
(0.1,1)	0.17	0.18	0.91	0.92	(0.1,1)	0.18	0.17	0.91	0.92
(0.5,1)	0.33	0.33	0.83	0.83	(0.5,1)	0.32	0.32	0.84	0.83
(0.75,1)	0.58	0.58	0.69	0.70	(0.75,1)	0.58	0.57	0.69	0.69
(0.4,0.6)	0.27	0.28	0.86	0.86	(0.4,0.6)	0.27	0.27	0.87	0.86
(0.1,0.9)	0.18	0.17	0.92	0.92	(0.1,0.9)	0.18	0.17	0.91	0.92

p=14, n=75, lambda=0.7 or 0, psi=0.51, 100 samples per DGP

Table A8

Simulation Results - Simple Three-Factor Models

	$E[(\hat{f}^* - f)^2]$			$\rho(\hat{f}^*, f)$				$E[(\hat{f}^* - f)^2]$			$\rho(\hat{f}^*, f)$		
(a,b)	f_1	f_2	f_3	f_1	f_2	f_3	(a,b)	f_1	f_2	f_3	f_1	f_2	f_3
			(categories = 2)							*(categories = 3)*			
(0.1,1)	0.78	0.76	0.79	0.58	0.59	0.57	(0.1,1)	0.24	0.24	0.24	0.88	0.88	0.88
(0.5,1)	0.31	0.33	0.31	0.83	0.83	0.84	(0.5,1)	0.27	0.25	0.26	0.86	0.87	0.87
(0.75,1)	0.41	0.43	0.42	0.79	0.77	0.77	(0.75,1)	0.44	0.43	0.43	0.77	0.78	0.77
(0.4,0.6)	-	-	-	-	-	-	(0.4,0.6)	0.26	0.26	0.26	0.86	0.87	0.87
(0.1,0.9)	-	-	-	-	-	-	(0.1,0.9)	0.33	0.33	0.33	0.83	0.82	0.82
			(categories = 7)							*(categories = 15)*			
(0.1,1)	0.19	0.18	0.19	0.91	0.91	0.91	(0.1,1)	0.18	0.18	0.17	0.92	0.91	0.92
(0.5,1)	0.29	0.28	0.29	0.85	0.85	0.85	(0.5,1)	0.31	0.32	0.31	0.84	0.84	0.85
(0.75,1)	0.51	0.51	0.50	0.73	0.73	0.73	(0.75,1)	0.55	0.54	0.56	0.71	0.71	0.70
(0.4,0.6)	0.27	0.28	0.27	0.86	0.86	0.86	(0.4,0.6)	0.28	0.27	0.27	0.86	0.86	0.86
(0.1,0.9)	0.17	0.17	0.18	0.92	0.91	0.91	(0.1,0.9)	0.18	0.18	0.18	0.91	0.91	0.91
			(categories = 25)							*(categories = 50)*			
(0.1,1)	0.19	0.19	0.19	0.91	0.91	0.91	(0.1,1)	0.19	0.18	0.19	0.91	0.91	0.91
(0.5,1)	0.32	0.32	0.31	0.84	0.84	0.84	(0.5,1)	0.33	0.33	0.33	0.84	0.83	0.83
(0.75,1)	0.54	0.59	0.56	0.71	0.68	0.70	(0.75,1)	0.58	0.60	0.57	0.69	0.68	0.69
(0.4,0.6)	0.27	0.27	0.28	0.86	0.86	0.86	(0.4,0.6)	0.28	0.28	0.27	0.86	0.86	0.86
(0.1,0.9)	0.18	0.17	0.18	0.91	0.91	0.91	(0.1,0.9)	0.18	0.17	0.19	0.91	0.91	0.91
			(categories = 250)							*(categories = 500)*			
(0.1,1)	0.18	0.18	0.18	0.91	0.91	0.91	(0.1,1)	0.18	0.18	0.18	0.91	0.91	0.91
(0.5,1)	0.34	0.33	0.32	0.83	0.83	0.83	(0.5,1)	0.34	0.32	0.34	0.83	0.84	0.83
(0.75,1)	0.60	0.59	0.59	0.68	0.68	0.68	(0.75,1)	0.58	0.60	0.56	0.69	0.68	0.70
(0.4,0.6)	0.27	0.28	0.27	0.86	0.86	0.86	(0.4,0.6)	0.28	0.27	0.28	0.86	0.86	0.86
(0.1,0.9)	0.18	0.18	0.19	0.92	0.91	0.91	(0.1,0.9)	0.19	0.18	0.18	0.91	0.92	0.91

p=21, n=75, lambda=0.7 or 0, psi=0.51, 100 samples per DGP

Table A9

Simulation Results - Complex Two-Factor Models

(a,b)	$E[(\hat{f}^* - f)^2]$ f_1	f_2	$\rho(\hat{f}^*, f)$ f_1	f_2	(a,b)	$E[(\hat{f}^* - f)^2]$ f_1	f_2	$\rho(\hat{f}^*, f)$ f_1	f_2
		(categories = 2)					*(categories = 3)*		
(0.1,1)	0.97	1.02	0.48	0.47	(0.1,1)	0.45	0.44	0.76	0.77
(0.5,1)	0.56	0.58	0.70	0.69	(0.5,1)	0.46	0.46	0.76	0.76
(0.75,1)	0.70	0.71	0.63	0.62	(0.75,1)	0.70	0.72	0.62	0.62
(0.4,0.6)	-	-	-	-	(0.4,0.6)	0.48	0.48	0.74	0.75
(0.1,0.9)	-	-	-	-	(0.1,0.9)	0.63	0.63	0.66	0.65
		(categories = 7)					*(categories = 15)*		
(0.1,1)	0.32	0.32	0.84	0.84	(0.1,1)	0.30	0.29	0.85	0.85
(0.5,1)	0.47	0.48	0.75	0.74	(0.5,1)	0.52	0.51	0.72	0.73
(0.75,1)	0.77	0.76	0.59	0.60	(0.75,1)	0.82	0.81	0.56	0.58
(0.4,0.6)	0.49	0.50	0.74	0.74	(0.4,0.6)	0.49	0.49	0.74	0.74
(0.1,0.9)	0.29	0.30	0.85	0.85	(0.1,0.9)	0.29	0.28	0.85	0.86
		(categories = 25)					*(categories = 50)*		
(0.1,1)	0.31	0.31	0.84	0.84	(0.1,1)	0.30	0.31	0.84	0.84
(0.5,1)	0.52	0.51	0.73	0.73	(0.5,1)	0.52	0.51	0.73	0.73
(0.75,1)	0.81	0.86	0.57	0.54	(0.75,1)	0.87	0.85	0.54	0.56
(0.4,0.6)	0.51	0.52	0.72	0.72	(0.4,0.6)	0.50	0.53	0.74	0.72
(0.1,0.9)	0.29	0.29	0.85	0.85	(0.1,0.9)	0.29	0.30	0.85	0.85
		(categories = 250)					*(categories = 500)*		
(0.1,1)	0.29	0.28	0.86	0.86	(0.1,1)	0.30	0.30	0.85	0.84
(0.5,1)	0.51	0.53	0.73	0.73	(0.5,1)	0.53	0.51	0.73	0.73
(0.75,1)	0.84	0.86	0.56	0.55	(0.75,1)	0.85	0.86	0.55	0.55
(0.4,0.6)	0.50	0.51	0.74	0.73	(0.4,0.6)	0.52	0.53	0.72	0.71
(0.1,0.9)	0.30	0.30	0.85	0.85	(0.1,0.9)	0.30	0.30	0.84	0.85

p=14, n=75, lambda=0.7 or 0.4, psi=0.4, 100 samples per DGP

Table A10

Simulation Results - Complex Three-Factor Models

	$E[(\hat{f}^{*}-f)^2]$			$\rho(\hat{f}^{*},f)$				$E[(\hat{f}^{*}-f)^2]$			$\rho(\hat{f}^{*},f)$		
(a,b)	f_1	f_2	f_3	f_1	f_2	f_3	(a,b)	f_1	f_2	f_3	f_1	f_2	f_3
			(categories = 2)							*(categories = 3)*			
(0.1,1)	1.08	1.08	1.09	0.44	0.44	0.43	(0.1,1)	0.46	0.46	0.43	0.76	0.76	0.77
(0.5,1)	0.61	0.61	0.62	0.68	0.68	0.67	(0.5,1)	0.47	0.46	0.44	0.75	0.76	0.76
(0.75,1)	0.74	0.78	0.75	0.60	0.59	0.60	(0.75,1)	0.71	0.70	0.71	0.62	0.62	0.62
(0.4,0.6)	-	-	-	-	-	-	(0.4,0.6)	0.48	0.48	0.47	0.75	0.75	0.75
(0.1,0.9)	-	-	-	-	-	-	(0.1,0.9)	0.69	0.64	0.62	0.63	0.65	0.67
			(categories = 7)							*(categories = 15)*			
(0.1,1)	0.28	0.29	0.27	0.85	0.85	0.86	(0.1,1)	0.28	0.28	0.28	0.86	0.86	0.85
(0.5,1)	0.46	0.47	0.45	0.76	0.75	0.77	(0.5,1)	0.49	0.52	0.51	0.74	0.73	0.74
(0.75,1)	0.79	0.80	0.78	0.58	0.58	0.60	(0.75,1)	0.87	0.87	0.84	0.55	0.54	0.56
(0.4,0.6)	0.49	0.49	0.48	0.74	0.74	0.75	(0.4,0.6)	0.51	0.50	0.53	0.73	0.74	0.72
(0.1,0.9)	0.28	0.27	0.28	0.86	0.86	0.86	(0.1,0.9)	0.28	0.27	0.27	0.86	0.86	0.86
			(categories = 25)							*(categories = 50)*			
(0.1,1)	0.30	0.28	0.29	0.85	0.86	0.85	(0.1,1)	0.28	0.30	0.30	0.86	0.85	0.85
(0.5,1)	0.52	0.51	0.51	0.73	0.73	0.74	(0.5,1)	0.52	0.51	0.53	0.73	0.74	0.72
(0.75,1)	0.85	0.83	0.89	0.56	0.56	0.52	(0.75,1)	0.89	0.90	0.88	0.53	0.53	0.54
(0.4,0.6)	0.52	0.52	0.48	0.73	0.72	0.74	(0.4,0.6)	0.51	0.51	0.54	0.73	0.73	0.71
(0.1,0.9)	0.27	0.28	0.29	0.86	0.86	0.86	(0.1,0.9)	0.28	0.28	0.29	0.86	0.86	0.85
			(categories = 250)							*(categories = 500)*			
(0.1,1)	0.29	0.27	0.28	0.85	0.86	0.86	(0.1,1)	0.28	0.29	0.27	0.86	0.85	0.86
(0.5,1)	0.52	0.50	0.52	0.73	0.73	0.73	(0.5,1)	0.52	0.53	0.53	0.73	0.72	0.71
(0.75,1)	0.88	0.87	0.87	0.54	0.54	0.54	(0.75,1)	0.88	0.87	0.87	0.54	0.55	0.54
(0.4,0.6)	0.52	0.51	0.51	0.73	0.73	0.73	(0.4,0.6)	0.50	0.50	0.51	0.73	0.73	0.73
(0.1,0.9)	0.29	0.29	0.30	0.85	0.85	0.85	(0.1,0.9)	0.28	0.28	0.29	0.86	0.86	0.86

p=21, n=75, lambda=0.7,0.4 or 0.3, psi=0.51, 100 samples per DGP

Table A11

Simulation Results - Complex Five-Factor Models

	$E[(\hat{f}^* - f)^2]$					$\rho(\hat{f}^*, f)$						$E[(\hat{f}^* - f)^2]$					$\rho(\hat{f}^*, f)$				
(a,b)	f_1	f_2	f_3	f_4	f_5	f_1	f_2	f_3	f_4	f_5	(a,b)	f_1	f_2	f_3	f_4	f_5	f_1	f_2	f_3	f_4	f_5
					(categories = 2)											*(categories = 3)*					
(0.1,1)	0.88	0.93	0.80	1.02	1.08	0.53	0.60	0.62	0.48	0.47	(0.1,1)	0.36	0.44	0.40	0.49	0.81	0.83	0.78	0.80	0.74	0.56
(0.5,1)	0.51	0.61	0.59	0.69	1.17	0.74	0.70	0.70	0.64	0.37	(0.5,1)	0.37	0.42	0.41	0.46	0.76	0.82	0.79	0.80	0.76	0.60
(0.75,1)	0.65	0.73	0.73	0.86	1.33	0.67	0.63	0.63	0.54	0.31	(0.75,1)	0.60	0.67	0.69	0.75	1.15	0.70	0.67	0.65	0.61	0.40
(0.4,0.6)	-	-	-	-	-	-	-	-	-	-	(0.4,0.6)	0.42	0.47	0.46	0.57	0.73	0.79	0.77	0.76	0.70	0.61
(0.1,0.9)	-	-	-	-	-	-	-	-	-	-	(0.1,0.9)	0.54	0.63	0.66	0.80	1.16	0.72	0.69	0.67	0.58	0.37
					(categories = 7)											*(categories = 15)*					
(0.1,1)	0.17	0.18	0.17	0.27	0.75	0.93	0.92	0.92	0.86	0.61	(0.1,1)	0.16	0.17	0.18	0.30	0.82	0.92	0.92	0.92	0.84	0.58
(0.5,1)	0.38	0.42	0.42	0.49	0.62	0.82	0.80	0.79	0.74	0.66	(0.5,1)	0.43	0.46	0.42	0.54	0.70	0.79	0.77	0.79	0.72	0.62
(0.75,1)	0.70	0.75	0.77	0.90	1.20	0.65	0.63	0.60	0.54	0.38	(0.75,1)	0.73	0.78	0.83	0.94	1.25	0.63	0.61	0.58	0.52	0.35
(0.4,0.6)	0.45	0.51	0.50	0.54	0.85	0.78	0.75	0.75	0.72	0.54	(0.4,0.6)	0.47	0.53	0.53	0.54	0.91	0.77	0.74	0.73	0.72	0.52
(0.1,0.9)	0.18	0.19	0.17	0.24	0.39	0.92	0.91	0.92	0.88	0.80	(0.1,0.9)	0.19	0.20	0.19	0.25	0.37	0.92	0.91	0.91	0.88	0.80
					(categories = 25)											*(categories = 50)*					
(0.1,1)	0.19	0.18	0.19	0.27	1.17	0.91	0.92	0.91	0.87	0.39	(0.1,1)	0.19	0.17	0.20	0.30	1.02	0.92	0.92	0.90	0.85	0.48
(0.5,1)	0.44	0.47	0.48	0.54	0.73	0.78	0.77	0.76	0.72	0.61	(0.5,1)	0.45	0.46	0.48	0.56	0.72	0.78	0.77	0.76	0.72	0.62
(0.75,1)	0.72	0.78	0.83	0.95	1.18	0.63	0.60	0.59	0.51	0.39	(0.75,1)	0.78	0.82	0.85	1.01	1.28	0.62	0.59	0.57	0.48	0.33
(0.4,0.6)	0.47	0.54	0.52	0.55	0.95	0.77	0.74	0.73	0.71	0.50	(0.4,0.6)	0.46	0.53	0.53	0.57	0.85	0.77	0.74	0.73	0.70	0.55
(0.1,0.9)	0.21	0.20	0.20	0.26	0.39	0.91	0.90	0.91	0.87	0.80	(0.1,0.9)	0.20	0.22	0.20	0.27	0.38	0.91	0.90	0.91	0.87	0.80
					(categories = 250)											*(categories = 500)*					
(0.1,1)	0.19	0.18	0.17	0.25	0.72	0.91	0.92	0.92	0.87	0.62	(0.1,1)	0.19	0.18	0.18	0.30	0.81	0.91	0.92	0.91	0.85	0.59
(0.5,1)	0.47	0.47	0.49	0.55	0.74	0.77	0.77	0.76	0.71	0.61	(0.5,1)	0.47	0.50	0.49	0.56	0.68	0.77	0.76	0.76	0.72	0.64
(0.75,1)	0.76	0.80	0.84	0.98	1.31	0.61	0.60	0.58	0.49	0.32	(0.75,1)	0.76	0.80	0.85	0.97	1.23	0.62	0.59	0.56	0.50	0.36
(0.4,0.6)	0.46	0.55	0.52	0.57	0.80	0.78	0.73	0.73	0.70	0.57	(0.4,0.6)	0.48	0.55	0.51	0.59	0.92	0.77	0.73	0.74	0.69	0.52
(0.1,0.9)	0.19	0.21	0.20	0.28	0.38	0.91	0.90	0.91	0.86	0.81	(0.1,0.9)	0.20	0.23	0.20	0.27	0.40	0.91	0.89	0.91	0.87	0.79

p=45, n=75, lambda, psi - see Appendix, 100 samples per DGP

Table A12

Simulation Results - Complex, Negative Two-Factor Models

	$E[(\hat{f}^* - f)^2]$		$\rho(\hat{f}^*, f)$			$E[(\hat{f}^* - f)^2]$		$\rho(\hat{f}^*, f)$	
(a,b)	f_1	f_2	f_1	f_2	(a,b)	f_1	f_2	f_1	f_2
	(categories = 2)					*(categories = 3)*			
(0.1,1)	0.90	1.42	0.52	0.22	(0.1,1)	0.49	0.94	0.73	0.48
(0.5,1)	0.55	1.07	0.70	0.41	(0.5,1)	0.54	0.87	0.71	0.52
(0.75,1)	0.63	1.17	0.65	0.36	(0.75,1)	0.67	1.11	0.64	0.40
(0.4,0.6)	-	-	-	-	(0.4,0.6)	0.53	0.88	0.72	0.53
(0.1,0.9)	-	-	-	-	(0.1,0.9)	0.57	1.09	0.69	0.40
	(categories = 7)					*(categories = 15)*			
(0.1,1)	0.51	0.74	0.73	0.60	(0.1,1)	0.49	0.75	0.74	0.60
(0.5,1)	0.58	0.87	0.69	0.54	(0.5,1)	0.59	0.90	0.68	0.51
(0.75,1)	0.75	1.22	0.60	0.34	(0.75,1)	0.78	1.20	0.59	0.36
(0.4,0.6)	0.54	0.89	0.71	0.51	(0.4,0.6)	0.54	0.98	0.71	0.46
(0.1,0.9)	0.48	0.76	0.75	0.59	(0.1,0.9)	0.50	0.74	0.74	0.60
	(categories = 25)					*(categories = 50)*			
(0.1,1)	0.50	0.75	0.74	0.60	(0.1,1)	0.49	0.74	0.74	0.61
(0.5,1)	0.61	0.93	0.68	0.49	(0.5,1)	0.60	0.98	0.68	0.47
(0.75,1)	0.79	1.23	0.58	0.35	(0.75,1)	0.78	1.30	0.59	0.31
(0.4,0.6)	0.53	0.91	0.71	0.50	(0.4,0.6)	0.53	0.93	0.71	0.49
(0.1,0.9)	0.50	0.75	0.74	0.59	(0.1,0.9)	0.49	0.74	0.73	0.62
	(categories = 250)					*(categories = 500)*			
(0.1,1)	0.49	0.77	0.74	0.59	(0.1,1)	0.49	0.74	0.74	0.61
(0.5,1)	0.60	0.99	0.68	0.47	(0.5,1)	0.59	0.94	0.68	0.50
(0.75,1)	0.78	1.29	0.58	0.30	(0.75,1)	0.82	1.25	0.57	0.33
(0.4,0.6)	0.54	0.95	0.72	0.50	(0.4,0.6)	0.55	0.97	0.72	0.46
(0.1,0.9)	0.49	0.77	0.74	0.60	(0.1,0.9)	0.50	0.75	0.74	0.60

p=14, n=75, lambda=0.7, 0.4, 0 or -0.3; psi = various; 100 samples per DGP

Table A13

Simulation Results - Complex, Negative Three-Factor Models

	$E[(\hat{f}^* - f)^2]$			$\rho(\hat{f}^*, f)$				$E[(\hat{f}^* - f)^2]$			$\rho(\hat{f}^*, f)$		
(a,b)	f_1	f_2	f_3	f_1	f_2	f_3	(a,b)	f_1	f_2	f_3	f_1	f_2	f_3
			(categories = 2)							*(categories = 3)*			
(0.1,1)	0.93	0.92	1.02	0.50	0.52	0.46	(0.1,1)	0.61	0.43	0.79	0.68	0.77	0.57
(0.5,1)	0.68	0.52	0.82	0.63	0.72	0.56	(0.5,1)	0.44	0.43	0.71	0.77	0.77	0.62
(0.75,1)	0.52	0.61	0.80	0.72	0.68	0.57	(0.75,1)	0.55	0.60	0.77	0.71	0.68	0.58
(0.4,0.6)	-	-	-	-	-	-	(0.4,0.6)	0.59	0.45	0.76	0.69	0.77	0.59
(0.1,0.9)	-	-	-	-	-	-	(0.1,0.9)	0.73	0.53	0.83	0.62	0.72	0.55
			(categories = 7)							*(categories = 15)*			
(0.1,1)	0.50	0.36	0.82	0.74	0.82	0.57	(0.1,1)	0.52	0.37	0.76	0.73	0.81	0.60
(0.5,1)	0.41	0.43	0.66	0.79	0.78	0.65	(0.5,1)	0.43	0.46	0.67	0.77	0.77	0.64
(0.75,1)	0.63	0.68	0.87	0.67	0.65	0.55	(0.75,1)	0.63	0.72	0.89	0.67	0.61	0.54
(0.4,0.6)	0.66	0.46	0.78	0.65	0.76	0.58	(0.4,0.6)	0.61	0.47	0.81	0.67	0.76	0.58
(0.1,0.9)	0.56	0.35	0.79	0.70	0.83	0.58	(0.1,0.9)	0.47	0.35	0.78	0.75	0.83	0.59
			(categories = 25)							*(categories = 50)*			
(0.1,1)	0.54	0.35	0.78	0.71	0.82	0.58	(0.1,1)	0.55	0.36	0.75	0.71	0.82	0.60
(0.5,1)	0.45	0.48	0.71	0.76	0.75	0.63	(0.5,1)	0.45	0.49	0.72	0.77	0.75	0.62
(0.75,1)	0.68	0.73	0.90	0.64	0.61	0.53	(0.75,1)	0.70	0.77	0.90	0.63	0.60	0.52
(0.4,0.6)	0.57	0.46	0.76	0.70	0.77	0.58	(0.4,0.6)	0.60	0.46	0.77	0.68	0.76	0.59
(0.1,0.9)	0.46	0.36	0.80	0.76	0.82	0.58	(0.1,0.9)	0.50	0.36	0.77	0.73	0.82	0.60
			(categories = 250)							*(categories = 500)*			
(0.1,1)	0.42	0.34	0.73	0.77	0.83	0.61	(0.1,1)	0.49	0.35	0.74	0.73	0.82	0.61
(0.5,1)	0.46	0.49	0.71	0.76	0.75	0.62	(0.5,1)	0.46	0.47	0.73	0.76	0.75	0.62
(0.75,1)	0.70	0.74	0.91	0.63	0.61	0.53	(0.75,1)	0.67	0.73	0.91	0.65	0.62	0.51
(0.4,0.6)	0.63	0.46	0.80	0.67	0.76	0.57	(0.4,0.6)	0.55	0.45	0.75	0.71	0.76	0.59
(0.1,0.9)	0.53	0.35	0.74	0.71	0.82	0.61	(0.1,0.9)	0.50	0.35	0.78	0.74	0.82	0.60

p=21, n=75, lambda=0.7,0.4,,0.3,0,-0.3, or -0.4; psi=various; 100 samples per DGP

Table A14

3.2 Monte Carlo Parameters

The following matrices were used to generate simulated data in the Monte Carlo exercises.

3.2.1 One-Factor Models

$$\Lambda_{simple} = \begin{bmatrix} 0.7 \\ 0.7 \\ 0.7 \\ 0.7 \\ 0.7 \\ 0.7 \\ 0.7 \end{bmatrix}$$

3.2.2 Two-Factor Models

$$\Lambda_{simple} = \begin{bmatrix} 0.7 & 0 \\ 0.7 & 0 \\ 0.7 & 0 \\ 0.7 & 0 \\ 0.7 & 0 \\ 0.7 & 0 \\ 0.7 & 0 \\ 0 & 0.7 \\ 0 & 0.7 \\ 0 & 0.7 \\ 0 & 0.7 \\ 0 & 0.7 \\ 0 & 0.7 \\ 0 & 0.7 \end{bmatrix}, \Lambda_{complex} = \begin{bmatrix} 0.7 & 0.4 \\ 0.7 & 0.4 \\ 0.7 & 0.4 \\ 0.7 & 0.4 \\ 0.7 & 0.4 \\ 0.7 & 0.4 \\ 0.7 & 0.4 \\ 0.4 & 0.7 \\ 0.4 & 0.7 \\ 0.4 & 0.7 \\ 0.4 & 0.7 \\ 0.4 & 0.7 \\ 0.4 & 0.7 \\ 0.4 & 0.7 \end{bmatrix}, \Lambda_{comp.,neg.} = \begin{bmatrix} 0.7 & -0.3 \\ 0.7 & -0.3 \\ 0.7 & -0.3 \\ 0.7 & -0.3 \\ 0.7 & -0.3 \\ 0.7 & 0.4 \\ 0.7 & 0.4 \\ 0.7 & 0.4 \\ 0.7 & 0.4 \\ 0.7 & 0.4 \\ 0.7 & 0 \\ 0.7 & 0 \\ 0.7 & 0 \\ 0.7 & 0 \end{bmatrix}$$

3.2.3 Three-Factor Models

$$\Lambda_{simple} = \begin{bmatrix} 0.7 & 0 & 0 \\ 0.7 & 0 & 0 \\ 0.7 & 0 & 0 \\ 0.7 & 0 & 0 \\ 0.7 & 0 & 0 \\ 0.7 & 0 & 0 \\ 0.7 & 0 & 0 \\ 0 & 0.7 & 0 \\ 0 & 0.7 & 0 \\ 0 & 0.7 & 0 \\ 0 & 0.7 & 0 \\ 0 & 0.7 & 0 \\ 0 & 0.7 & 0 \\ 0 & 0.7 & 0 \\ 0 & 0 & 0.7 \\ 0 & 0 & 0.7 \\ 0 & 0 & 0.7 \\ 0 & 0 & 0.7 \\ 0 & 0 & 0.7 \\ 0 & 0 & 0.7 \\ 0 & 0 & 0.7 \end{bmatrix}, \Lambda_{complex} = \begin{bmatrix} 0.7 & 0.4 & 0.3 \\ 0.7 & 0.4 & 0.3 \\ 0.7 & 0.4 & 0.3 \\ 0.7 & 0.4 & 0.3 \\ 0.7 & 0.4 & 0.3 \\ 0.7 & 0.4 & 0.3 \\ 0.7 & 0.4 & 0.3 \\ 0.3 & 0.7 & 0.4 \\ 0.3 & 0.7 & 0.4 \\ 0.3 & 0.7 & 0.4 \\ 0.3 & 0.7 & 0.4 \\ 0.3 & 0.7 & 0.4 \\ 0.3 & 0.7 & 0.4 \\ 0.3 & 0.7 & 0.4 \\ 0.4 & 0.3 & 0.7 \\ 0.4 & 0.3 & 0.7 \\ 0.4 & 0.3 & 0.7 \\ 0.4 & 0.3 & 0.7 \\ 0.4 & 0.3 & 0.7 \\ 0.4 & 0.3 & 0.7 \\ 0.4 & 0.3 & 0.7 \end{bmatrix}, \Lambda_{comp.,neg.} = \begin{bmatrix} 0.7 & 0 & -0.3 \\ 0.7 & 0 & -0.3 \\ 0.7 & 0 & -0.3 \\ 0.7 & 0 & -0.3 \\ 0.7 & 0 & -0.3 \\ 0.7 & 0 & -0.3 \\ 0.7 & 0 & -0.3 \\ 0 & 0.7 & 0.4 \\ 0 & 0.7 & 0.4 \\ 0 & 0.7 & 0.4 \\ 0 & 0.7 & 0.4 \\ 0 & 0.7 & 0.4 \\ 0 & 0.7 & 0.4 \\ 0 & 0.7 & 0.4 \\ -0.4 & 0.3 & 0.7 \\ -0.4 & 0.3 & 0.7 \\ -0.4 & 0.3 & 0.7 \\ -0.4 & 0.3 & 0.7 \\ -0.4 & 0.3 & 0.7 \\ -0.4 & 0.3 & 0.7 \\ -0.4 & 0.3 & 0.7 \end{bmatrix}$$

3.2.4 Five-Factor Models

$$\Lambda_{complex} = \begin{bmatrix}
0.9 & 0.1 & 0.2 & 0.2 & 0.2 \\
0.9 & 0.1 & 0.2 & 0.2 & 0.2 \\
0.9 & 0.1 & 0.2 & 0.2 & 0.2 \\
0.9 & 0.1 & 0.2 & 0.2 & 0.2 \\
0.9 & 0.1 & 0.2 & 0.2 & 0.2 \\
0.9 & 0.1 & 0.2 & 0.2 & 0.2 \\
0.9 & 0.1 & 0.2 & 0.2 & 0.2 \\
0.9 & 0.1 & 0.3 & 0.2 & 0.1 \\
0.9 & 0.1 & 0.3 & 0.2 & 0.1 \\
0.5 & 0.8 & 0.2 & 0.2 & 0.1 \\
0.5 & 0.8 & 0.2 & 0.2 & 0.1 \\
0.5 & 0.8 & 0.2 & 0.2 & 0.1 \\
0.5 & 0.8 & 0.2 & 0.2 & 0.1 \\
0.5 & 0.8 & 0.2 & 0.2 & 0.1 \\
0.5 & 0.8 & 0.2 & 0.2 & 0.1 \\
0.5 & 0.8 & 0.2 & 0.2 & 0.1 \\
0.5 & 0.8 & 0.2 & 0.2 & 0.1 \\
0.5 & 0.8 & 0.2 & 0.2 & 0.1 \\
0.5 & 0.3 & 0.7 & 0.2 & 0.2 \\
0.5 & 0.3 & 0.7 & 0.2 & 0.2 \\
0.5 & 0.3 & 0.7 & 0.2 & 0.2 \\
0.5 & 0.3 & 0.7 & 0.2 & 0.2 \\
0.5 & 0.3 & 0.7 & 0.2 & 0.2 \\
0.5 & 0.3 & 0.7 & 0.2 & 0.2 \\
0.5 & 0.3 & 0.7 & 0.2 & 0.2 \\
0.5 & 0.3 & 0.7 & 0.2 & 0.2 \\
0.5 & 0.3 & 0.7 & 0.2 & 0.2 \\
0.5 & 0.3 & 0.3 & 0.6 & 0.2 \\
0.5 & 0.3 & 0.3 & 0.6 & 0.1 \\
0.5 & 0.3 & 0.3 & 0.6 & 0.1 \\
0.5 & 0.3 & 0.3 & 0.6 & 0.1 \\
0.5 & 0.3 & 0.3 & 0.6 & 0.1 \\
0.5 & 0.3 & 0.3 & 0.6 & 0.1 \\
0.5 & 0.3 & 0.3 & 0.6 & 0.1 \\
0.5 & 0.3 & 0.3 & 0.6 & 0.1 \\
0.3 & 0.3 & 0.2 & 0.6 & 0.1 \\
0.3 & 0.3 & 0.2 & 0.2 & 0.5 \\
0.3 & 0.3 & 0.2 & 0.2 & 0.5 \\
0.3 & 0.3 & 0.2 & 0.2 & 0.5 \\
0.3 & 0.3 & 0.2 & 0.2 & 0.5 \\
0.3 & 0.3 & 0.2 & 0.2 & 0.5 \\
0.3 & 0.3 & 0.2 & 0.2 & 0.5 \\
0.3 & 0.3 & 0.2 & 0.2 & 0.5 \\
0.3 & 0.3 & 0.2 & 0.2 & 0.5 \\
0.3 & 0.3 & 0.2 & 0.2 & 0.5
\end{bmatrix}$$

Bibliography

Acemoglu, D., S. Johnson and J. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review.* Vol. 1(5)

Acemoglu, D., S. Johnson and J. Robinson. 2002. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics.* 1231-1294,

Acemoglu, D., S. Johnson and J. Robinson. 2004. "Institutions as the Fundamental Cause of Long-Run Growth." prepared chapter for *Handbook of Economic Growth.* Aghion and Durlauf, eds.

Acemoglu, D., and S. Johnson. 2005. "Unbundling Institutions." *Journal of Political Economy.* 113:5.

Adcock, R. and D. Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *The American Political Science Review.* 95:3, pp. 529-546.

Akaike, H. 1987. "Factor Analysis and AIC." *Psychometrika.* 52:3, pp. 317-332.

Anderson, J.C. and D. W. Gerbing. 1984. "The Effect of Sampling Error on Convergence, Improper Solutions, and Goodness-of-Fit Indices for Maximum Likelihood Confirmatory Factor Analysis." *Psychometrika.* 49:2, pp. 155-173.

Anderson, T.W. 2003. *An Introduction to Multivariate Statistical Analysis, Third ed.* Hoboken: John Wiley and Sons.

Anderson, T.W. and H. Rubin. 1956. "Statistical Inference in Factor Analysis." *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability.*

Armstrong, J. Scott. - 1967. "Derivation of Theory by Means of Factor Analysis or Tom Swift and His Electric Factor Analysis Machine." *The American Statistician.*

Arndt, C. and C. Oman. 2006. "Uses and Abuses of Governance Indicators." *OECD Development Center Studies.*

Bai, J. and S. Ng. - 2002. "Determining the Number of Factors in Approximate Factor Models." *Econometrica.* 70:1, pp. 191-221.

Bartholomew, D.J. 1980. "Factor Analysis for Categorical Data." *Journal of the Royal Statistical Society. Series B.* 42:3.

Bartlett, M.S. 1950. "Tests of Significance in Factor Analysis." *British Journal of Psychology, Statistics Section.* 3, pp. 77-85.

Beck, T. and L. Laeven. 2005. "Institution Building and Growth in Transition Economies." World Bank WPS3657.

Bernaards, C. and R. I. Jennrich. 2003. "Orthomax Rotation and Perfect Simple Structure." *Psychometrika.* 68:4, pp. 585-588.

Bernanke, B., J. Boivin. 2003. "Monetary Policy in a Data-Rich Environment." *Journal of Monetary Economics.* 50:3, pp. 525-546.

Boivin, J. and S. Ng. 2006. "Are more data always better for factor analysis?" *Journal of Econometrics.* 132, pp. 169-194.

Bound, J., D. A. Jaeger and R. M. Baker. 1995. "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak." *Journal of the American Statistical Association.* 90:430, pp. 443-450.

Brown, M. W. 1969. "Fitting the Factor Analysis Model." *Psychometrika.* 34:3, pp. 375-394.

Brown, M. W. 2001. "An Overview of Analytic Rotations in Exploratory Factor Analysis." *Multivariate Behavioral Research.* 36:1, pp. 111-150.

Cattell, R. B. 1966. "The Scree Test for the Number of Factors." *Multivariate Behavioral Research.* 1, pp. 245-276.

Cattell, R. B. and S. Vogelmann. 1977. "A Comprehensive Trial of the Scree and KG Criteria for Determining the Number of Factors." *Multivariate Behavioral Research.* 12, pp. 289-325.

Chamberlain, G. and M. Rothschild. 1983. "Factor Structure, and Mean-Variance Analysis on Large Asset Markets." *Econometrica.* 51:5, pp. 1281-1304.

Chao, J. C. and N. R. Swanson. 2005. "Consistent Estimation with a Large Number of Weak Instruments." *Econometrica.* 73:5, pp. 1673-1692.

Chen, D.H.C. and K. Gawande. 2007. "Underlying Dimensions of Knowledge Assessment: Factor Analysis of the Knowledge Assessment Methodology Data." World Bank WPS 4216.

Connor, G. and R. A. Korajczyk. 1993. "A Test for the Number of Factors in an Approximate Factor Model." *The Journal of Finance.* 48:4.

Cragg, J.G. and S. G. Donald. 1993. "Testing Identifiability and Specification in Instrumental Variable Models." *Econometric Theory.* 9:2.

Cragg, J.G. and S. G. Donald. 1997. "Inferring the Rank of a Matrix." *Journal of Econometrics.* 76, pp. 223-250.

Cristadoro, R., M. Forni, L. Reichlin and G. Veronese. 2005. "A Core Inflation Indicator for the Euro Area." *Journal of Money, Credit and Banking.* 37:3, pp. 539-560.

Durham, J. 2004. "Economic Growth and Institutions: Some Sensitivity Analyses, 1961-2000." *International Organization.* 58, pp. 485-529.

DeSoto, H. 2000. *The Mystery of Capital.* Basic Books.

Dillon, W.R., A. Kumar and N. Mulani. 1987. "Offending Estimates in Covariance Structure Analysis: Comments on the Causes of and Solutions to Heywood Cases." *Psychological Bulletin.* 101:1, pp. 126-135.

Dolan, C. V. 1994. "Factor analysis of variables with 2,3,5, and 7 response categories: A comparison of categorical variable estimators using simulated data." *British Journal of Mathematical and Statistical Psychology.* pp. 309-326.

Dollar, D. and A. Kraay. 2003. "Institutions, Trade and Growth: Revisiting the Evidence." World Bank Policy Research Working Paper 3004.

Easterly, W. and R. Levine. 2003. "Tropics, germs and crops: how endowments influence economic development." *Journal of Monetary Economics.* 50, pp. 3-39.

Flora, D.B. 2002. "Evaluation of Categorical Variable Methodology for Confirmatory Factor Analysis with Likert-Type Data." Doctoral Dissertation, University of North Carolina, Chapel Hill.

Gallup, J.L., J. Sachs and A.D. Mellinger. 1999. "Geography and International Development." Harvard CID Working Paper No. 1.

Gerbing D. W. and J. C. Anderson. 1987. "Improper Solutions in the Analysis of Covariance Structures: Their Interpretability and a Comparison of Alternate Respecifications." *Psychometrika.* 52:1, pp. 99-111.

Geweke, J.F. and K.J. Singleton. 1980. "Interpreting the likelihood ratio test statistic in factor models when sample size is small." *Journal of the American Statistical Association.* 75, pp. 133-137.

Glaeser E.L., R. La Porta, F. Lopez-de-Silanes. and A. Shleifer. 2004. "Do Institutions Cause Growth?" *Journal of Economic Growth.* 9:3.

Greene, W. 2003. *Econometric Analysis.* 5th ed. Upper Saddle River, NJ: Prentice Hall.

Greif, A. 1993. "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition." *American Economic Review.* 83:3.

Hair, J. F., B. Black, B. Babin, R. E. Anderson, and R. Tatham. 2005. *Multivariate Data Analysis*, 6th ed. Prentice Hall.

Hakstian, A. R., W.T. Rogers and R.B. Cattell. 1982. "The Behavior of Numbers-of-Factors Rules with Simulated Data." *Multivariate Behavioral Research.* 17, pp. 193-219.

Hall, R. and C. Jones. 1999. "Why Do Some Countries Produce so Much More Output per Worker than Others?" *Quarterly Journal of Economics.* 114:1, pp. 83-116.

Hayashi, K. and G.A. Marcoulides. 2006. "Examining Identification Issues in Factor Analysis." *Structural Equation Modeling.* 13:4, pp. 631-645.

Hayton, J. C., D. G. Allen and V. Scarpello. 2005. "Factor Retention Decisions in Exploratory Factor Analysis: A Tutorial on Parallel Analysis." *Organizational Research Methods.* 7, pp. 191-205.

Hendley, K., P. Murrell and R. Ryterman. 2000. "Law, Relationships and Private Enforcement: Transactional Strategies of Russian Enterprises." *Europe-Asia Studies.*

Heston, A. , R. Summers and B. Aten. 2009. Penn World Table Version 6.3, Center for International Comparisons of Production, Income and Prices. University of Pennsylvania.

Hogg, R. and A. Craig. 1995. *Introduction to Mathematical Statistics.* 5th ed. Upper Saddle River, NJ: Pearson Education.

Horn, J. L. 1965. "A Rationale and Test for the Number of Factors in Factor Analysis." *Psychometrika.* 30:2, pp. 179-185.

Humphreys, L. G. and R. G. Montanellis, Jr. 1975. "An Investigation of the Parallel Analysis Criterion for Determining the Number of Common Factors." *Multivariate Behavioral Research.* 10:2, pp. 193-205.

Jennrich, R. I. 2004. "Rotation to Simple Loadings Using Component Loss Functions: The Orthogonal Case." *Psychometrika.* 69:2, pp. 257-273.

Johnson, S, J. McMillan and C. Woodruff. 1999. "Contract Enforcement in Transition." CESifo Working Paper No. 211.

Johnson, S., J. McMillan and C. Woodruff. 2002. "The central role of entrepreneurs in transition economies." *Journal of Economic Perspectives.* 16:3, pp. 153-170.

Kaiser, H. F. 1960. "The Application of Electronic Computers to Factor Analysis." *Educational and Psychological Measurement.* 20, pp. 141-151.

Kamakura, W.A., and M. Wedel. 2001. "Exploratory Tobit Factor Analysis for Multivariate Censored Data." *Multivariate Behavioral Research.* 36:1.

Karvonen, M. J., and M. Niemi. 1953. "Factor analysis of performance in track and field events." *Europen Journal of Applied Physiology and Occupational Physiology.* 15:2.

Kaufmann, D., A. Kraay and P. Zoido-Lobaton. 1999a. "Governance Matters." World Bank Policy Research Working Paper, No. 2196. World Bank Institute.

Kaufmann, D., A. Kraay and P. Zoido-Lobaton. 1999b. "Aggregating Governance Indicators." World Bank Policy Research Working Paper, No. 2195. World Bank.

Kaufmann, D. and A. Kraay. 2004. "Governance Matters III: New Indicators for 1996-2002 and Addressing Methodological Challenges." World Bank Policy Research Working Paper, No. 3106. World Bank.

Kaufmann, D., A. Kraay and M. Mastruzzi. 2006a. "Governance Matters V: Aggregate and Individual Governance Indicators for 1996-2005." World Bank.

Kaufmann, D., A. Kraay and M. Mastruzzi. 2006b. "Governance Matters V: Appendices" World Bank.

Kaufmann, D., and A. Kraay. 2007. Worldwide Governance Indicators dataset*: <http://info.worldbank.org/governance/wgi/index.asp>*. Data used in this paper were downloaded between Dec, 2006-Jan,2007.

Kaufmann, D., and A. Kraay. 2007. "Governance Indicators: Where are we? Where should we be going?" World Bank Policy Research Working Paper 4370. World Bank.

Kaufmann, D., A. Kraay and M. Mastruzzi. 2007. "The Worldwide Governance Indicators Project: Answering the Critics." World Bank.

Kendall, M.G., and D.N. Lawley. 1956. "The Principles of Factor Analysis." *Journal of the Royal Statistical Society, Series A (General).* 119:1.

Kennedy, P. 2003. *A Guide to Econometrics.* 5th ed. Cambridge: MIT Press.

Kim,K. H., and P. M. Bentler. 2002. "Tests of homogeneity of means and covariance matrices for multivariate incomplete data." *Psychometrika.* 67:4.

Kleibergen, F. 2004. "Testing Subsets of Structural Parameters in the Instrumental Variables Regression Model." *The Review of Economics and Statistics.* 86:1.

Knack, S. and P. Keefer. 1995. "Institutions and Economic Performance: Cross-Country Tests using Alternative Institutional Measures." *Economics and Politics.* 7:3.

Knack, S. and P. Keefer. 1997. "Does Social Capital Have an Economic Payoff: A cross-country Invetigation." *Quarterly Journal of Economics.* 112:4.

Knack, S. 2006. "Measuring Corruption in Eastern Europe and Central Asia: A Critique of the Cross-Country Indicators." World Bank Policy Research Department, Working Paper 3968.

Kurtz, M. and A. Schrank. 2007. "Growth and Governance: Models, Measures, and Mechanisms." *Journal of Politics.*

La Porta, R., F. Lopez-de-Silanes, A. Shleifer, R. Vishny. 1999. "The Quality of Government." *Journal of Law, Economics & Organization.*" 15:1.

Lawley, D.N. and A. E. Maxwell. 1971. - *Factor Analysis as a Statistical Method*. London: Butterworth.

Linden, M. 1977. "Factor Analytical Study of Olympic Decathlon Data." *Research Quarterly.* 48:3.

Little, R.J.A. and D. B. Rubin. 2002. *Statistical Analysis with Missing Data*, 2nd ed. Hoboken: John Wiley and Sons.

Mauro, P. 1995. "Corruption and Growth." *Quarterly Journal of Economics.* 110:3, pp. 681-712.

McDonald, R.P. 1970. "The Theoretical Foundations of Principal Factor Analysis, Factor Analysis and Alpha Factor Analysis." *British Journal of Mathematical and Statistical Psychology.* 23, pp. 1-21.

Mikusheva, A. and B. Poi. 2002. "Tests and confidence sets with correct size in the simultaneous equations model with potentially weak instruments." *The Stata Journal.*

Mislevy, R. J. 1986. "Recent Developments in the Factor Analysis of Categorical Variables." *Journal of Educational Statistics.* 11:1.

Moreira, M.J. 2003. "A conditional likelihood ratio test for structural models." *Econometrica.* 71:4.

Muthen, B.O. 1989. "Tobit factor analysis." *British Journal of Mathematical and Statistical Psychology.* 42, pp. 241-250.

Nelson, C. R., and R. Startz. 1990. "Some Further Results on the Exact Small Sample Properties of the Instrumental Variable Estimator." *Econometrica.* 58:4, pp. 967-976.

North, D. C. 1990. *Institutions, Institutional Change and Economic Performance*, Cambridge, UK: Cambridge University Press.

Presbitero, A.F. 2006. - "Institutions and Geography as Sources of Economic Development." *Journal of International Development.* 18:3.

Rodrik, D., A. Subramanian and F. Trebbi. 2004. - "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development." *Journal of Economic Growth.* 9, pp. 131-165.

Roll, R. and S. A. Ross. 1980. "An Empirical Investigation of the Arbitrage Pricing Theory." *The Journal of Finance.* 35:5.

Sachs, J. .2003."Institutions Don't Rule: Direct effects of Geography on Per Capita Income." NBER Working Paper No. 9490.

Schwarz, G. 1978. "Estimating The Dimension of a Model." *Annals of Statistics.* 6:2, pp. 461-464.

Shea, J. 1997. "Instrument Relevance in Multivariate Linear Models: A Simple Measure." *The Review of Economics and Statistics.* 79:2.

Shi, J and S. Lee. 1997. "A Bayesian Estimation of Factor Score in Confirmatory Factor Model with Polytomous, Censored or Truncated Data." *Psychometrika.* 62:1.

Staiger, D. and J. H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica.* 65:3.

Stewart, D. W. 1981. "The Application and Misapplication of Factor Analysis in Marketing Research." *Journal of Marketing Research.* 18:1, pp. 51-62.

Thompson, J. 2004. "A Monte Carlo Comparison of Tests for the Number of Factors under Alternative Factor Models." Unpublished doctoral dissertation. University of California, Davis.

Thomson, G. H. 1951. *The Factorial Analysis of Human Ability* (5th Ed.). London: University of London.

Schafer, J. L. 1997. *Analysis of Incomplete Multivariate Data.* London: Chapman and Hall.

Stanghellini, E. 1997. "Identification of a single-factor model using graphical gaussian rules." *Biometrika.* 84:1.

Stock, J. H. and M. W. Watson. 1998. "Diffusion Indexes." NBER working paper number 6702.

Stock, J. H. and M. W. Watson. 2002. "Forecasting Using Principal Components from a Large Number of Predictors." *Journal of the American Statistical Association.* 97:460, pp. 1167-1179.

Stock, J. H. and M. Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression." in D.W.K. Andrews and J.H. Stock, eds., *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg.* Cambridge: Cambridge University Press, pp. 80–108.

Svensson, J. - 2005. "Eight Questions about Corruption." *Journal of Economic Perspectives.* 19:3.

Thomas, M. 2006. "What Do The Worldwide Governance Indicators Measure?" Manuscript, Johns Hopkins University.

Van Driel, O. P. 1978. "On Various Causes of Improper Solutions In Maximum Likelihood Factor Analysis." *Psychometrika.* 43:2, pp. 225-243.

Waller N.G. and B.O. Muthen. 1992. "Genetic Tobit Factor Analysis: Quantitative Genetic Modeling with Censored Data." *Behavior Genetics.* 22:3.

Wansbeek, T. and E. Meijer. 2000. - *Measurement Error and Latent Variables in Econometrics.* Amsterdam: North-Holland.

Weingast, B.P. 1995. "The Economic Role of Political Institutions: Market-Preserving Federalism and Economic Development." *Journal of Law, Economics and Organization.* 11:1.

Wooldridge, J. W. 2002. *Econometric Analysis of Cross Section and Panel Data.* Cambridge: MIT Press.

Zwick, W. R. and W.F. Velicer. 1982. "Factors Influencing Four Rules for Determining the Number of Components to Retain." *Multivariate Behavioral Research.* 17, pp. 253-269.

Zwick, W. R. and W.F. Velicer. 1986. "Comparison of Five Rules for Determining the Number of Components to Retain." *Psychological Bulletin.* 99:3, pp.

432-442.