

Article

Not peer-reviewed version

Instrument Selection in Panel Data Models with Endogeneity: A Bayesian Approach

[Álvaro Herce Postigo](#) * and [Manuel Salvador](#)

Posted Date: 23 September 2024

doi: 10.20944/preprints202409.1805.v1

Keywords: endogeneity; bayesian methods; MCMC; instrumental variables selection; international bank flows



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Instrument Selection in Panel Data Models with Endogeneity: A Bayesian Approach

Herce Álvaro * and Salvador Manuel

Department of Applied Economy, Faculty of Business and Economics, University of Zaragoza

* Correspondence: 577134@unizar.es

Abstract: This paper proposes the use of Bayesian inference techniques to search for and obtain valid instruments in dynamic panel data models where endogenous variables may exist. The use of Principal Component Analysis (PCA) allows for obtaining a reduced number of instruments in comparison to the high number of instruments commonly used in the literature, and Monte Carlo Markov Chain (MCMC) methods enable efficient exploration of the instrument space, deriving accurate point estimates of the elements of interest. The proposed methodology is illustrated in a simulated case and in an empirical application, where the partial effect of a series of determinants on the attraction of international bank flows is quantified. The results highlight the importance of promoting and developing the private sector in these economies, as well as the importance of maintaining good levels of credit worthiness

Keywords: endogeneity; bayesian methods; MCMC; instrumental variables selection; international bank flows

1. Introduction

Endogeneity is a common issue in economic research that can lead to biased and inconsistent estimates if not properly addressed. Instrumental variables are typically employed to mitigate this problem. These variables must satisfy two key conditions: they must be uncorrelated with the error term of the regression model (exogeneity condition) and correlated with the endogenous variables (relevance condition). By using these instruments as regressors instead of the endogenous variables, it is possible to obtain asymptotically unbiased estimates of the regression coefficients.

In panel data modelling, estimation techniques based on the Generalized Method of Moments (GMM) are commonly used (Arellano & Bond, 1991; Arellano & Bover, 1995). To apply GMM, a system of equations is specified. The first equation captures the relationship between the dependent variable and a set of explanatory variables, which may include both strictly exogenous and endogenous variables. The remaining equations describe the interdependencies between the endogenous variables and a set of regressors or instrumental variables. These instruments are intended to maximize the information about the endogenous variables (relevance condition) while ensuring they do not introduce bias into the estimation of the first equation due to their orthogonality with the error term (exogeneity condition).

Under certain regularity conditions, the GMM estimation procedure is asymptotically unbiased and efficient (Roodman, 2009b). However, this method faces several limitations. One significant challenge is the high dimensionality of the instrument set. As GMM methodology constructs instruments based on orthogonality conditions between the idiosyncratic error and the instruments, the number of instruments increases quadratically with the length of the time series, leading to a trade-off between efficiency gains from incorporating additional moments and the loss of degrees of freedom. Moreover, while the orthogonality conditions generally ensure the exogeneity of instruments, they do not guarantee the relevance condition, which is a critical issue addressed in this study.

To reduce the number of instruments, Roodman (2009 a, b) proposed collapsing the instrument matrix. While effective in many cases, this approach can still result in high-dimensional matrices, particularly in contexts with a large number of time observations, and it does not differentiate between relevant and irrelevant instruments. Nevertheless, the approach of constructing instrumental matrices using lags of the endogenous variables, combined with dimensional reduction through matrix collapsing, provides a useful starting point for applying GMM in panel models, including non-dynamic ones, where suitable instruments may be scarce.

In this paper, we propose the use of Principal Component Analysis (PCA) combined with Bayesian methods to select a reduced number of valid instruments without sacrificing precision in estimating the parameters of interest. First, we perform PCA on the instrument matrix to address multicollinearity among the instruments. Next, we model the endogenous variable with respect to each transformed instrument, ensuring orthogonality. We estimate each relationship using Gibbs Sampling with a non-informative prior on the parameters of interest. This approach allows us to identify relevant instruments that are significant at a minimum credibility level of 90%. The application of Bayesian techniques facilitates the selection of a parsimonious set of instruments while maintaining the goodness of fit.

The remainder of the paper is organized as follows. Section 2 introduces the problem, assuming the presence of an endogenous regressor. Section 3 outlines the process of identifying valid instruments that satisfy the relevance and exogeneity conditions while minimizing their number. Section 4 details the posterior inference procedure for the model parameters following the identification of the necessary instruments. In Section 5, we evaluate the performance of the proposed methodology using a simulated example. Section 6 presents the empirical study, where we analyse the partial effects of governance indicators on bank capital inflows. Section 7 concludes the paper. Supplementary material includes three appendices: Appendix A provides a detailed mathematical derivation of key results, Appendix B offers a description of the variables, and Appendix C presents additional supporting information.

2. Setting-Up the Problem

Let $\{y_{it}, \{x_{jit}, j = 1, 2, \dots, K_x\}, \{w_{jit}, j = 1, 2, \dots, K_w\}; i = 1, \dots, N; t = 1, \dots, T\}$ be a balanced panel dataset corresponding to N individual units and T periods, where y is the dependent variable, $X = (X_1, \dots, X_{K_x})$ are explanatory strictly exogenous variables and $W = (W_1, \dots, W_{K_w})$ are endogenous explanatory variables. Without loss of generality, we assume that $K_w = 1$, with an extension to the case of $K_w > 1$ being straightforward (see section 4.4. below).

Let us assume a dynamic panel model with individual effects given by:

$$y_{it} = \beta_0 y_{i,t-1} + \sum_{j=1}^{K_x} x_{jit} \beta_{1j} + w_{it} \beta_2 + \alpha_i + u_{it} \quad (2.1)$$

which describes the evolution of the variable of interest y , where $\{\alpha_i, i = 1, 2, \dots, N\}$ are individual fixed effects and $\{u_{it}; i = 1, \dots, N; t = 1, \dots, T\}$ are i.i.d. homoscedastic random normal disturbances, with $u_{it} \sim N(0, \sigma_u^2)$. The main concern is the estimation of β_2 , the direct effect of W on y .

To eliminate bias in the estimation of β_2 , we consider a set of K_z instrumental variables $Z = (Z_1, \dots, Z_{K_z})$ with observed values $\{z_{jit}, j = 1, 2, \dots, K_z; t = 1, \dots, T; i = 1, \dots, N\}$, which are assumed to be related to the variable W (relevance condition) by means of the linear regression:

$$w_{it} = \sum_{j=1}^{K_z} z_{jit} \phi_j + \epsilon_{it} \quad (2.2)$$

with $\{\epsilon_{it} \sim N(0, \sigma_\epsilon^2); t=1, \dots, T; i=1, \dots, N\}$. The instrumental variables Z are supposed to be uncorrelated with the random disturbances $\{u_{it}; t=1, \dots, T; i=1, \dots, N\}$ in such a way that $\text{cov}(z_{jit}, u_{it}) = 0, \forall i, t$ (exogeneity condition), and it is possible to consider including those variables that are marked as strictly exogenous in the model. In the best-case scenario, the use of instrumental variables that meet the aforementioned conditions and are available to the researcher is considered. However, in reality, these variables are often unknown and not externally accessible, making it necessary to use the available sample information to generate them.

In addition, it is assumed that:

$$\begin{aligned} \text{cov}(u_{it}, x_{jit}) &= \text{cov}(\epsilon_{it}, z_{jit}) = \text{cov}(\epsilon_{it}, x_{jit}) = 0 \quad \forall i, j, t \\ \text{cov}((u_{it}, \epsilon_{it}), (u_{jt}, \epsilon_{jt})) &= 0 \quad \forall i \neq j \\ (u_{it}, \epsilon_{it}) &\sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \sigma_{u,\epsilon} \\ \sigma_{u,\epsilon} & \sigma_\epsilon^2 \end{bmatrix}\right) \end{aligned}$$

where $\sigma_{u,\epsilon} = \text{cov}(u_{it}, \epsilon_{it})$. The endogeneity problem arises when $\sigma_{u,\epsilon} = \text{cov}(u_{it}, \epsilon_{it}) \neq 0$, which lead to bias in the inference about the parameter β_2 , which does not disappear asymptotically. This bias can be overcome using instrumental variables that verify the conditions of relevance ($\exists j \in \{1, \dots, K_z\} : \phi_j \neq 0$), and exogeneity ($\text{cov}(z_{jit}, u_{it}) = 0; \forall j$). However, it is usual to choose the instrumental variables Z using doubtful ad hoc subjective value judgements. Commonly used procedures, even those proposed by (Roodman, 2009 a, b), can generate a large number of instruments, increasing the computational complexity of the procedure, as well as the existence of multicollinearity problems between the instruments, reducing the precision of the inferences.

In this paper, we propose a Bayesian methodology to estimate β_2 providing, on one hand, a general methodology to obtain valid instruments generated from sample information, using relevant moment conditions and, on the other hand, a procedure to reduce the size of the generated matrix of instruments by mitigating the likelihood of quasi-multicollinearity among instruments.

3. Methodology

3.1. General Setting and Forward Orthogonal Deviations

We present a general methodology for dealing with endogenous variables using only one endogenous regressor. Application to more than one endogenous regressor is straightforward, and general guidelines are provided below (see section 4.4). The starting point is the equation system defined by:

$$\begin{aligned} y_{it} &= \beta_0 y_{i,t-1} + \sum_{j=1}^{K_x} x_{jit} \beta_{1j} + w_{it} \beta_2 + \alpha_i + u_{it} \\ w_{it} &= \sum_{j=1}^{K_z} z_{jit} \phi_j + \epsilon_{it} \end{aligned} \quad (3.1)$$

In panel data settings, a source of potential endogeneity is the value of $\text{cov}(y_{i,t-1}, \alpha_i)$, but one can apply some of the common transformations to eliminate individual effects while retaining sample information. Our proposal is to employ a Forward Orthogonal Deviation Transformation (Arellano

& Bover, 1995) to guarantee homoscedasticity and uncorrelation of the regression errors after the application under the assumptions listed above. For each element included in the model, the transformation is given by:

$$\begin{aligned}
 \tilde{y}_{it} &= \left[\frac{T-t}{T-t+1} \right]^{\frac{1}{2}} \left[y_{it} - \frac{1}{T-t} \sum_{s=t+1}^T y_{is} \right] \\
 \tilde{x}_{jit} &= \left[\frac{T-t}{T-t+1} \right]^{\frac{1}{2}} \left[x_{jit} - \frac{1}{T-t} \sum_{s=t+1}^T x_{jis} \right] \\
 \tilde{z}_{jit} &= \left[\frac{T-t}{T-t+1} \right]^{\frac{1}{2}} \left[z_{jit} - \frac{1}{T-t} \sum_{s=t+1}^T z_{jis} \right] \\
 \tilde{w}_{it} &= \left[\frac{T-t}{T-t+1} \right]^{\frac{1}{2}} \left[w_{it} - \frac{1}{T-t} \sum_{s=t+1}^T w_{is} \right] \\
 \tilde{u}_{it} &= \left[\frac{T-t}{T-t+1} \right]^{\frac{1}{2}} \left[u_{it} - \frac{1}{T-t} \sum_{s=t+1}^T u_{is} \right] \\
 \tilde{\varepsilon}_{it} &= \left[\frac{T-t}{T-t+1} \right]^{\frac{1}{2}} \left[\varepsilon_{it} - \frac{1}{T-t} \sum_{s=t+1}^T \varepsilon_{is} \right]
 \end{aligned} \tag{3.2}$$

which, applied to the system (3.1) resulting in:

$$\begin{aligned}
 \tilde{y}_{it} &= \beta_0 \tilde{y}_{i,t-1} + \sum_{j=1}^{K_x} \tilde{x}_{jit} \beta_{1j} + \tilde{w}_{it} \beta_2 + \tilde{u}_{it} \\
 \tilde{w}_{it} &= \sum_{j=1}^{K_z} \tilde{z}_{jit} \phi_j + \tilde{\varepsilon}_{it}
 \end{aligned} \tag{3.3}$$

Thus, the fixed effects α_i are removed from (3.1). When dealing with an unbalanced panel, as pointed out by (Roodman, 2009b), the Forward Orthogonal Deviation Transformation fills the gaps with the average of all future available observations of a variable, minimizing data loss. Practically, this is equivalent to replace the missing value as a zero value.

3.2. Over-Dimensionality Problems

From a frequentist perspective, the Generalized Method of Moments (GMM) approach has usually been considered for the inference process because of its flexibility and because it requires adopting very few assumptions about the data-generating process. In addition, most proposed versions provide estimates that prevent the well-known (Nickell, 1981) bias from appearing. However, this method has some shortcomings. The main one is the way the instrumental variables are introduced. Using the approach of (Holtz-Eakin et al., 1988), the number of valid instruments is

$NT \times \frac{T(T-1)}{2}$, which can cause problems in the finite samples. One of problems is the dimension

of the resultant variance matrix of the moments, which is a quadratic function of T, as the number of instruments increases. This ultimately leads to the matrix of instruments becoming singular, and it is necessary to perform GMM. In a Bayesian framework, given that the inference is exact or based on Monte Carlo methods, this is not problematic. However, the second problem is that a large instrument collection can overfit the endogenous variables in equation (2.2). Unfortunately, Bayesian procedures can have a similar problem, but in a lesser way, because of their tendency to select parsimonious models.

A first proposal to establish a valid methodology to reduce the number of instruments was proposed by Roodman (2009 a, b), which is based on collapsing the resultant instrumental matrix. Although this construction implies a significant reduction in the number of instruments to be incorporated, this perspective still preserves some relevant aspects to be corrected. On the one hand, in panels where the size of period T is high, the loss of degrees of freedom, even in collapsed matrices, will be relevant in frequentist cases. From a Bayesian inference perspective, working with a collapsed matrix with a large T may mean incorporating correlated elements in the instrument matrix, which could decrease the precision of posterior inferences. However, and highly related to the above, there is a possibility of adding instruments that are not very relevant in the inference process and can cause potential multicollinearity problems. Thus, the optimal procedure would be to introduce only those instruments relevant to the explanation of endogenous variables. Our proposal allows us to reduce the number of instruments incorporated while guaranteeing the absence of multicollinearity among the valid instruments selected.

3.3. Instrumental Matrices

We start by establishing valid instruments keeping the required orthogonality condition invariant:

$$E\left[Z'_{ji}u_i\right]=0; i=1,\dots,N, j=1,2,\dots,K_z$$

where $u_i=(u_{i1},\dots,u_{iT})'$ and $Z_{ji}=(z_{ji1},\dots,z_{jiT})'$. In our case we take the collapsed form proposed by Roodman (2009 a, b), after applying Forward Orthogonal Deviations (3.2), and we create

the collapsed instrumental matrix for endogenous variable W as $\tilde{Z}=\begin{bmatrix} \tilde{Z}_1 \\ \vdots \\ \tilde{Z}_N \end{bmatrix}$ where

$$\tilde{Z}_i=\begin{bmatrix} \tilde{x}_{1i1} & \cdots & \tilde{x}_{K_x i1} & 0 & 0 & \cdots & 0 \\ \tilde{x}_{1i2} & \cdots & \tilde{x}_{K_x i2} & \tilde{w}_{i1} & 0 & \cdots & 0 \\ \tilde{x}_{1i3} & \cdots & \tilde{x}_{K_x i3} & \tilde{w}_{i2} & \tilde{w}_{i1} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \tilde{x}_{1iT} & \cdots & \tilde{x}_{K_x iT} & \tilde{w}_{iT-1} & \tilde{w}_{iT-2} & \cdots & \tilde{w}_{i1} \end{bmatrix}; i=1,\dots,N$$

Once the instrumental matrix $\tilde{Z}$ is established, our proposal consists of carrying out a principal component analysis on $\tilde{Z}$, to select the components with better explanatory power of. Let the matrix form of the endogenous component be defined as:

$$\tilde{w}=Z\phi+\tilde{\varepsilon}$$

where $\tilde{w}\in\Re^{NT}\times 1$, $\tilde{Z}\in\Re^{NT}\times\Re^{K_{\tilde{Z}}}$, $\phi\in\Re^{K_{\tilde{Z}}\times 1}$, $\tilde{\varepsilon}\in\Re^{NT}\times 1$ and $K_{\tilde{Z}}=K_x+T-1$. Note that in the general equation (3.3), the dimension of the instrumental variables would be now $K_Z=K_{\tilde{Z}}=K_x+T-1$, in accordance with the methodology.

Our proposal for selecting a reduced set of valid instruments for $\tilde{w}$ is based on the extraction of the principal components of $\tilde{Z}$ more relevant to explain $\tilde{w}$. Reformulating the instrumental equation in terms of the principal component scores of $\tilde{Z}$, $\{\tilde{s}_{1it},\dots,\tilde{s}_{K_{\tilde{Z}}it}\}$ for i,t we have that:

$$\tilde{w}_{it}=\tilde{s}_{1it}\delta_1+\tilde{s}_{2it}\delta_2+\dots+\tilde{s}_{K_{\tilde{Z}}it}\delta_{K_{\tilde{Z}}}+\tilde{v}_{it}$$

with $\tilde{v}_{it}\sim N(0,\sigma_{\tilde{v}}^2)\forall i,t$, or, in matrix terms:

$$\tilde{w} = \tilde{S}\delta + \tilde{v} \quad (3.4)$$

To search for valid instruments, we relied on Bayesian procedures. Specifically, we propose a search for relevant instruments using a non-informative Jeffrey's prior on the parameters of interest. Given the orthogonality of the principal components, a relationship can be established between the endogenous variable and a given instrument, allowing for an individualized analysis of its representativeness. Moreover, this process can be carried out in parallel with all available instruments to improve computational efficiency. The employment of this kind of prior try to reflect our prior ignorance about the parameters, being this distribution invariant to reparameterizations (Jeffreys, 1961).

3.4. Search for Valid Instruments

The likelihood function is based on the distribution:

$$\tilde{w} | \tilde{S}, \delta, \sigma_v^2 \sim N_{NT}(\tilde{S}\delta, \sigma_v^2 I_{NT}) \quad (3.5)$$

with a non-informative Jeffrey's prior given by:

$$\pi(\sigma_v^2, \delta) \propto \left(\frac{1}{\sigma_v^2} \right) I_{(0,\infty)}(\sigma_v^2) \quad (3.6)$$

The posterior kernel, that is, the posterior distribution up to a normalizing constant, is given then by:

$$p(\delta, \sigma_v^2 | \tilde{w}) \propto (\sigma_v^2)^{-\left(\frac{NT}{2}+1\right)} \exp \left\{ -\frac{1}{2\sigma_v^2} (\tilde{w} - \tilde{S}\delta)' (\tilde{w} - \tilde{S}\delta) \right\} I_{(0,\infty)}(\sigma_v^2) \quad (3.7)$$

This posterior kernel leads to the following posterior conditional distributions:

$$\begin{aligned} \sigma_v^2 | \delta, \tilde{w} &\sim IG(a_1, b_1); a_1 = \frac{NT}{2}, b_1 = \frac{(\tilde{w} - \tilde{S}\delta)' (\tilde{w} - \tilde{S}\delta)}{2} \\ \delta | \sigma_v^2, \tilde{w} &\sim N_{K_Z}(\hat{\delta}, V_1); \hat{\delta} = (\tilde{S}'\tilde{S})^{-1} \tilde{S}'\tilde{w}, V_1 = \sigma_v^2 (\tilde{S}'\tilde{S})^{-1} \end{aligned} \quad (3.8)$$

We use these posterior conditional distributions in a Gibbs Sampling scheme (Geman & Geman, 1984) to obtain point estimates and credible intervals for each instruments. After an initial burning period, we discarded the initial part of the simulated chain, and we took the remaining iterations as an approximate dependent sample from this distribution. This sample was used to make posterior inferences regarding the model parameters. We select those instruments that are significant at a credibility level of at least 90%. The details about the mathematical derivation of the full conditional distribution employed are shown in Appendix A of the Supplementary Material.

4. Estimation of the Equation System

Once we have selected the instruments, we infer the parameters of Model (3.3). Again, we used the Bayesian approach. In this case, the likelihood function and prior distribution of parameters are described as in (Greenberg, 2012).

4.1. Likelihood Function

The equation system will be given by:

$$\begin{aligned} \tilde{y}_{NT \times 1} &= \tilde{y}_{-1} \beta_0 + \tilde{X}_{NT \times K_x} \beta_1 + \tilde{w}_{NT \times 1} \beta_2 + \frac{\sigma_{\tilde{u}, \tilde{v}}}{\sigma_{\tilde{v}}^2} \tilde{v} + \tilde{u} - \frac{\sigma_{\tilde{u}, \tilde{v}}}{\sigma_{\tilde{v}}^2} \tilde{v} \\ \tilde{w}_{NT \times 1} &= \tilde{S}_{NT \times K_{\tilde{z}_w}} \delta + \tilde{v}_{\tilde{z}_w \times 1} \end{aligned}$$

i.e.,

$$\begin{aligned} \tilde{y} &= \tilde{y}_{-1} \beta_0 + \tilde{X} \beta_1 + \tilde{w} \beta_2 + \tilde{v} \beta_3 + \tilde{u} \\ \tilde{w} &= \tilde{S} \delta + \tilde{v} \end{aligned} \quad (4.1)$$

where $\beta_3 = \frac{\sigma_{\tilde{u}, \tilde{v}}}{\sigma_{\tilde{v}}^2}$, $\hat{u} = \tilde{u} - \frac{\sigma_{\tilde{u}, \tilde{v}}}{\sigma_{\tilde{v}}^2} \tilde{v}$ and

$$\begin{pmatrix} \hat{u}_{it} \\ \hat{v}_{it} \end{pmatrix} \sim N_2 \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma = \begin{bmatrix} \omega_{11} & 0 \\ 0 & \sigma_{\tilde{v}}^2 \end{bmatrix} \right) \text{ with } \omega_{11} = \sigma_{\hat{u}}^2 = \sigma_{\tilde{u}}^2 - \frac{\sigma_{\tilde{u}, \tilde{v}}^2}{\sigma_{\tilde{v}}^2}$$

are independent $\forall i, t$. The joint likelihood of the model (4.1) is given by:

$$p(\tilde{y}, \tilde{w} | \tilde{y}_{-1}, \tilde{X}, \tilde{S}, \delta, \beta_0, \beta_1, \beta_2, \beta_3, \sigma_{\tilde{v}}^2, \omega_{11}) = p(\tilde{w} | \tilde{S}, \delta, \sigma_{\tilde{v}}^2) \prod_{t=2}^T p(\tilde{y}_{-t} | \tilde{y}_{-(t-1)}, \tilde{X}_t, \tilde{w}_t, \tilde{S}_t, \delta, \beta_0, \beta_1, \beta_2, \beta_3, \omega_{11})$$

where $\tilde{y}_{-t} = (\tilde{y}_{1t}, \dots, \tilde{y}_{Nt})'$, $\tilde{X}_t = \begin{pmatrix} \tilde{X}_{1t}' \\ \vdots \\ \tilde{X}_{Nt}' \end{pmatrix}$, $\tilde{w}_t = \begin{pmatrix} \tilde{w}_{1t} \\ \vdots \\ \tilde{w}_{Nt} \end{pmatrix}$, $\tilde{S}_t = \begin{pmatrix} \tilde{S}_{1t}' \\ \vdots \\ \tilde{S}_{Nt}' \end{pmatrix}$ so that:

$$\begin{aligned} \tilde{w} | S, \delta, \sigma_{\tilde{v}}^2 &\sim N_{NT}(\tilde{S} \delta, \sigma_{\tilde{v}}^2 I_{NT}) \\ \tilde{y}_{-t} | \tilde{y}_{-(t-1)}, \tilde{X}_t, \tilde{w}_t, \tilde{S}_t, \delta, \beta_0, \beta_1, \beta_2, \beta_3, \omega_{11} &\sim N_N(\tilde{y}_{-(t-1)} \beta_0 + \tilde{X}_t \beta_1 + \tilde{w}_t \beta_2 + \tilde{v}_t \beta_3, \omega_{11} I_N) \end{aligned} \quad (4.2)$$

Therefore, the joint density function of $(\tilde{y}, \tilde{w})$ is given by:

$$\begin{aligned} p(\tilde{y}, \tilde{w} | \tilde{y}_{-1}, \tilde{X}, \tilde{S}, \delta, \beta_0, \beta_1, \beta_2, \beta_3, \sigma_{\tilde{v}}^2, \omega_{11}) &\propto (\sigma_{\tilde{v}}^2)^{-(NT/2)} \times \\ \exp \left\{ \frac{1}{2\sigma_{\tilde{v}}^2} \left\{ (\tilde{w} - \tilde{S} \delta)' (\tilde{w} - \tilde{S} \delta) \right\} \right\} &I_{(0, \infty)}(\sigma_{\tilde{v}}^2) \times (\omega_{11})^{-(NT/2)} \times \\ \exp \left\{ \frac{1}{2\omega_{11}} \left\{ (\tilde{y} - \tilde{y}_{-1} \beta_0 + \tilde{X} \beta_1 - \tilde{w} \beta_2 - \tilde{v} \beta_3)' (\tilde{y} - \tilde{y}_{-1} \beta_0 + \tilde{X} \beta_1 - \tilde{w} \beta_2 - \tilde{v} \beta_3) \right\} \right\} &I_{(0, \infty)}(\omega_{11}) \end{aligned} \quad (4.3)$$

4.2. Prior Distribution

The prior distributions of regression coefficients $\delta, \beta_0, \beta_1, \beta_2, \beta_3$ are the usual conjugates, which are given by:

$$\begin{aligned} \delta &\sim N_{K_s}(\delta_0, V_0) \\ \beta_0 &\sim N(\beta_{00}, \sigma_{\beta_0}^2) \\ \beta_1 &\sim N_{K_x}(\beta_{01}, V_0 \beta_1) \\ \beta_2 &\sim N(\beta_{02}, V_0 \beta_2) \\ \beta_3 &\sim N(\beta_{03}, V_0 \beta_3) \end{aligned} \quad (4.4)$$

while, employing the same conjugates distributions, the prior distributions for variance components are given by:

$$\begin{aligned}\sigma_v^2 &\sim IG\left(\frac{a_0}{2}, \frac{b_0}{2}\right) \\ \omega_{11} &\sim IG\left(\frac{\mu_0}{2}, \frac{c_0}{2}\right)\end{aligned}\quad (4.5)$$

Therefore, the full prior distribution is specified as:

$$\begin{aligned}\pi(\delta, \beta_0, \beta_1, \beta_2, \beta_3, \sigma_v^2, \omega_{11}) &= \pi(\delta) \pi(\beta_0) \pi(\beta_1) \pi(\beta_2) \pi(\beta_3) \pi(\sigma_v^2) \pi(\omega_{11}) \\ &\propto \exp\left\{-\frac{1}{2}(\delta - \delta_0)' V_0^{-1}(\delta - \delta_0)\right\} \times \exp\left\{-\frac{1}{2}(\beta_0 - \beta_{00})' V_{0\beta_0}^{-1}(\beta_0 - \beta_{00})\right\} \times \\ &\exp\left\{-\frac{1}{2}(\beta_1 - \beta_{01})' V_{0\beta_1}^{-1}(\beta_1 - \beta_{01})\right\} \times \exp\left\{-\frac{1}{2}(\beta_2 - \beta_{02})' V_{0\beta_2}^{-1}(\beta_2 - \beta_{02})\right\} \\ &\times \exp\left\{-\frac{1}{2}(\beta_3 - \beta_{03})' V_{0\beta_3}^{-1}(\beta_3 - \beta_{03})\right\} \times (\sigma_v^2)^{-\left(\frac{a_0}{2}+1\right)} \exp\left\{-\frac{b_0}{2\sigma_v^2}\right\} I_{(0,\infty)}(\sigma_v^2) \\ &\times (\omega_{11})^{-\left(\frac{\mu_0}{2}+1\right)} \exp\left\{-\frac{c_0}{2\omega_{11}}\right\} I_{(0,\infty)}(\omega_{11})\end{aligned}\quad (4.6)$$

4.3. Posterior Distribution

Applying Bayes' Theorem, the posterior distribution is given by:

$$\begin{aligned}p(\delta, \beta_0, \beta_1, \beta_2, \beta_3, \sigma_v^2, \omega_{11} | \tilde{y}, \tilde{y}_{-1}, \tilde{X}, \tilde{w}, \tilde{S}) &\propto p(\tilde{w} | \tilde{S}, \delta, \sigma_v^2) p(\tilde{y} | \tilde{y}_{-1}, \tilde{X}, \tilde{w}, \tilde{S}, \delta, \beta_0, \beta_1, \beta_2, \beta_3, \sigma_v^2, \omega_{11}) \times \\ \pi(\delta, \beta_0, \beta_1, \beta_2, \beta_3, \sigma_v^2, \omega_{11}) &\propto (\sigma_v^2)^{-(NT/2)} \exp\left\{\frac{1}{2\sigma_v^2}(\tilde{w} - \tilde{S}\delta)'(\tilde{w} - \tilde{S}\delta)\right\} I_{(0,\infty)}(\sigma_v^2) \times \\ (\omega_{11})^{-(NT/2)} &\exp\left\{\frac{1}{2\omega_{11}}(\tilde{y} - \tilde{y}_{-1}\beta_0 + \tilde{X}\beta_1 - \tilde{w}\beta_2 - \tilde{v}\beta_3)'(\tilde{y} - \tilde{y}_{-1}\beta_0 + \tilde{X}\beta_1 - \tilde{w}\beta_2 - \tilde{v}\beta_3)\right\} I_{(0,\infty)}(\omega_{11}) \times \\ \exp\left\{-\frac{1}{2}(\delta - \delta_0)' V_0^{-1}(\delta - \delta_0)\right\} &\times \exp\left\{-\frac{1}{2}(\beta_0 - \beta_{00})' V_{0\beta_0}^{-1}(\beta_0 - \beta_{00})\right\} \times \\ \exp\left\{-\frac{1}{2}(\beta_1 - \beta_{01})' V_{0\beta_1}^{-1}(\beta_1 - \beta_{01})\right\} &\times \exp\left\{-\frac{1}{2}(\beta_2 - \beta_{02})' V_{0\beta_2}^{-1}(\beta_2 - \beta_{02})\right\} \times \\ (\sigma_v^2)^{-\left(\frac{a_0}{2}+1\right)} \exp\left\{-\frac{b_0}{2\sigma_v^2}\right\} &I_{(0,\infty)}(\sigma_v^2) \times (\omega_{11})^{-\left(\frac{\mu_0}{2}+1\right)} \exp\left\{-\frac{c_0}{2\omega_{11}}\right\} I_{(0,\infty)}(\omega_{11})\end{aligned}\quad (4.7)$$

This distribution (4.7) is not analytically tractable, and for this reason, we employ MCMC methods to obtain valuable information about it. Specifically, we use the No-U-Turn sampler algorithm introduced by (Hoffman & Gelman, 2014) in its version modified by the STAN probabilistic programming language (Carpenter et al., 2017). The main advantage of this algorithm with respect to Gibbs or Metropolis-Hastings is its ability to generate iterations capable of exploring the state space more efficiently, a virtue enhanced in high-dimensional parametric spaces, as may be the case for models with dynamic panel data or with endogeneity problems.

4.4. Extension to Several Endogenous Regressors

This procedure can be extended to consider the existence of two or more endogenous variables. In this case, we set up an instrumental regression model (3.1) for each endogenous variable and separately apply the stochastic search of valid instruments for each variable. The selected instruments will be incorporated in the equation system (4.1) from which inferences about the regression coefficients β_{2j} ; $j = 1, \dots, K_w$ of each endogenous variables in the regression model:

$$y_{it} = \alpha_i + \beta_0 y_{i,t-1} + \sum_{j=1}^{K_x} x_{jit} \beta_{1j} + \sum_{j=1}^{K_w} w_{jit} \beta_{2j} + u_{it}$$

can be obtained using an adaptation of the algorithm described in Subsection 4.3. An example can be found in the empirical example of Section 7.

5. Simulated Data

5.1. Model with Endogeneity Problems

We first check the performance of the proposed methodology in a simulated model, whose specification is given by:

$$\begin{aligned} y_{it} &= \beta_0 y_{i,t-1} + \beta_1 x_{it} + \beta_2 w_{it} + \alpha_i + u_{it} \\ w_{it} &= \phi_1 w_{i,t-1} + \phi_2 z_{it} + \varepsilon_{it} \end{aligned} \quad (5.1)$$

where $K_X = K_W = 1$ and $K_Z = 2$. We establish $\beta_0 = 0.8, \beta_1 = 3, \beta_2 = 2, \phi_1 = 0.5, \phi_2 = 1.5$. The correlation between the endogenous variable and the error term $\text{corr}(w_{it}, u_{it}); \forall i, t$ was chosen to be 0.75. To achieve the desired correlation, we randomly take samples from a Multivariate Normal as:

$$\begin{aligned} \begin{pmatrix} u_{it} \\ w_{it} \end{pmatrix} &\sim N \left(\begin{bmatrix} \mu_u \\ \mu_w \end{bmatrix}, \Sigma_{u,w} \right); \forall i, t \\ \Sigma_{u,w} &= \begin{bmatrix} \sigma_u & 0 \\ 0 & \sigma_w \end{bmatrix} \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \begin{bmatrix} \sigma_u & 0 \\ 0 & \sigma_w \end{bmatrix} \end{aligned} \quad (5.2)$$

where:

$$\begin{aligned} \mu_u &= 0, \mu_w = \phi_1 w_{i,t-1} + \phi_2 z_{it}; \forall i, t; \\ \sigma_u &\sim (U(2,5))^{\frac{1}{2}}, \sigma_w \sim (U(1,4))^{\frac{1}{2}}, \rho = 0.75 \end{aligned} \quad (5.3)$$

We also randomly select values for the remaining components $\varepsilon \sim N_{NT}(0, \sigma_w^2 I_{NT}); \alpha_i \sim U(-1,1)$

. For variables x and z , we take random samples from $x_{ij} \sim N(0,1); \forall i, j$
 $z_{ij} \sim N(0,1); \forall i, j$. The panel data

dimension corresponds to $N = 60$ and $T = 15$, as it is typical to employ the GMM framework for short T and large N or macro panels.

Stage 1: PCA analysis for relevant and exogenous instruments

To search for relevant instruments, we take the following prior values to start the Gibbs sampling:

$$\begin{aligned} (\tilde{w}_{it} | \tilde{s}_{jit}) &\sim N \left(\sum_{j=1}^{14} \tilde{s}_{jit} \delta_j, \sigma_v^2 \right); j = 1, \dots, 14 \\ \sigma_v^{2(0)} &\sim IG \left(\frac{a_0}{2} = 2, \frac{b_0}{2} = 2 \right) \\ \delta^{(0)} &\sim N_{K_Z}(0,1) \end{aligned} \quad (5.4)$$

The total number of possible components included was 14. It was necessary to delete the last component as its zero in all its values. After running four parallel chains of 10,000 iterations each and discarding the first 20% of the sample as burning, the results are listed in Table 5.1. The results yield a total of 6 components that are significant at a credibility level of at least 90%, so they will be candidates for the next step. The statistics used to report the results include the posterior mean (Pos. Mean), the posterior standard deviation (Pos. SD), the Monte Carlo Standard Error (MCSE = Pos. SD/ (ESS)^{1/2}), the 2.5th percentile (Q 2.5), the median or 50th percentile (Q 50), and the 97.5th percentile (Q 97.5). Additionally, the convergence measures, ESS and $\hat{R}$, are used to evaluate the convergence of the chains and the quality of the results. Notice that $\hat{R}$ represents the potential scale reduction factor, and ESS the Effective Sample Size, which are both defined in (Gelman et al., 2014). The MCSE serves also as a measure of the standard deviation of the estimates obtained from multiple independent simulations.

Table 5. 1. Results for Posterior PCA Analysis. (in bold signalled the selected principal components).

Component	Pos. Mean	Pos. SD	MCSE	Q 2.5	Q 50	Q 97.5	ESS	$\hat{R}$
Comp. 1	-0.1778***	0.0251	0.0001	-0.2271	-0.1776	-0.1285	42837	1.0001
Comp. 2	-0.1943***	0.0295	0.0001	-0.2522	-0.1943	-0.1363	42313	1.0001
Comp. 3	-0.1859***	0.0345	0.0002	-0.2534	-0.1857	-0.1186	42537	1.0000
Comp. 4	-0.2200***	0.0389	0.0002	-0.2959	-0.2200	-0.1448	42858	1.0001
Comp. 5	-0.2601***	0.0470	0.0002	-0.3526	-0.2604	-0.1685	42673	1.0001
Comp. 6	-0.1406**	0.0561	0.0003	-0.2518	-0.1402	-0.0314	42701	1.0001
Comp. 7	0.0022	0.0619	0.0003	-0.1190	0.0021	0.1251	43031	1.0000
Comp. 8	0.0133	0.0663	0.0003	-0.1159	0.0131	0.1436	42890	1.0001
Comp. 9	0.0608	0.0723	0.0004	-0.0795	0.0603	0.2033	41819	1.0000
Comp. 10	0.0331	0.0799	0.0004	-0.1239	0.0334	0.1903	42513	1.0001
Comp. 11	-0.1260	0.0821	0.0004	-0.2860	-0.1257	0.0339	42709	1.0000
Comp. 12	-0.0009	0.0921	0.0004	-0.1791	-0.0010	0.1799	42411	1.0000
Comp. 13	-0.0280	0.1064	0.0005	-0.2382	-0.0285	0.1798	42713	1.0001
Comp. 14	0.2278	0.1469	0.0007	-0.0598	0.2278	0.5155	42410	1.0001

Coefficients marked with (*) represent significant variables at 90% credible level, (**) at 95% and (***) at 99%.

Stage 2: Posterior Inference on the full model

In the second step, we specify the following setting:

$$\begin{aligned} \left(\tilde{w}_{it} \mid \tilde{s}_{jit} \right) &\sim N \left(\sum_{j=1}^6 \tilde{s}_{jit} \delta_j, \sigma_v^2 \right); \forall i, t; j = 1, \dots, 6 \\ \left(\tilde{y}_{it} \mid \tilde{y}_{i,t-1}, \tilde{x}_{it}, \tilde{w}_{it}, \tilde{v}_{it} \right) &\sim N \left(\beta_0 \tilde{y}_{i,t-1} + \beta_1 \tilde{x}_{it} + \beta_2 \tilde{w}_{it} + \beta_3 \tilde{v}_{it}, \omega_{11} \right); \forall i, t \\ \tilde{v}_{it} &= \tilde{w}_{it} - \sum_{j=1}^6 \tilde{s}_{jit} \delta_j \\ \left(\beta_0^{(0)}, \beta_1^{(0)}, \beta_2^{(0)}, \beta_3^{(0)} \right) &\sim N(0,1) \\ \sigma_v^{2(0)} &\sim IG(2,2) \\ \omega_{11}^{(0)} &\sim IG(2,2) \end{aligned} \tag{5.5}$$

In this case, we launched four chains of 10,000 iterations each, using the NUTS algorithm in STAN, for a model with and without endogeneity treatment, that is, we compare an estimated model where it is kept the assumption of no correlation between, or a Fixed Effects Model (FE), and a model in which we suppose some degree of correlation, so Instrumental Variables are employed (IVFE). Mean

Squared Error (MSE) was used to compare the point estimates with real values. Our main goal is to demonstrate that proposed methodology can provide better, and more accurate estimations of parameters affected by endogeneity problems than simple Fixed Effects models. Table 5.2 presents the results.

Table 5. 2. Results for Posterior Exploration of Simulated model with $\varrho = 0.75$.

[illegible]

As can be seen, the chains seem to converge properly, and there are no problems during the algorithm execution, as $\hat{R}$ is near 1 in all estimations. Furthermore, the ESS statistic suggests that, in all cases, there has been a sufficient volume of statistically independent replicates to have confidence in the quality of the approximations. Results are also quite successful both in terms of the posterior inference of the parameter of interest, β_2 , and those less important for the exposition, β_0 and β_1 . The point estimate presents a better approximation of the real value, reducing the Mean Squared Error (MSE) from 0.1522 to 0.0003. Additionally, the 95% credible intervals contain the real values of all the parameters, which makes the process more reliable. However, this is not the case for β_2 in the FE model. The value of the WAIC model selection criterion (Watanabe, 2013), and BIC criterion (Schwarz, 1978), was calculated to select between the two methods. Both criteria are minimized between the two alternatives for the IVFE case, which provides additional support for the use of the methodology proposed in this study in the simulated case.

5.2. Model without Endogeneity Problems

We now verify that the application of the methodology when there is no endogeneity problem, or at least when this problem is not sufficiently relevant, does not cause significant changes in the posterior inference of the parameters of interest. In this case, we use the model (5.1) but now we assume that $\rho = \text{corr}(w_{it}, u_{it}) = 0; \forall i, t$ and, therefore, there are not endogeneity problems between u and w . The results of the estimation process are shown in Table 5.3.

In this case, good precision in the estimation of the parameter of interest β_2 in the IVFE model is still observed, even surpassing the estimation in the FE model in terms of MSE. In addition, the true value is still not contained within the limits of the 95% credibility interval. Regarding model selection criteria, the reported values are very similar with respect to both WAIC and BIC criteria, indicating that the methodology is robust enough to be used in applied cases, even when true endogeneity problems are not present.

Table 5. 3. Results for Posterior Exploration of Simulated model with $\varrho = 0$.

[illegible]

6. Empirical Study: World Governance Indicators and Bank Capital Flows

In this example, we are interested in quantifying the partial effect of a set of economic, financial, and governance indicators on the International Bank Inflows received by a set of emerging economies, motivated by the work of Kim & Wu (2008). We mainly try to study the role played by public sector as policy maker to favour international bank inflows. In general, it's considered that some country risk indicators behave as a negative pull factor on banking flows, while others could be considered as positive. However, to guide economic policy, we should also attempt to discard the effects that actually present significant effects from those that do not, instead of considering the entire set of governance indicators as a homogeneous cluster. To this end, the possibility of an endogenous regressor among governance determinants is raised, motivated by the fact that there may be unobservable effects strongly correlated with indicators that affect the quality of internal development of the economy.

6.1. Variables and Data

Our dataset corresponds to N=50 economies considered as emerging economies in the original work and T=26 years, which ranges from 1996 to 2021. However, the panel contains missing values, around 30% of the initial sample, for some exogenous variables, so the number of observations was reduced to 915 once these values were discarded. The presence of missing data does not affect the implementation of the proposed methodology, as the forward orthogonal deviation transformation must be adapted to this setting. We consider three main groups of variables (see Appendix B for a description of the variables). On the one hand, TRADE and LGDPPC reflect the behaviour of the economic cycle. The second group corresponds to financial variables, such as SMCAPLISTED and PRIVCREDIT. The third group comprises a set of government indicators to quantify the institutional performance and socio-economic features of the countries analysed. These indicators were ACCOUNT, RULAW, REGQLTY, CORRUPTION, GOVEFF, and POLSTA. Finally, we included only the Foreign Currency Long Term Credit Rating (SPRATING), issued by the S&P agency, incorporated through a rating discretization process at each point in time, coupled with their perspective to change. According to the S&P scale, for long-term issues the ratings range from the highest investment grade (AAA) to the Default category (D or SD). The outlook for rating changes varies from a positive credit review outlook to a negative credit review outlook. To convert the ratings issued, a numerical value is assigned to each rating, ranging from 21 (AAA) to 1 (C), and a numerical value is added according to the outlook for change, ranging from 0.5 (watch positive) to -0.5 (negative watch). If there have been several issues in the same time period, the simple average of the quantifications of all of them is taken as the value. As an illustrative example, if a country has a rating of (BBB+) and a positive revision outlook, the value of the variable will be the sum of 13 and 0.25, according to the proposed scale.

The data set has been extracted from the World Bank's World Development Indicators (WDI) and World Governance Indicators (WGI) databases, and from information provided by the credit rating agency Standar & Poors. Appendix B contains detailed information on the variables and the data used in this study. Furthermore, to maintain the spirit of the original proposal, the target variable is lagged one period.

The target variable, LCBLOAN, captures the total amount of bank capital inflows in logarithmic terms in the analysed economies, taking as a reference the flows reported by the Bank for International Settlement (BIS), issued by a set of banking institutions. We prefer this definition of the target variable mainly because we are interested in quantifying the effects on attract capital, rather than simply incorporating the net effect, as we consider that inflows and outflows could respond to different economic and political incentives.

6.2. Model Specification

Based on the above, the model under study is as follows:

$$\begin{aligned}
LCBLOAN_{it} = & \alpha_i + \beta_{10}LCBLOAN_{i,t-1} + \beta_{11}TRADE_{it} + \beta_{12}LGDPPEC_{it} \\
& + \beta_{13}SMCAPLISTED_{it} + \beta_{14}PRIVCRED_{it} + \beta_{15}SPRATING \\
& + \beta_{16}ACCOUNT_{it} + \beta_{17}RULAW_{it} + \beta_{18}POLST_{it} + \\
& \beta_{21}REGQLTY_{it} + \beta_{22}CORRUPTION_{it} + \beta_{23}GOVEFF_{it} + u_{it}
\end{aligned} \tag{6.1}$$

In line with what has been established throughout the paper, the model perturbations are assumed to be homoscedastic, independent and normally distributed, so that $u_{it} \sim N(0, \sigma_u^2)$, $i = 1, 2, \dots, 50$, $t \in T_i \subseteq \{1, 2, \dots, 26\}$. The endogenous variables considered were REGQLTY, CORRUPTION, and GOVEFF, while the rest were considered strictly exogenous.

We consider these variables to be endogenous primarily because they represent key aspects of government performance that directly influence the promotion and development of the private sector. The variable REGQLTY measures citizen perceptions regarding the effectiveness of government efforts to foster private sector growth. Similarly, the level of corruption, as captured by CORRUPTION, can be viewed as an indicator of the degree of extractive impunity exercised by the public sector over the private sector. This, in turn, reflects the public sector's ability to generate revenue without engaging in productive activities. Additionally, the inclusion of GOVEFF is justified by the perception that public goods and services are provided efficiently and that public policies are well-designed and implemented. This efficiency in public sector performance reduces the need for private sector intervention. For example, if the public sector efficiently provides hospital infrastructure and health services, there will be limited private incentives to establish new facilities or offer additional services, leading private capital to be allocated to other areas of need. These variables, as a last resort, can be found to be related to latent aspects that affect the performance of our target variable, such as degree of liberalization of domestic markets, the competitiveness of the internal economy, or the degree of the shadow economy and, therefore, generate an endogeneity problem in the model. Therefore, according to the specifications, it is established that $K_x = 8$ and $K_w = 3$.

6.3. Instruments Search

Once the Forward Orthogonal Deviation transformation has been applied, we establish:

$$\begin{aligned}
(\tilde{w}_{jit} | \tilde{s}_{jlit}) & \sim N\left(\sum_{l=1}^{33} \tilde{s}_{jlit} \delta_{jl}, \sigma_{\tilde{v}_j}^2\right); j = 1, 3; l = 1, \dots, 33 \\
\sigma_{\tilde{v}_j}^2 & \sim IG(2, 2); j = 1, 2, 3 \\
\delta_{jl} & \sim N(0, 1); \forall j, l
\end{aligned} \tag{6.2}$$

For ease of notation, $w_{jit}; j = 1, 2, 3$ represents REGQLTY, CORRUPTION, and GOVEFF. The model employs 33 instruments since last component, the number 34, compound by the sum of 8 exogenous variables plus 26 lagged collapsed instruments, is discarded as it 0 in all values.

We ran, again, four parallel chains of 10,000 iterations each, discarding the 20% of the first sample as burning. The SSVS method applied to Principal Components in our contest yields a total of sixteen components for REGQLTY, whose posterior significance is greater than 50%, a total of eight for CORRUPTION, and three for GOVEFF. Compared to Roodman's matrix, which does not include exogenous variables, and collapsed the instrument matrix, we use only sixteen variables for posterior inference to the first endogenous variable, and eight and three for the second and third one respectively, against 26 of the main proposal. Not only do we include more information, but we also significantly reduce the number of variables to be employed without losing relevant information.

6.4. Posterior Inference on the Full Model

We now estimate the model proposed using a simple fixed effects model with Forward Orthogonal Deviation transformation, which we call the Fixed Effects Model (FE), and the same

proposal using the methodology proposed in this paper, called the instrumental variable fixed effect model (IVFE). For FE, we provide the following prior and likelihood specifications:

$$\begin{aligned} (\tilde{y}_{it} | \tilde{y}_{i,t-1}, \tilde{x}'_{it}, \tilde{w}'_{it}) &\sim N\left(\beta_{10}\tilde{y}_{i,t-1} + \sum_{j=1}^8 \beta_{1j}\tilde{x}_{jit} + \sum_{j=1}^3 \beta_{2j}\tilde{w}_{jit}, \sigma^2\right); \forall i, t \\ (\beta_{1j}^{(0)}, \beta_{2j}^{(0)}) &\sim N(0,1); \forall j \\ \sigma^{2(0)} &\sim IG(2,2) \end{aligned} \quad (6.3)$$

Notice that $y_{t-1} = \text{LCBLOAN_1}$,
 $X' = [\text{TRADE, LGDPPC, SMCAPLISTED, PRIVCRED, SPRATING, RULAW, ACCOUNT}]'$,
 $W' = [\text{REGQLTY, CORRUPTION, GOVEFF}]'$.

For the IVFE model, we establish:

$$\begin{aligned} (\tilde{w}_{lit} | \tilde{s}_{lit}) &\sim N\left(\sum_{j=1}^{16} \tilde{s}_{lit}\delta_j, \sigma_{1\tilde{v}}^2\right); \forall i, t \\ (\tilde{w}_{2it} | \tilde{s}_{2it}) &\sim N\left(\sum_{j=1}^8 \tilde{s}_{2it}\delta_j, \sigma_{2\tilde{v}}^2\right); \forall i, t \\ (\tilde{w}_{3it} | \tilde{s}_{3it}) &\sim N\left(\sum_{j=1}^3 \tilde{s}_{3it}\delta_j, \sigma_{3\tilde{v}}^2\right); \forall i, t \\ (\tilde{y}_{it} | \tilde{y}_{i,t-1}, \tilde{x}'_{it}, \tilde{w}'_{it}) &\sim N\left(\beta_{10}\tilde{y}_{i,t-1} + \sum_{j=1}^8 \beta_{1j}\tilde{x}_{jit} + \sum_{j=1}^3 \beta_{2j}\tilde{w}_{jit} + \sum_{j=1}^3 \beta_{3j}\tilde{v}_{jit}, \omega_{11}\right); \forall i, t \\ \tilde{v}_{lit} &= \tilde{w}_{lit} - \sum_j \tilde{s}_{lit}\delta_j; l=1,2,3 \\ (\beta_{1j}^{(0)}, \beta_{2j}^{(0)}, \beta_{3j}^{(0)}) &\sim N(0,1); \forall j \\ \sigma_{j\tilde{v}}^{2(0)} &\sim IG(2,2); j=1,2,3 \\ \omega_{11}^{(0)} &\sim IG(2,2) \end{aligned} \quad (6.4)$$

We launched four chains of 10,000 iterations each using the NUTS algorithm in STAN, discarding 20% of the first samples as burning. The results are shown in Table 6.1. The convergence measure, $\hat{R}$, and ESS present reasonable values, and the Monte Carlo standard error (MCSE) is small enough to rely on the results. Results report favourable values for WAIC and BIC model selection criterion to our proposal, therefore, the IVFE model has a better goodness of fit to observed data. In both models, LGDPCC, PRIVCRED, SPRATING, and LCBLOAN_1 are significant variables at the 99% credible level. Variables TRADE and SPRATING was found significant at 95% credibility level in the IVFE model, and variables SMCAPLISTED, POLSTA and GOVEFF were found significant in the IVFE model at 99% credibility level.

Economically, results reported by IVFE model seem to be more consistent with our prior beliefs. TRADE and LGDPPC are both positive in FE and IVFE models, being also significant in the IVFE model, which means that economic cycle plays a positive and relevant partial effect on attracting international bank flows. Internal private credit (PRIVCRED) is significant and positive in both models; therefore, internal private indebtedness acts as a pull factor. In the same way, credit ratings reflect that better trustworthiness of debtors leads to receiving international bank flows. Market

Capitalization plays also its positive role in attracting inflows. The main differences arise when comparing country-risk drivers.

Table 6. 1. Results for FE and IVFE Model.

Variable	Mean	SD	Q 2.5	Q 50	Q 97.5	MCSE	ESS	$\hat{R}$	Mean	SD	Q 2.5	Q 50	Q 97.5	MCSE	ESS	$\hat{R}$
FE Model									IVFE Model							
TRADE	0.0004	0.0008	-0.0011	0.0004	0.0019	0.0000	37541	1.0001	0.0022**	0.0009	0.0005	0.0022	0.0040	0.0000	25919	1.0001
LGDPPC	0.2301***	0.0193	0.1926	0.2301	0.2681	0.0001	24886	0.9999	0.3578***	0.0347	0.2902	0.3573	0.4273	0.0002	27661	0.9999
SMCAPLISTED	0.0007	0.0006	-0.0005	0.0007	0.0018	0.0000	32424	1.0002	0.0053***	0.0017	0.0024	0.0052	0.0091	0.0000	17254	1.0002
PRIVCRED	0.0073***	0.0006	0.0061	0.0073	0.0085	0.0000	32767	1.0001	0.0106***	0.0014	0.0082	0.0105	0.0136	0.0000	17701	1.0001
SPRATING	0.0289***	0.0063	0.0165	0.0289	0.0411	0.0000	28172	1.0001	0.0167**	0.0069	0.0031	0.0167	0.0302	0.0000	43398	1.0001
RULAW	0.0088	0.0044	0.0001	0.0088	0.0175	0.0000	33904	0.9999	0.0064	0.0044	-0.0023	0.0064	0.0151	0.0000	70333	0.9999
ACCOUNT	-0.0453	0.0285	-0.1010	-0.0455	0.0109	0.0002	26567	1.0001	-0.0102	0.0472	-0.1016	-0.0107	0.0839	0.0003	18431	1.0001
POLSTA	0.0002	0.0002	-0.0002	0.0002	0.0006	0.0000	31112	1.0000	-0.0029***	0.0010	-0.0050	-0.0029	-0.0012	0.0000	17485	1.0000
LCBLOAN_1	0.2721***	0.0138	0.2451	0.2721	0.2993	0.0001	26924	1.0000	0.2478***	0.0143	0.2195	0.2480	0.2757	0.0001	47862	1.0000
REGQLTY	-0.0094	0.0152	-0.0390	-0.0093	0.0205	0.0001	28274	0.9999	-0.0322	0.0176	-0.0668	-0.0322	0.0019	0.0001	34787	0.9999
CORRUPTION	-0.012	0.0312	-0.0732	-0.0122	0.0500	0.0002	26016	1.0001	-0.0244	0.1790	-0.3916	-0.0209	0.3184	0.0014	15343	1.0001
GOVEFF	0.0062	0.0236	-0.0403	0.0062	0.0520	0.0001	27062	1.0002	-1.8944***	0.4160	-2.7785	-1.8689	-1.1469	0.0037	12751	1.0002
Coefficients marked with (*) represent significant variables at 10% credible level, (**) at 5% and (***) at 1%																
WAIC(FE) = 1133.9		BIC(FE) = 1181.35														
WAIC(IVFE) = 1089.4		BIC(IVFE) = 1137.45														

Comparing our results with those obtained by Kim and Wu (2008), and accounting for the potential endogeneity in the aforementioned government components, certain observations stand out. In the original study, the variables ACCOUNT and GOVEFF were the only ones with positive signs, indicating that marginal increases in both indicators led to net positive inflows of external banking capital, although only ACCOUNT had a significant estimate. In our estimations, the FE model captures a partial positive effect in the variables RULAW, POLSTA, and GOVEFF, although none of these effects are significant. On the other hand, the IVFE model estimates a positive marginal effect, though not significant, only for the variable RULAW. This finding supports the argument made by (Koepke, 2019) that government indicators are deterrents in attracting external banking flows. Moreover, the IVFE model identifies significant estimates in the variables POLSTA and GOVEFF, effects that are not captured by the FE model. We consider these findings to be economically relevant because both indicators are similar in that they help identify the degree of political stability, its independence, and citizens' perceptions of the public services provided. Our results also corroborate the idea that foreign banks use a lower perception of the internal development of the private economy to locate financing flows. In fact, the idea that in emerging economies, pressure on the private sector in the form of regulation generates business opportunities for the entry of external bank capital gains strength, in view of our results. Additionally, it's relevant the fact that IVFE can isolate the relevant partial effect when using governance indicators from those that don't provide useful added value. In summary, there is some evidence of the positive economic role played by the government sector in attracting foreign bank capital, mainly through the credit rating issued, although the model estimates that an internal capitalized economy is less dependent on foreign financial flows. The combination of good credit quality, along with a favorable development of the private sector of the economy, seems to be a determining factor in the attraction of external bank funds by the emerging economies analyzed.

7. Discussion

In this paper, we propose a general procedure to build instrumental matrices of reduced dimensions in a panel regression model with endogenous explanatory variables. The procedure is based on the collapsed version proposed by Roodman (2009 a, b) and the use of the principal components of instrumental variables, the number of which is selected according to our Bayesian proposal. Our method improves upon existing alternatives by allowing the selection of empirically relevant instruments, preventing issues of multicollinearity among these instruments from causing overfitting in the point estimates of the parameters of interest, while also facilitating the exploration of the sample space due to the orthogonality of the principal components. Once the instrumental matrix has been built, the No-U-Turn sampling algorithm by (Hoffman & Gelman, 2014) is used to make more accurate inferences about the regression coefficients of the endogenous variables. The

performance of the methodology is analysed by employing a simulated example, with one endogenous variable among the set of regressors and one empirical application to the partial effect of World Governance Indicators on international bank flows. The existence of potential endogeneity in measuring some of these indicators justifies the use of instrumental variables and, therefore, the methodology proposed in this paper. The results obtained are more coherent from an economic perspective, and statistically are more consistent.

Although the methodology seems promising for dealing with endogeneity problems in dynamic and static panel data settings, with one or more potential endogenous regressors, some questions must be raised to improve the proposal and overcome some limitations. Mainly, it is based on the assumptions of normality and homoscedasticity of the error structure. We are currently working to eliminate this limitation. Additionally, the methodology should be tested in other simulation scenarios to corroborate the promising results obtained in this work.

Data Availability Statement: The datasets used and/or analysed during the current study are available from the corresponding au-thor on reasonable request.

Funding: This research received no external funding.

Author Contributions: Álvaro Herce (AH) and Manuel Salvador (MS) contributed equally to the development of this paper. AH, as part of his PhD studies, led the conception and design of the study, including the formulation of the Bayesian econometric model and the application of the Monte Carlo Markov Chain (MCMC) methods. He was also responsible for the simulation experiments and the empirical analysis regarding international bank flows. Manuel Salvador contributed to the methodological framework, particularly in the selection of instrumental variables using Principal Component Analysis (PCA). He also provided significant input into the interpretation of the results and the refinement of the final model. Both authors participated in writing and revising the manuscript and approved the final version.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Deriving Full-Conditional Densities in the Selection of Principal Components

In obtaining the conditional densities used in the development of the Gibbs algorithm for the selection of principal components as valid instruments for instrumental regression, the following developments must be considered, starting from the following specifications:

$$\begin{aligned}\tilde{w} | \tilde{S}, \delta, \sigma_v^2 &\sim N_{NT}(\tilde{S}\delta, \sigma_v^2 I_{NT}) \\ \pi(\sigma_v^2, \delta) &\propto \left(\frac{1}{\sigma_v^2}\right) I_{(0,\infty)}(\sigma_v^2) \\ p(\delta, \sigma_v^2 | \tilde{w}) &\propto (\sigma_v^2)^{-\left(\frac{NT}{2}+1\right)} \exp\left\{-\frac{1}{2\sigma_v^2}(\tilde{w}-\tilde{S}\delta)'(\tilde{w}-\tilde{S}\delta)\right\} I_{(0,\infty)}(\sigma_v^2)\end{aligned}$$

Starting from the posterior distribution above, we obtain:

$$\begin{aligned}p(\sigma_v^2 | \delta, \tilde{w}) &\propto (\sigma_v^2)^{-\left\{\frac{NT}{2}+1\right\}} \exp\left\{-\frac{(\tilde{w}-\tilde{S}\delta)'(\tilde{w}-\tilde{S}\delta)}{2\sigma_v^2}\right\} I_{(0,\infty)}(\sigma_v^2) \\ &\propto (\sigma_v^2)^{-\{a_1+1\}} \exp\left\{-\frac{b_1}{\sigma_v^2}\right\} I_{(0,\infty)}(\sigma_v^2)\end{aligned}$$

where

$$a_1 = \frac{NT}{2}$$
$$b_1 = \frac{\left(\tilde{w}-\tilde{S}\delta\right)'\left(\tilde{w}-\tilde{S}\delta\right)}{2}$$

The conditional posterior density of the structural parameters δ would be:

$$\begin{aligned} p\left(\delta|\tilde{w},\sigma_v^2\right) &\propto \exp\left\{-\frac{1}{2\sigma_v^2}\left[\left(\tilde{w}-\tilde{S}\delta\right)'\left(\tilde{w}-\tilde{S}\delta\right)\right]\right\}I_{(0,\infty)}\left(\sigma_v^2\right) \\ &\propto \exp\left\{-\frac{1}{2\sigma_v^2}\left[\delta'\tilde{S}'\tilde{S}\delta-2\delta'\tilde{S}'\tilde{w}\right]\right\}I_{(0,\infty)}\left(\sigma_v^2\right) \\ &\propto \exp\left\{-\frac{1}{2}\left[\delta'\left(\frac{1}{\sigma_v^2}\tilde{S}'\tilde{S}\right)\delta-2\delta'\left(\frac{1}{\sigma_v^2}\tilde{S}'\tilde{S}\right)\sigma_v^2\left(\tilde{S}'\tilde{S}\right)^{-1}\tilde{S}'\tilde{w}\right]\right\}I_{(0,\infty)}\left(\sigma_v^2\right) \\ &\propto \exp\left\{-\frac{1}{2}\left[\delta'V_1^{-1}\delta-2\delta'V_1^{-1}\hat{\delta}+\hat{\delta}'V_1^{-1}\hat{\delta}\right]\right\}I_{(0,\infty)}\left(\sigma_v^2\right) \\ &\propto \exp\left\{-\frac{1}{2}\left[\left(\delta-\hat{\delta}\right)'V_1^{-1}\left(\delta-\hat{\delta}\right)\right]\right\}I_{(0,\infty)}\left(\sigma_v^2\right) \end{aligned}$$

where:

$$V_1 = \sigma_v^2\left(\tilde{S}'\tilde{S}\right)^{-1}$$
$$\hat{\delta} = \left(\tilde{S}'\tilde{S}\right)^{-1}\tilde{S}'\tilde{w}$$

This demonstrates that:

$$\sigma_v^2\mid\delta,\tilde{w}\sim IG\left(a_1,b_1\right);a_1=\frac{NT}{2},\,b_1=\frac{\left(\tilde{w}-\tilde{S}\delta\right)'\left(\tilde{w}-\tilde{S}\delta\right)}{2}$$
$$\delta\mid\sigma_v^2,\tilde{w}\sim N_{K_Z}\left(\hat{\delta},V_1\right);\hat{\delta}=\left(\tilde{S}'\tilde{S}\right)^{-1}\tilde{S}'\tilde{w},\,V_1=\sigma_v^2\left(\tilde{S}'\tilde{S}\right)^{-1}$$

Appendix B. Description of Countries and Applicable Variables

Variable	Description	Obs	Origin
LCBLOAN	Natural log of total loans of BSI reporting banks vis-à-vis individual surveyed countries (in billions of dollars)	1300	BIS, Locational Statistics
TRADE	Percentage of the economy’s GDP of the total value of commercial exchanges	1285	World Bank, WDI
LGDP	Natural log of GDP per Capita	1293	World Bank, WDI
SMCAPLISTED	Market capitalization of listed companies, taking the year-end value as a result (% of GDP)	830	World Bank, WDI
PRIVCREDIT	Percentage of credit granted to the private sector as a percentage of GDP	1208	World Bank, Global Financial Development
ACCOUNT	Degree of perception by which a country’s citizens participate in the election of governments, freedom of expression, freedom of association, and freedom of communication	1272	World Bank, Worldwide Governance Indicators
RULAW	Captures citizen’s perception of compliance with the law and social rules	1300	World Bank, Worldwide Governance Indicators

REGQTY	Captures citizen’s perceptions of the government’s ability to enact and implement sound policies and regulations, to promote private sector development	1300	World Bank, Worldwide Governance Indicators
CORRUPTION	It captures the degree of perception of the power of the public sector to exert pressure on the private sector, including any form of corruption	1300	World Bank, Worldwide Governance Indicators
GOVEFF	Captures the perception of the quality of public services, the quality of civil services and the degree of their independence from public authorities, the quality of policy formulation and its implementation	1300	World Bank, Worldwide Governance Indicators
POLSTA	Measures the perception of the plausibility of political instability and/or political violence, including the possibility of terrorist acts	1300	World Bank, Worldwide Governance Indicators
SPRATING	Long-term credit quality indicator or rating, provided by Standard & Poor’s and denominated in foreign currency	1300	Standard & Poor’s

References

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2), 277–297. <https://doi.org/10.2307/2297968>

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1), 29–51. [https://doi.org/10.1016/0304-4076\(94\)01642-D](https://doi.org/10.1016/0304-4076(94)01642-D)

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76. <https://doi.org/10.18637/jss.v076.i01>

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2014). *Bayesian data analysis*. Chapman and Hall/CRC.

Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721–741.

Greenberg, E. (2012). *Introduction to Bayesian econometrics*. Cambridge University Press.

Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.*, 15(1), 1593–1623.

Holtz-Eakin, D., Newey, W., & Rosen, H. S. (1988). Estimating vector autoregressions with panel data. *Econometrica: Journal of the Econometric Society*, 1371–1395. <https://doi.org/10.2307/1913103>

Jeffreys, H. (1961). *The theory of probability* (3rd ed.). Clarendon Press, Oxford.

Kim, S.-J., & Wu, E. (2008). Sovereign credit ratings, capital flows and financial sector development in emerging markets. *Emerging Markets Review*, 9(1), 17–39. <https://doi.org/10.1016/j.ememar.2007.06.001>

Koepke, R. (2019). What Drives Capital Flows to Emerging Markets? A Survey of the Empirical Literature. *Journal of Economic Surveys*, 33(2), 516–540. <https://doi.org/10.1111/joes.12273>

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, 1417–1426. <https://doi.org/10.2307/1911408>

Roodman, D. (2009a). A Note on the Theme of Too Many Instruments. *Oxford Bulletin of Economics and Statistics*, 71(1), 135–158. <https://doi.org/10.1111/j.1468-0084.2008.00542.x>

Roodman, D. (2009b). How to do xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal*, 9(1), 86–136.

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2), 461–464.

Watanabe, S. (2013). A widely applicable Bayesian information criterion. *The Journal of Machine Learning Research*, 14(1), 867–897.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.