Semiparametric Forecasting Problem in High Dimensional Dynamic Panel with Correlated Random Effects: A Hierarchical Empirical Bayes Approach

Pacifico, Antonio

LUISS Guido Carli University, CEFOP-LUISS

2021

Semiparametric Forecasting Problem in High Dimensional Dynamic Panel with Correlated Random Effects: A Hierarchical Empirical Bayes Approach

Abstract

This paper aims to address semiparametric forecasting problem when studying high dimensional data in multivariate dynamic panel model with correlated random effects. A hierarchical empirical Bayesian perspective is developed to jointly deal with incidental parameters, structural framework, unobserved heterogeneity, and model misspecification problems. Methodologically, an `ad-hoc` model selection on a mixture of normal distributions is addressed to obtain the *best* combination of outcomes to construct empirical Bayes estimator and then investigate ratio-optimality and posterior consistency for better individual–specific forecasts. Simulated and empirical examples are conducted to highlight the performance of the estimating procedure.

JEL classification: A1; C01; E02; H3; N01; O4

Keywords: Dynamic Panel Data; Ratio-Optimality; Bayesian Methods; Forecasting; MCMC Simulations; Tweedie Correction.

1 Introduction

This paper aims to construct and develop a methodology to improve the recent literature on Dynamic Panel Data (DPD) models when dealing with (*i*) individual–specific forecasts, (*ii*) Bayesian analyses with parametric priors on heterogeneous parameters, (*iii*) ratio-optimality and posterior consistency in dynamic panel setups, (*iv*) empirical Bayes estimator and alternative Tweedie corrections, and (*v*) the curse of dimensionality when estimating time-varying data.

DPD models are widely used in empirical economics for forecasting individuals' future outcomes (see, e.g., Hirano (2002), Gu and Koenker (2017b), Liu (2018), and Liu et al. (2020)) and allowing the possibility of controlling for unobserved time-invariant individual heterogeneity (see, e.g., Chamberlain (1984) and Arellano and Bond (1991) (linear case); and Chamberlain (2010) and Arellano and Bonhomme (2011) (non-linear case)). Such heterogeneity is an important issue and failure to control for it results in misleading inferences. That problem is even more severe when the unobserved heterogeneity may be correlated with covariates.

Consider a simple DPD model:

$$y_{it} = w_{i,t-1}\mu_i + \beta y_{i,t-1} + u_{it} \tag{1}$$

where $i = 1, \ldots, N$, $t = 1, \ldots, T$, y_{it} and $y_{i,t-1}$ denote the outcomes and their first lags, μ_i refers to individual-specific intercept with $w_{i,t-1} = 1$, and $u_{it} \sim N(0, \sigma^2)$ is an independent and identically distributed (*i.i.d.*) shock.

In the dynamic panel literature, the focus is to find a consistent estimate of β in the presence of the incidental parameters μ_i to avoid the incidental parameters problem and then perform better forecasts of the outcomes in period $T+1$ (y_{T+1}). In the context of panel data, the incidental parameters problem typically arises from the presence of individual–specific factors. The challenges because of incidental parameters are highly severe in dynamic panels where behavioural effects over time are jointly measured with individual–specific effects. Whereas the incidental parameters to be estimated are consistent in least squares methods, maximum likelihood estimation leads to inconsistent estimates of them affecting the dynamics of data (see, for instance, Nickell (1981)). Both fixed and random effects have been used to evaluate these individual–specific factors. The former treats them as parameters to be estimated, leaving the distribution of unobserved heterogeneity relatively unrestricted at the cost of introducing a large number of nuisance parameters; random effects typically assume that their distributions belong to a known parametric family indexed by a finite dimensional parameter.

However, in these traditional methods, with large cross-sectional dimension (N) and short (fixed) time-series (T), the estimators of the common parameters (β, σ^2) would result biased and inconsistent due to the incidental parameter problems. Indeed, leaving the individual heterogeneity unrestricted, the number of individual–specific effects would grow with the sample size. Moreover, with short T, the estimates of the heterogeneous parameters (μ_i) would be highly contaminated from the shock u_{it} obtaining inaccurate forecasts. Last but not least, when dealing with time-varying and high dimensional data, problems concerning (*i*) overshrinkage/undershrinkage, (*ii*) functional forms of misspecification, (*iii*) endogeneity issues, and (*iv*) variable selection problem also matter in DPD models involving inconsistent estimates.

The methodology proposed in this paper focuses on the aforementioned issues and takes the name of Dynamic Panel Bayesian model with Correlated Random Effects (DPB-CRE). It develops a hierarchical structural empirical Bayes approach for inference in multivariate dynamic panel setup with cross-sectional heterogeneity; where 'structural' stands for designing a more conventional empirical procedure to provide reduced-form causal relationships. Methodologically, a Finite Mixture approximation of Multivariate (FMM) distributions is used to construct the Empirical Bayes (EB) estimator for alternative Tweedie corrections to avoid the impossibility of the oracle forecast of computing the correlated random effect distribution (or prior). The multivariate panel considered in this paper is unbalanced and includes large cross-sectional dimension N and sufficiently large time-series T. Let the framework be hierarchical, Conjugate Informative Proper Mixture (CIPM) priors are used to select promising model fitting the data, acting as a strong model selection in high dimensional model classes[1]. The CIPM priors are an implementation of the conjugate informative proper priors in Pacifico (2020c) when studying DPD with correlated random effects. Markov Chain Monte Carlo (MCMC) algorithms and implementations are used to design posterior distributions and then perform cross-country forecasts and policy issues. Theoretically, ratio-optimality and posterior consistency are also investigated modelling heterogeneous parameters for better individual–specific forecasts.

The contributions of this paper are fourfold. First, I develop a hierarchical structural Bayes approach to deal with potential features in real-world data such as non-linearity, incidental parameters (because of individual–specific heterogeneity), endogeneity issues (because of omitted factors and unobserved heterogeneity), and structural model uncertainty[2] (because of one or more parameters are posited as the source of model misspecification problems). CIPM priors and MCMC-based Posterior Model Probabilities (PMPs) are used in oder to: (*i*) include all of the information from the whole panel, acting as a strong model selection in high dimensional model classes; (*ii*) impose choice and specification strategy of the informative priors concerning the outcomes of interest and the distribution of unobserved heterogeneous effects; and (*iii*) deal with overfitting[3] and model uncertainty when addressing variable selection problems.

Second, I build on and implement the Pacifico (2020c)'s analysis, who develops a Robust Open Bayesian (ROB) procedure in two stages for implementing Bayesian Model Averaging (BMA) and Bayesian Model Selection (BMS) in multiple linear regression models when accounting for dynamics of the economy in either time-invariant moderate data or time-varying high dimensional multivariate data. More precisely, I

[1] In Bayesian statistics, variable and model selection procedures are performed to deal with the complexity of the model, where the 'complexity' stands (for example) for the number of unknown parameters.

[2] See, for instance, Gelfand and Dey (1994) for related works.

[3] It refers to the overestimation of effect sizes since more complex models always provide a somewhat better fit to the data than simpler models, where the 'complexity' stands – for example – for the number of unknown parameters.

implement the prior specification strategy in multivariate dynamic panel data in order to make inference on multivariate high dimensional panel setups and then obtain a reduced subset containing the *best*[4] combination of predictors (or the *best* model solution) that mainly explain and thus fit the data. In the second stage, further shrinkage is performed in order to obtain a smallest final subset of *top best* submodels containing the only *significant* solutions[5]. Finally, the submodel with the highest Bayes Factor in logarithm (lBF) would correspond to the final solution containing a potential subset of candidate predictors with higher *significant* overall F value and sufficiently *strong* adjusted-R^2 ($\bar{R}^2$) measure, where *strong* refers to $\bar{R}^2$ value equal to or bigger than 30%. However, interdependencies and inter-linkages – across units and time periods – matter and would strongly affect all (potential) covariates contained in the reduced parameter space. Thus, the main nolvelty of that implementation– taking the name of Multivariate Panel ROB (MPROB) procedure – consists of adding an additional stage to select a smaller pool of *top best* candidate predictors specifying the final solution model. More precisely, I perform a further shrinkage based on the Granger (Non-)Causality test in multivariate dynamic panel data (see, for instance, Dumitrescu and Hurlin (2012)). The idea is to exclude the predictors when no causal link holds across units within the panel (homogeneity under the null hypothesis); conversely, whether highly strong causal links matter for a subgroup of units (heterogeneity under the alternative), the same parameters should be taken into account in order to deal with overestimation of effect sizes (or individual contributions). In this study, the optimal lag length testing Granger-causality is set using the Arellano's test (see, for instance, Arellano (2003) and Arellano and Honore (2001)).

Third, MCMC algorithms and implementations are addressed to account for relative regrets dealing with semiparametric forecasting problem. Better evidence-based forecasting is involved in DPB-CRE because of two main features: the use of a Bayesian hierarchical approach with informative mixture priors and correlated random coefficients.

Fourth, I also build on and implement the Arellano and Bond (1991)'s strategy, where lagged-based values of the instrumented variables of interest are included within the system as *internal* necessary instruments[6], but with a novelty. More precisely, the DPB-CRE allows the inclusion of *external* instruments. Here, the *instruments* refer to univariate processes and correspond to all available lags of the *top best* candidate predictors obtained in the second stage; *external*, because of all lagged parameters are included before the estimation method, but after the MPROB procedure. Finally, a CRE approach is used in which the unobserved individual heterogeneities are treated as random variables that are possibly correlated with some of the predictors within the system. In this way, possible biases in the estimated coefficients of lagged outcomes will be avoided as well.

An empirical application on a pool of advanced and emerging economies is assessed describing the functioning and the performance of the methodology. It aims to identify and analyze a set of potential

[4]In BMA and BMS, *best* stands for the model providing the most accurate predictive performance over all candidate models.

[5]Here, *top best* stands for the model providing the most accurate predictive performance over all candidate submodels obtained in the first stage, and *significant* stands for models having statistically significant predictive capability.

[6]See, for instance, Arellano (2003), Arellano and Bonhomme (2011), and Arellano and Hahn (2016) for some other relevant applications.

socioeconomic–demographic factors, policy tools, and economic–financial issues during the pandemic crisis. The estimation sample refers to the period $1990 - 2020$, covering a sufficiently large sample to address possible causal links and interdependency between variables of interest (e.g., outcomes of economic growth), predetermined variables[7] (e.g., lagged values of the outcomes and control variables), strictly exogenous factors (e.g., dummy variables capturing structural effects), directly observed (endogenous) variables (e.g., socioeconomic–demographic and economic–financial factors), and time-invariant effects (e.g., heterogeneous individual-specific parameters possibly correlated with potential predictors within the system). Furthermore, the empirical strategy is also able to investigate and thus design better forecasts and strategic policy measures to contain the socioeconomic challenges of COVID-19 pandemic and ensure more resilient and robust health systems safeguarding against future epidemic diseases.

A simulated experiment – compared to related works – is also addressed to highlight the performance of the estimating procedure developed in this study using some Monte Carlo simulations.

The remainder of this paper is organized as follows. Section 2 discusses related works. Section 3 introduces the econometric model and the estimating procedure. Section 4 displays prior specification strategy and posterior distributions accounting for FMM-based Empirical Bayes estimator (Tweedie Correction), ratio-optimality, and theoretical properties. Section 5 describes the data and the empirical analysis. Section 6 presents the simulated experiment dealing with relative regrets for Tweedie Correction through Monte Carlo algorithms. The final section contains some concluding remarks.

2 Related Literature

This paper is related to several strands of the literature in dynamic panel setups. As regards the frequentist literature, closely related studies addressing similar deconvolution problem and estimates of μ_i's distribution are Anderson and Hsiao (1981), Arellano and Bond (1991), Arellano and Bover (1995), Blundell and Bond (1998), and Alvarez and Arellano (2003) (Instrumental Variables (IV) and Generalized Method of Moments (GMM) estimators); Hahn and Newey (2004), Carro (2007), Arellano and Hahn (2007, 2016), Bester and Hansen (2009), Fernandez-Val (2009), and Hahn and Kuersteiner (2011) (fixed effects approach in non-linear panel data); and Compiani and Kitamura (2016) (mixture models-based approach). However, the frequentist approach is not able to deal with model uncertainty and overfitting in performing individual–specific forecasts in high dimensional data.

Earlier works regarding empirical Bayes methods with parametric priors on heterogeneous parameters refer to Robbins (1964), Robert (1994), Brown and Greenshtein (2009), and Jiang and Zhang (2009), and – more recently – Liu et al. (2019, 2020) (henceforth LMS) and Gu and Koenker (2017a,b) (henceforth GK). LMS aim to forecast a collection of short time-series using cross-sectional information. Then,

[7]In econometrics, predetermined variables denote covariates uncorrelated with contemporaneous errors, but not for their past and future values.

they construct point forecasts predictors using Tweedie's formula[8] for the posterior mean of heterogeneous individual–specific factors under a correlated random effects distribution. They show that the ratio optimality of point forecasts asymptotically converge to the one based on a nonparametric kernel estimate of the Tweedie correction. However, they replace the μ_i's distribution with a kernel density estimator performing less accurate forecasts than alternative estimates of the Tweedie correction (e.g., nonparametric maximum likelihood estimation and finite mixture of normal distributions). Then, they estimate relative regrets for these two alternative Tweedie corrections via Markov chains simulations, but specifying bounds for the domain of μ_i and partitioning it into default setting bins. It would compromise the estimates because of weak empirical forecast optimality limited to restrictive and constrained classes of models. GK use Tweedie's formula to construct an approximation to the posterior mean of the heterogeneous parameters. They build on Kiefer and Wolfowitz (1956) and implement the empirical Bayes predictor based on a nonparametric maximum likelihood estimator of the cross-sectional distribution of the sufficient statistics. However, no theoretical optimality results are provided.

The methodology proposed in this paper makes three contributions: (*i*) correlated random effects are allowed for cross-sectional heterogeneity interacting with the initial conditions; (*ii*) a MPROB procedure is used in a dynamic panel setup to deal with some open issues related to BMA for multiple competing models classes in high dimension (such as overfitting, model uncertainty, endogeneity issues, and model misspecification problems); and (*iii*) a FMM distribution – in accordance with the MPROB procedure – is used to minimize relative regrets for a Tweedie Correction performing better forecasts and policy strategies.

Finally, this paper is also related to several frequentist statistical, dynamic, and multicountry approaches concerning the current COVID-19 pandemic crisis, investigated in details in Section 5.1.

3 Empirical Model and Estimation Procedure

3.1 Multivariate Dynamic Panel Data

The baseline Dynamic Panel Data (DPD) model is:

$$y_{it} = \beta_l y_{i,t-l} + \alpha x_{it} + \gamma_l z_{i,t-l} + \mu_i + u_{it} \tag{2}$$

where the subscripts $i = 1, 2, \ldots, N$ are country indices, $t = 1, 2, \ldots, T$ denotes time, y_{it} is a $N \cdot 1$ vector of outcomes, $y_{i,t-l}$ and $z_{i,t-l}$ are $N \cdot 1$ vectors of predetermined and directly observed (endogenous) variables for each i, respectively, with $l = 0, 1, 2, \ldots, \lambda$, $\beta_{\tilde{l}}$ and $\gamma_{\tilde{l}}$ are the autoregressive coefficients to be estimated for each i, with $\tilde{l} = 1, \ldots, \lambda$, $x_{i,t}$ is a $N \cdot 1$ vector of strictly exogenous factors for each i, with α denoting the regression coefficients to be estimated, μ_i is a $N \cdot 1$ heterogeneous intercept containing – for example

[8]The formula is attributed to the astronomer Arthur Eddington and the statistician Maurice Tweedie.

– time-constant differences (such as territorial competitiveness, infrastructural system, competitiveness developments, macroeconomic imbalances), and $u_{it} \sim i.i.d.N(0, \sigma_u^2)$ is a $N \cdot 1$ vector of unpredictable shock (or idiosyncratic error term), with $E(u_{it}) = 0$ and $E(u_{it} \cdot u_{js}) = \sigma_u^2$ if $i = j$ and $t = s$, and $E(u_{it} \cdot u_{js}) = 0$ otherwise. In this study, I consider the same lag order (or optimal lag length) for both predetermined $(y_{i,t-l})$ and observed variables $(z_{i,t-l})$.

Here, some considerations are in order: (i) the predetermined variables contain the lagged values of the outcomes y_{it}, capturing – for example – the persistence, and control variables; (ii) the μ_i's denote cross-sectional heterogeneity affecting the outcomes y_{it}; (iii) correlated random effects matter and then μ_i's are treated as random variables and possibly correlated with some of the covariates within the system; (iv) the roots of $\tilde{l}(L) = 0$ lie outside the unit circle so that the AR processes involved in the model (2) are stationaries, with L denoting the lag operator; (v) the strictly exogenous factors x_{it} contain dummy variables to test – for example – the presence of structural breaks or policy shifts, and (vi) the instruments are fitted values from AutoRegressive (AR) parameters based on all of the available lags of the time-varying variables. In this study, the optimal lag length and the order of integration have been set using the Arellano's test and the Augmented Dickey-Fuller (ADF) test for each i, respectively.

Let the stationarity hold in (2), the time-series regressions are valid and the estimates feasible. Thus, moment restrictions[9] need to hold in order to address exact identification in a context of correlated random effects and estimate $\beta_{\tilde{l}}$ and $\gamma_{\tilde{l}}$ for $T \geq 3$. More precisely, I assume that μ_i and $u_{i,t}$ are independently distributed across i and have the familiar error components structure:

$$E(\mu_i) = 0, E(u_{it}) = 0, E(u_{it} \cdot \mu_i) = 0 \qquad for \quad i = 1, \dots, N \quad and \quad t = 2, \dots T \tag{3}$$

and

$$E(u_{it} \cdot u_{is}) = 0 \qquad for \quad i = 1, \dots, N \quad and \quad t \neq s \tag{4}$$

Then, I also assume the standard assumption concerning the initial conditions $y_{i,t=1}$:

$$E(y_{i,t=1} \cdot u_{it}) = 0 \qquad for \quad i = 1, \dots, N \quad and \quad t = 2, \dots T \tag{5}$$

3.2 Bayesian Analysis

The main thrust of MPROB procedure in multivariate DPB-CRE in (2) is threefold. First, it provides for the *best* model solution (or combination of predictors) better explaining and thus fitting the data among high dimensional panel setups. It is very useful when studying causality and interdependency between different

[9]See, e.g., Anderson and Hsiao (1981), Arellano and Honore (2001), and Blundell and Bond (1998).

events affecting outcomes. Second, the use of CIPM priors allows for shrinking the dataset via an `ad-hoc` model selection since the common and heterogeneous coefficients (β_l, α, γ_l) and their distribution change in a corresponding fashion in accordance with different model solutions. Third, better individual–specific forecast can be performed assigning more weight according to model size so as to deal with overfitting and model uncertainty.

In this study, forecasts account for good and consistent estimates of both common and heterogeneous coefficients (β_l, α, γ_l) in the presence of the incidental parameters (μ_i) with $N \to \infty$ and large fixed T. More precisely, forecasts are based on the knowledge of the common parameters (θ, σ_u^2) and the distribution $\pi(\mu_i|\cdot)$ of the heterogeneous factors μ_i, but not the values μ_i themselves. It takes the name of oracle forecast and replaces μ_i by its posterior mean under the prior distribution $\pi(\mu_i|\cdot)$. Thus, neither the common parameters nor the distribution of the individual–specific coefficients are known. Here, I follow two steps: firstly, I replace the unknown common parameters by a consistent estimator and then I use the Tweedie's formula[10] that involves in evaluating the posterior mean of μ_i through a function of the cross-sectional density of certain sufficient statistics rather than through the likelihood function and an estimate of $\pi(\mu_i|\cdot)$. This density is estimated from the whole cross-sectional information by using an EB estimate of μ_i and an EB predictor of the optimal forecast of y_{it} at time T $(y_{i,T+k})$, with k denoting the k-step-ahead forecast. The main difference between an empirical and fully Bayesian approach is that the former picks the μ_i distribution by maximizing the Maximum Likelihood (ML) of the data[11], whereas a fully Bayesian method constructs a prior for the correlated random effects and then evaluates it in view of the observed panel data[12]. Even if the fully Bayesian approach tends to be more suitable for density forecasting and more easily extended to non-linear case, it would be a lot more computationally intensive. In this study, I implement the EB predictor by maximizing the log likelihood function and Tweedie correction using an Expectation-Maximization (EM) algorithm. The information is uploaded from the whole cross-section via a strong model selection implicit in the MPROB procedure, where the forecast evaluation criterion is the Mean Squared Error (MSE) computed across countries.

Given the DPB-CRE in (2), I decompose the vectors of the observed endogenous variables: $y_{i,t-l} = \left[y^{o'}_{i,t-l}, y^{c'}_{i,t-l}\right]'$, with $y^{o'}_{i,t-l}$ denoting lagged outcomes to capture the persistence and $y^{c'}_{i,t-l}$ including lagged control variables such as general economic conditions; and $z_{i,t-l} = \left[z^{s'}_{i,t-l}, z^{p'}_{i,t-l}\right]'$, referring to other lagged factors such as socioeconomic conditions $(z^{s'}_{i,t-l})$ and policy implications $(z^{p'}_{i,t-l})$. Then, I combine the (non-)homogeneous parameters into the vector $\theta = \left(\beta^{o'}_l, \beta^{c'}_l, \alpha', \gamma^{s'}_l, \gamma^{p'}_l\right)'$.

In order to model the key latent heterogeneities (μ_i) and observed determinants ($y_{i,t-l}$, x_{it}, $z_{i,t-l}$) when dealing with high dimensional analysis, I define the conditioning set at period t (c_{it}) and the structural density ($D(y_{it}|\cdot)$) as:

[10]See, e.g., Brown and Greenshtein (2009), Efron (2011), and Gu and Koenker (2017b) for some applications in big data analytics.

[11]See, e.g., Chamberlain and Hirano (1999), Hirano (2002), Lancaster (2002), Jiang and Zhang (2009), and Gu and Koenker (2017a,b) concerning some studies on the empirical Bayes methods in dynamic panel data models.

[12]See, for instance, Liu (2018) and Liu et al. (2020) (linear case); and Liu et al. (2019) (non-linear case).

$$c_{it} = \Big(y^{o}_{i,0:t-l}, y^{c}_{i,0:t-l}, z^{s}_{i,0:t-l}, z^{p}_{i,0:t-l}, x_{i,0:t}\Big) \tag{6}$$

and

$$D\Big(y_{it}|y_{i,t-l}, x_{it}, z_{i,t-l}, \mu_i\Big) = D\Big(y_{it}|y_{i,t-l}, x_{it}, z_{i,t-l}, y_{i0}, \mu_i\Big) \tag{7}$$

The error terms (u_{it}) are individual-time-specific shocks characterized by zero mean and homoskedastic Gaussian innovations. In a unified and hierarchical framework, I combine the individual heterogeneity into the vector $\phi_i = \Big(\mu_i, \sigma^2_u\Big)$ under cross-sectional heterogeneity and homoskedasticity. Assuming correlated random coefficients model, ϕ_i and c_{i0} could be correlated with each other, with:

$$c_{i0} = \Big(y^{o}_{i,0}, y^{c}_{i,0}, z^{s}_{i,0}, z^{p}_{i,0}, x_{i,0:T}\Big) \tag{8}$$

Given these primary specifications, the DPB-CRE model in (2) would be less parsimonious and harder to implement. Thus, I adopt an EB approach using cross-sectional information to estimate the prior distribution of the correlated random effects and then the conditions on these estimates. Moreover, it is not necessary to include all initial values of the outcomes ($y^{o}_{i,0}$), the control ($y^{c}_{i,0}$), the socioeconomic ($z^{s}_{i,0}$), and the policy ($z^{p}_{i,0}$) factors. Indeed, according to the MPROB procedure, only a subset of c_{i0} – relevant for the analysis – will be accounted for obtaining an advantage of highly larger feasibility.

Let $\mathcal{F}$ be the full panel set containing all (potential) model solutions, the variable selection problem is addressed by imposing an auxiliary indicator variable χ_h, with $h = 1, 2, \ldots, m$, containing every possible 2^m subset choices, where $\chi_h = 0$ if θ_h is small (absence of h-th covariate in the model) and $\chi_h = 1$ if θ_h is sufficiently large (presence of h-th covariate in the model). According to the Pacifico (2020c)'s framework, I run the MPROB procedure by matching all potential candidate models to shrink both the model space and the parameter space. The shrinking jointly deals with overestimation of effect sizes (or individual contributions) and model uncertainty (implicit in the procedure) by using Posterior Model Probabilities (PMPs) for every candidate models[13]. It can be defined as:

$$\pi\Big(y|\theta_h\Big) = \int_{\mathcal{B}} \pi\Big(y, \mu_i|\theta_h, M_h\Big) \cdot d\mu \tag{9}$$

where $\mathcal{B}$ denotes the multidimensional (natural) parameter space for θ_h, $M_h = (M_1, \ldots, M_m)$ denotes a countable collection of all (potential) model solutions given the data. The integrand in (9) is defined as:

[13] The PMP denotes the probability of each candidate model performing the data.

$$\int_{\mathcal{B}} \pi\Big(y, \mu_i | \theta_h, M_h\Big) = \pi\Big(\theta_h, \mu_i, M_h | y\Big) \cdot \pi\Big(y | M_h\Big) \tag{10}$$

where $\pi(\theta_h, \mu_i, M_h|y)$ denotes the joint likelihood and $\pi(y|M_h) = \int \pi\Big(y|M_h, \theta_h, \mu_i\Big) \cdot \pi\Big(\theta_h, \mu_i|M_h\Big) d\theta_h$ is the marginal likelihood, with $\pi(\theta_h, \mu_i|M_h)$ referring to the conditional prior distribution of θ_h and μ_i. In this first stage, with N high dimensional and T sufficiently large, the calculation of the integral $\pi(y|M_h)$ in unfeasible and then a fully enumerated Markov Chain Monte Carlo (MC^F) implementation is conducted[14].

The subset containing the *best* model solutions will correspond to:

$$\mathcal{S} = \Big\{ M_j : M_j \subset \mathcal{S}, \mathcal{S} \in \mathcal{F}, \Theta_j \subset \Theta_h, \ \sum_{j=1}^{\varpi} \pi\Big(M_j | y_i = y_i, \chi\Big) \geq \tau \Big\} \tag{11}$$

where M_j denotes the submodel solutions of the DPB-CRE in (2), with $M_j < M_h$, $j \ll h$, $\{1 \leq j < h\}$, and τ is a threshold chosen arbitrarily for an enough posterior consistency[15]. In this study, I use $\tau = 0.5\%$ with N high dimensional (predictors ≥ 15). In this study, I am able to jointly manage all equations within the system (through the conditioning set c_{it}), their (potential) interactions (through AR coefficients), and their possible causal links (through Granger (Non-)Causality test).

The second stage consists of reducing the model space $\mathcal{S}$ to obtain a smaller subset of *top best* submodel solutions:

$$\mathcal{E} = \Big\{ M_\xi : M_\xi \subset \mathcal{E}, \mathcal{E} \in \mathcal{S}, \ \sum_{j=1}^{\varpi} \pi\Big(M_j | y_i = y_i, \dot{\chi}\Big) \geq \dot{\tau} \Big\} \tag{12}$$

where $M_\xi \ll M_j$, $\pi(M_j|y_i = y_i, \dot{\chi})$ denotes the PMPs, with $\dot{\chi}$ denoting a new auxiliary variable containing the only *best* model solutions in the subset $\mathcal{S}$ and $\dot{\tau}$ referring to a new arbitrary threshold to evaluate the probability of the model solutions in $\mathcal{S}$ performing the data (PMPs). In this study, I still use $\tau = 0.5\%$ – independently of N – for a sufficient prediction accuracy in explaining the data.

The MPROB procedure comes to a conclusion once a further shrinkage – based on the panel Granger (Non-)Causality test[16] – is conducted to obtain the smallest final subset of *top best* submodel solutions ($M_{\xi^*} \subset \mathcal{E}$). More precisely, that stage consists of including the only candidate predictors displaying highly strong causal links for at least a subgroup of units (heterogeneity under the alternative) with p-value $\leq \dot{\tau}$. To deal with endogeneity issues and misspecified dynamics, all available lags of the *top best* candidate predictors – obtained in the second stage – are included as *external* instruments. In this study, the optimal lag length testing Granger-causality is set using the Arellano's test[17].

The final model solution to be considered performing forecasting and policy-making will correspond to

[14] MC^F integration is used to move through the model space and the parameter space at the same time in order to obtain a reduced set containing the best combination of predictors. See, for instance, Pacifico (2020c) in linear static case.

[15] In Bayesian analysis, posterior concistency ensures that the posterior probability (PMP) concentrates on the true model.

[16] See, for instance, Dumitrescu and Hurlin (2012).

[17] See, for instance, Arellano (2003) and Arellano and Honore (2001).

one of the submodels M_{ξ^*} with higher log natural Bayes Factor (lBF):

$$lBF_{\xi^*,\xi} = log\left\{\frac{\pi(M_{\xi^*}|y_i = y_i)}{\pi(M_\xi|y_i = y_i)}\right\} \tag{13}$$

In this analysis, the lBF is interpreted according to the scale evidence in Pacifico (2020c), but with more stringent conditions:

$$\begin{cases} 0 < lB_{\xi^*,\xi} \leq 5 & \text{no evidence for submodel } M_{\xi^*} \\ 6 < lB_{\xi^*,\xi} \leq 10 & \text{moderate evidence for submodel } M_{\xi^*} \\ 11 < lB_{\xi^*,\xi} \leq 15 & \text{strong evidence for submodel } M_{\xi^*} \\ lB_{\xi^*,\xi} > 15 & \text{very strong evidence for submodel } M_{\xi^*} \end{cases} \tag{14}$$

4 Hierarchical Framework and Empirical Bayes Approach

4.1 Prior Specification Strategy and Tweedie's Formula

According to the EB approach, the variable selection procedure entails estimating χ_h and θ_h as posterior means (the probability that a variable is *in* the model). All observal variables in c_{it} and individual heterogeneity in ϕ_i are hierarchically modelled via proper conjugate informative mixture priors:

$$\pi(\theta, \phi, \chi) = \pi(\theta|\chi) \cdot \pi(\mu_i|\chi, y_{i0}) \cdot \pi(\sigma_u^2|\chi) \cdot \pi(\chi) \tag{15}$$

where

$$\pi\Big(\theta|\mathfrak{F}_{-1}\Big) = N\Big(\bar{\theta}, \bar{\rho}\Big) \tag{16}$$

$$\pi(\mu_i|\theta) = N\Big(\delta_{\mu_i}, \Psi_{\mu_i}\Big) \quad \texttt{with} \quad \delta_{\mu_i} \sim N\Big(0, \zeta\Big) \quad \texttt{and} \quad \Psi_{\mu_i} \sim IG\Big(\frac{\varphi}{2}, \frac{\varepsilon}{2}\Big) \tag{17}$$

$$\pi(y_{i0}|\mu_i) = N(0, \kappa) \tag{18}$$

$$\pi(\chi) = w_{|\chi|} \cdot \binom{h}{|\chi|}^{-1} \tag{19}$$

$$\pi(\sigma_u^2) = IG\Big(\frac{\bar{\omega}}{2}, \frac{\nu}{2}\Big) \tag{20}$$

where $N(\cdot)$ and $IG(\cdot)$ stand for Normal and Inverse-Gamma distribution, respectively, $\mathfrak{F}_{-1}$ refers to the cross-sectional information available at time -1, and $w_{|\chi|}$ in (19) denotes the model prior choice related to the sum of the PMPs (or Prior Inclusion Probabilities) with respect to the model size $|\chi|$, through which the θ's will require a non-0 estimate or the χ's should be included in the model. In this way, one would weight more according to model size and – setting $w_{|\chi|}$ large for smaller $|\chi|$ – assign more weight to parsimonious models.

All hyperparameters are known. More precisely, collecting them in a vector $\tilde{\omega}$, where $\tilde{\omega} = \Big(\bar{\theta}, \bar{\rho}, \zeta, \varphi, \varepsilon, \kappa, w_{|\chi|}, \bar{\omega}, \nu\Big)$, they are treated as fixed and are either obtained from the data to tune the prior to the specific applications (such as $\bar{\theta}, \varphi, \kappa, w_{|\chi|}, \bar{\omega}$) or selected a priori to produce relatively loose priors (such as $\bar{\rho}, \zeta, \varepsilon, \nu$). Here, $w_{|\chi|}$ is restricted to a benchmark prior $max\Big(NT, |\chi|\Big)$ according to the non-0 components of χ.

Nevertheless, to accomodate the correlated random coefficients model where the individual–specific heterogeneity (μ_i) can be correlated with the conditioning variables c_{i0} and y_{i0}, I use an empirical Bayes procedure where the posterior mean of μ_i is expressed in terms of the marginal distribution of a sufficient statistic ($\hat{\mu}_i(\theta)$) estimated from the cross-sectional whole information (Tweedie's formula). Given the CIPM priors in (16) - (20), I define the compound risk and loss functions – under which the forecasts are evaluated – accounting for expectations over the observed trajectories $\mathcal{Y}_i = \Big(y_1^{0:T}, \ldots, y_N^{0:T}\Big)$, with $y_i^{0:T} = \Big(y_{i0}, y_{i1}, \ldots, y_{iT}\Big)$, the unobserved heterogeneity ($\mu_i = \mu_1, \ldots, \mu_N$), and the future shocks $u_{i,T+k} = \Big(u_{1,T+k}, \ldots, u_{N,T+k}\Big)$:

$$\mathcal{R}\Big(\hat{y}_{i,T+k}\Big) = \mathbb{E}_{\theta,\phi,\pi(\cdot)}^{(\mathcal{Y}_N,\mu_i,u_{i,T+k})}\Big[L_N\Big(\hat{y}_{i,T+k}, y_{i,T+k}\Big)\Big] \tag{21}$$

where $L_N\Big(\hat{y}_{i,T+k}, y_{i,T+k}\Big) = \sum_{i=1}^{N}\Big(\hat{y}_{i,T+k} - y_{i,T+k}\Big)^2$ denotes the compound loss obtained by summing over the units i the forecast error losses $(\hat{y}_{i,T+k} - y_{i,T+k})$, with $\hat{y}_{i,T+k} = (\hat{y}_{1,T+k}, \ldots, \hat{y}_{N,T+k})'$ is a vector of k-period-ahead forecasts.

In the compound decision theory, the infeasible oracle forecast (or benchmark forecast) implies that ϕ_i and the distribution of the unobserved heterogeneity ($\pi(\mu_i, y_{i0})$) are known, the trajectories ($\mathcal{Y}_i$) are observed, and the values of μ_i are unknown across units i. Moreover, the integrated risk in (21) is minimized performing individual–specific forecasting that minimizes the posterior risk for each $\mathcal{Y}_i$. Thus, according to the Liu et al. (2020)'s framework, the posterior risk can be defined as:

$$\mathbb{E}_{\theta,\phi,\pi(\cdot)}^{(\mathcal{Y}_N,\mu_i,u_{i,T+k})}\Big[L_N\Big(\hat{y}_{i,T+k}, y_{i,T+k}\Big)\Big] = \sum_{i=1}^{N}\Bigg\{\Big(\hat{y}_{i,T+k} - \mathbb{E}_{\theta,\phi,\pi(\cdot)}^{(\mathcal{Y}_i,\mu_i,u_{i,T+k})}[y_{i,T+k}]\Big)^2 + \mathbb{V}_{\theta,\phi,\pi(\cdot)}^{(\mathcal{Y}_i,\mu_i,u_{i,T+k})}[y_{i,T+k}]\Bigg\} \tag{22}$$

where $\mathbb{V}_{\theta,\phi,\pi(\cdot)}^{(\mathcal{Y}_i,\mu_i,u_{i,T+k})}[y_{i,T+k}]$ is the posterior predictive variance of $y_{i,T+k}$. The optimal predictor would be the mean of the posterior predictive distribution:

$$\hat{y}^{op}_{i,T+k} = \mathbb{E}^{(\mathcal{Y}_i,\mu_i,u_{i,T+k})}_{\theta,\phi,\pi(\cdot)}[y_{i,T+k}] = \mathbb{E}^{(\mathcal{Y}_i,\mu_i)}_{\theta,\phi,\pi(\cdot)}[\mu_i] + (\theta \cdot c_{it}) \tag{23}$$

where the acronym OP stands for 'Optimal'. Then, the compound risk in (21) associated with the infeasible oracle forecast can be rewritten as:

$$\mathcal{R}^{op} = \mathbb{E}^{(\mathcal{Y}_i,\mu_i,u_{i,T+k})}_{\theta,\phi,\pi(\cdot)}\left\{\sum_{i=1}^{N}\left(\mathbb{V}^{(\mathcal{Y}_i,\mu_i)}_{\theta,\phi,\pi(\cdot)}[\mu_i] + \sigma^2_u\right)\right\} \tag{24}$$

The optimal compound risk in (24) consists of two components: uncertainty concerning the individual–specific heterogeneity on the observations i and uncertainty with respect to the error terms. Because of infeasible benchmark forecast, the parameter vectors (θ, ϕ) and the CRE distribution ($\pi(\cdot)$) are unknown. Thus, the posterior mean $\mathbb{E}^{(\mathcal{Y}_i,\mu_i)}_{\theta,\phi,\pi(\cdot)}[\mu_i]$ in (23) is assessed through the Tweedie's formula by evaluating the marginal distribution of a sufficient statistic of the heterogeneous effects. The likelihood function associated with the multivariate DPB-CRE in (2) is:

$$\pi\left(y^{1:T}_i|y_{i0},\mu_i,\theta\right) \propto exp\left\{-\frac{1}{2\sigma^2_u}\sum_{t=1}^{T}\left(y_{it} - (c_{it-l}|\chi)\theta_t - \mu_i(\theta)\right)^2\right\} \propto \left\{-\frac{T}{2\sigma^2_u}\left(\hat{\mu}_i(\theta) - \mu_i\right)^2\right\} \tag{25}$$

where $\hat{\mu}_i(\theta)$ denotes the sufficient statistic and equals to:

$$\hat{\mu}_i(\theta) = \frac{1}{T}\sum_{t=1}^{T}\left(y_{it} - (\theta \cdot c_{it-l})\right) \tag{26}$$

According to Bayes's theorem, the posterior distribution of μ_i can be obtained as:

$$\pi\left(\mu_i|y^{0:T}_i,\theta\right) = \pi\left(\mu_i|\hat{\mu}_i,y_{i0},\theta\right) = \frac{\pi\left(\hat{\mu}_i|\mu_i,y_{i0},\theta\right)\cdot\pi\left(\mu_i|y_{i0}\right)}{exp\left\{ln\left(\pi(\hat{\mu}_i|y_{i0})\right)\right\}} \tag{27}$$

The last step to obtain the Tweedie's formula is to differentiate the equation $\pi\left(\mu_i|\hat{\mu}_i,y_{i0},\theta\right)$ in (27) with respect to $\hat{\mu}_i$ and solve the equation for the posterior mean $\mathbb{E}^{(\mathcal{Y}_i,\mu_i)}_{\theta,\phi,\pi(\cdot)}[\mu_i]$ in (23). Thus, the Tweedie's formula equals to:

$$\mathbb{E}^{(\mathcal{Y}_i,\mu_i)}_{\theta,\phi,\pi(\cdot)}[\mu_i] = \hat{\mu}_i(\theta) + \frac{\sigma^2_u}{T}\cdot\frac{\partial}{\partial\hat{\mu}_i(\theta)}ln\left(\hat{\mu}_i(\theta),y_{i0}\right) \tag{28}$$

where the second term in (28) denotes the correction term capturing heterogeneous effects of $\pi(\cdot)$ – the prior of μ_i – on the posterior. It is expressed as a function of the marginal density of $\hat{\mu}_i(\theta)$ conditional on y_{i0} and θ; contrarily to the full Bayesian approach, where one needs to avoid the deconvolution problem that disentangle the prior density $\pi(\mu_i|y_{i0})$ from the distribution of the error terms (u_{it}).

4.2 Tweedie Correction and Markov Chain Algorithms

The infeasible oracle forecast described in Section (4.1) is approximated through an EB method in order to substitute the unknows parameters θ and the joint distribution between the μ_i's sufficient statistic and individual outcome values $\pi\Big(\hat{\mu}_i(\theta), y_{i0}\Big)$ in (28) by estimates, uploading the cross-sectional information set into $\mathcal{E}$ (Tweedie correction).

In dynamic panel data, consistent estimates of the unkown parameters θ can be obtaining through Genelarized Method of Moments (GMM) estimators. In this study, they correspond to the AR(λ) coefficients related to predetermined and endogenous variables[18]. Let the stationarity and moment conditions in (3)-(5) hold in the system, the time-series regressions are valid (or computational) and GMM estimators are feasible. Concerning the density $\pi\Big(\hat{\mu}_i(\theta), y_{i0}\Big)$, I estimate it by using FMM distributions:

$$\pi_{mix}\Big(\hat{\mu}_i, y_{i0} \mid |\chi|, c_{i0}\Big) = |\chi| \cdot \pi_{\xi}\Big(\hat{\mu}_i, y_{i0} \mid c_{i0}\Big) \quad \texttt{with} \quad |\chi| > 0 \tag{29}$$

where $\pi_{\xi}(\cdot)$ is the conditional density distribution of heterogeneous effects with sample size $|\chi|$. In this way, I would able to account for the whole cross-sectional information so as to get estimates of (non-)homogenous parameters θ (`first stage`) and density $\pi_{\xi}(\cdot)$ (`second stage`). Here, I focus on the only *best* submodels achieved in the second stage of the MPROB procedure in order to work with sufficiently high posterior consistency.

The FMM distributions and their moments themselves (means and covariance matrices) are evaluated by maximizing the log likelihood function via an EM algorithm. More precisely, I suppose $\bar{m}$ regimes in which heterogeneous effects (ϕ) can vary in each submodel solution, where $\bar{m} = 0, 1, \ldots, \iota$ is close to $|\chi|$, with 0 indicating the uninformative model where heterogeneous effects do not affect outcomes (e.g., DPD with fixed effects), and $\bar{m} \subset \mathcal{E}$. Then, I use Metropolis-Hastings algorithm[19] to draw posteriors for $\hat{\mu}_i$ from the (proposal) joint density distribution $\pi^{\bar{m}} = |\chi| \cdot \pi^*_{\xi}\Big(\hat{\mu}_i^{\bar{m}}, y_{i0}^{\bar{m}} \mid c_{i0}^{\bar{m}}\Big)$, with probability $\mathfrak{p}_{\bar{m}}$ equals to:

$$\mathfrak{p}_{\bar{m}} = \frac{\pi\Big(\hat{\mu}_i^{\bar{m}}, y_{i0}^{\bar{m}} \mid \hat{\mu}_i^{\bar{m}-l}, \mathcal{Y}_i, \Big\{\theta_t\Big\}_{t=1}^{T}, c_{i0}^{\bar{m}}\Big) \cdot \pi^{\bar{m}-l}}{\pi\Big(\hat{\mu}_i^{\bar{m}-l}, y_{i0}^{\bar{m}} \mid \hat{\mu}_i^{\bar{m}-l}, \mathcal{Y}_i, \Big\{\theta_t\Big\}_{t=1}^{T}, c_{i0}^{\bar{m}}\Big) \cdot \pi^{\bar{m}}} \tag{30}$$

where π^*_{ξ} stands for the conditional density distribution of heterogeneous effects involved in the final model solution (`third stage`).

Let $|\chi|^*$ be the sample size according to the uninformative model in which neither (non-)homogeneous parameters nor unobserved effects achieve sufficient posterior consistency and $\theta_t^* = \mathbf{1}^i$ be a vector of ones, the probability function takes the form:

[18] See, e.g., Arellano (2003), Arellano and Honore (2001), Arellano and Bover (1995), and Blundell and Bond (1998).
[19] See, for instance, Levine and Casella (2014) for implementations of Monte Carlo algorithm.

$$\pi\Big(\theta_t \mid \mathcal{Y}_i\Big) \cdot \pi^*(\theta_t^* \mid \theta_t) \cdot \mathfrak{p}(\theta_t^*, \theta_t) = \pi\Big(\theta_t^* \mid \mathcal{Y}_i\Big) \cdot \pi^*(\theta_t \mid \theta_t^*) \tag{31}$$

where

$$\mathfrak{p}(\theta_t^*, \theta_t) = \min \left[\frac{\pi(\theta_t^* \mid \mathcal{Y}_i) \cdot \pi^*(\theta_t \mid \theta_t^*)}{\pi(\theta_t \mid \mathcal{Y}_i) \cdot \pi^*(\theta_t^* \mid \theta_t)}, 1\right] \cong \mathfrak{p}_{\bar{m}} \tag{32}$$

with $\mathfrak{p}(\theta_t^*, \theta_t)$ displaying the probability to accept or reject a draw[20] and $\pi^*(\cdot)$ denoting the density distribution according to sample size $|\chi|^*$. In this way, I am able to get the same probability that each submodel M_ξ would be true. In addition, since posterior distributions corresponds – by construction – to a FMM distribution, I define three possible intervals – displayed in (33) – in which the posterior predictive variance of $\mu_i\Big(\mathbb{V}^{(\mathcal{Y}_i,\mu_i)}_{\theta,\phi,\pi(\bar{m})}[\mu_i]\Big)$ can vary according to the model size ($|\chi|$). In this way, I am able to obtain exact posteriors on the predictive variance of μ_i taking into account both the model space and the parameter space. Thus, running an `ad-hoc` model selection through a mixture of normal distributions, it ensures that lower variability would be associated to less relative regrets during the estimating procedure, achieving more accurate forecasts.

$$\begin{cases} \underset{(high\ dimension)}{0.5 < \mathbb{V}^{(\mathcal{Y}_i,\mu_i)}_{\theta,\phi,\pi(\bar{m})}[\mu_i] \leq 1.0} \quad \text{with} \quad \xi^* > 10 \quad (\ \texttt{heterogeneity}\ ) \\ \\ \underset{(moderate\ dimension)}{0.1 \leq \mathbb{V}^{(\mathcal{Y}_i,\mu_i)}_{\theta,\phi,\pi(\bar{m})}[\mu_i] \leq 0.5} \quad \text{with} \quad 5 < \xi^* \leq 10 \quad (\ \texttt{sufficient-homogeneity}\ ) \\ \\ \underset{(small\ dimension)}{0.0 \leq \mathbb{V}^{(\mathcal{Y}_i,\mu_i)}_{\theta,\phi,\pi(\bar{m})}[\mu_i] < 0.1} \quad \text{with} \quad \xi^* \leq 5 \quad (\ \texttt{near-homogeneity}\ ) \end{cases} \tag{33}$$

4.3 Ratio-Optimality and Posterior Distributions

To obtain the EB predictor to generate forecasts on (28) given (29), I define the ratio optimality (or optimal point forecasts). Usual claims of empirical forecast optimality are limited to restrictive classes of models and thus it would be very weak. Nevertheless, through the BMS implicit in (2), I am able to work on a restricted set of submodels well specified in order to investigate – given the data – better available forecast models. In this context, the optimal point forecasts' objective is to predict the outcomes (y_{it}) given the data by minimizing the expected loss in (24). Methodologically, it means proving that the predictor $\hat{y}_{i,T+k}$ achieves ϑ_0-ratio-optimality uniformly for priors $\pi^{\bar{m}} \subset \mathcal{E}$, with $\vartheta_0 \geq 0$. Thus,

[20] See, for instance, Jacquier et al. (1994) and Pacifico (2021) for some applications to multicountry and multidimensional time-varying panel setups with stochastic and time-varying volatility, respectively.

$$\limsup_{N\to\infty} \sup_{\pi^{\bar{m}}\subset\mathcal{E}} \frac{\mathcal{R}\Big(\hat{y}_{i,T+k},\pi^{\bar{m}}\Big)-\mathcal{R}^{op}\Big(\pi^{\bar{m}}\Big)}{\Big\{N\xi^{*}\cdot\mathbb{E}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi^{\bar{m}}}\Big(\mathbb{V}^{(\mathcal{Y}_i,\mu_i)}_{\theta,\phi,\pi^{\bar{m}}}\,[\mu_i]\Big)\Big\}+N(\xi^{*})\vartheta_0}\leq 0 \tag{34}$$

In (34), some considerations are in order. (i) First, the predictor $\hat{y}_{i,T+k}$ – defined in (21) – is constructed by replacing θ with a consistent estimator $\hat{\theta}$ (estimated AR(λ) coefficients) and individual outcome values $\pi\Big(\hat{\mu}_i(\theta),y_{i0}\Big)$ in (28) with estimates using FMM distributions in (29). (ii) Second, taking expectations over $y^{0:T}$ in (24), it follows that optimal point forecasts aim to work well on average – through the BMA implicit in the MPROB procedure – rather than for a particular value (or single draw) of the outcomes. More precisely, the individual–specific forecasts consist of estimating not the realization of the outcomes themselves, but rather a function of their predictive conditional distributions. Thus, the optimal forecasts themselves will be a function of θ and then anything more than parameters of the conditional distribution, $\pi\Big(\theta_t|\mathcal{Y}_i\Big)$, in (31). ($iii$) Third, the prediction accuracy of optimal forecasts can be assessed through the Mean Squared Errors $\Big(MSE(\hat{\theta})=E_{\hat{\theta}}\Big[\sum_{i=1}^{N}(\hat{y}_{i,T+k}-y_{i,T+k})^2\Big]\Big)$, calculated as the average of the squared forecast errors for all observations assigned to the model class M_{ξ^*}. For high $\mathbb{V}^{(\mathcal{Y}_i,\mu_i)}_{\theta,\phi,\pi(\bar{m})}$ (e.g., with $\xi^*>10$), the further $\hat{\mu}_i$'s will be in the tails of their distribution, the larger the MSEs. Conversely, the MSEs will be smaller for less $\mathbb{V}^{(\mathcal{Y}_i,\mu_i)}_{\theta,\phi,\pi(\bar{m})}$ (e.g., with $\xi^*\leq 5$) and moderate for quite high $\mathbb{V}^{(\mathcal{Y}_i,\mu_i)}_{\theta,\phi,\pi(\bar{m})}$ (e.g., with $5\leq\xi^*\leq 10$). (iv) Four, in a semiparametric context, whether model classes in $\mathcal{E}$ are high dimensional (e.g., highly large heterogeneity among subgroups), the expected loss in (24) is minimized as $N\to\infty$ and $\pi^{\bar{m}}$ will converge to a limit that is optimal. More precisely, the final model solution will correspond to the oracle forecast of the prior (or correlated random effect distribution) that most favours the true model (Tweedie correction). Thus, for a sufficiently large sample size, the EB approach in Section (4.2) would give a solution that is close to the Bayesian oracle and thus would exploit information more efficiently than a fixed choice of μ_i (e.g., full Bayesian solutions)[21]. (v) Five, the ratio-optimality in (34) allows for the presence of estimated parameters in the sufficient statistic $\hat{\mu}_i$ and uniformity with respect to the correlated random effect density $\pi^{\bar{m}}$, which is allowed to have unbounded support.

For $\bar{m}>0$, the resulting predictor is:

$$\hat{y}_{i,T+k}=\Big[\hat{\mu}_i^{\bar{m}}(\theta)+\frac{\hat{\sigma}_u^2}{T}\cdot\frac{\partial}{\partial\hat{\mu}_i^{\bar{m}}(\theta)}ln\Big(\hat{\mu}_i^{\bar{m}}(\theta),y_{i0}^{\bar{m}}\Big)\Big]^{\bar{m}}+\hat{\theta}y_{it} \tag{35}$$

where $\bar{m}<\infty$ according to all possible model solutions $M_{\xi^*}\subset\mathcal{E}$.

The posterior distributions for $\ddot{\tilde{\omega}}=\Big(\ddot{\bar{\theta}},\ddot{\bar{\rho}},\tilde{\delta}_{\mu_i},\tilde{\Psi}_{\mu_i},\bar{\kappa},\tilde{w}_{|\chi|},\ddot{\bar{\omega}},\tilde{\nu}\Big)$ are calculated by combining the prior with the (conditional) likelihood for the initial conditions of the data. The resulting function is then proportional to

[21] See, for instance, George and Foster (2000) and Scott and Berger (2010).

$$L\Big(y_i^{0:T}|\ddot{\tilde{\omega}}\Big) \propto exp\Bigg\{-\frac{1}{2}\Bigg[\sum_{t=1}^{T}\Big(y_{it}-(c_{it}^{\bar{m}}|\dot{\chi})\hat{\theta}_t-\hat{\mu}_i^{\bar{m}}(\hat{\theta})\Big)'\Bigg]\cdot(\hat{\sigma}_u^2)^{-1}\cdot\Bigg[\sum_{t=1}^{T}\Big(y_{it}-(c_{it}^{\bar{m}}|\dot{\chi})\hat{\theta}_t-\hat{\mu}_i^{\bar{m}}(\hat{\theta})\Big)\Bigg]\Bigg\} \tag{36}$$

where $y_i^{0:T}=(y_{i0},y_{i1},\ldots,y_{iT})$ denotes the data and $\ddot{\tilde{\omega}}=\Big(\ddot{\bar{\theta}},\ddot{\bar{\rho}},\tilde{\delta}_{\mu_i},\tilde{\Psi}_{\mu_i},\bar{\kappa},\tilde{w}_{|\chi|},\ddot{\dot{\omega}},\tilde{\nu}\Big)$ refers to the unknowns whose joint distributions need to be found.

Despite the dramatic parameter reduction implicit in the MPROB procedure, the analytical computation of posterior distributions $(\ddot{\tilde{\omega}}|\hat{y}_{i,T+k})$ is unfeasible, where $\hat{y}_{i,T+k}$ denotes the expectations of outcomes associated with the infeasible oracle forecast to be estimated (equation (23)). Thus, I use MCMC implemetations to draw conditional posterior distributions of $(\theta_1,\theta_2,\ldots,\theta_T|\hat{y}_{i,T+k},\ddot{\tilde{\omega}}_{-\theta_t})$, with $\ddot{\tilde{\omega}}_{-\theta_t}$ standing the vector $\ddot{\tilde{\omega}}$, but excluding the parameter θ_t. More precisely, I include a variant of the Gibbs sampler approach – the Kalman-Filter technique – to analytically obtain forward recursions for posterior means and covariance matrix. Starting from $\bar{\theta}_{T|T}$ and $\bar{\rho}_{T|T}$, the marginal distributions of θ_t can be computed by averaging over draws in the nuisance dimensions, and the Kalman filter backwards can be run to characterise posterior distributions for $\ddot{\tilde{\omega}}$:

$$\theta_t|\theta_{t-l},\hat{y}_{i,T+k},\ddot{\tilde{\omega}}_{-\theta_t}\sim N\Big(\ddot{\bar{\theta}}_{t|T+k},\ddot{\bar{\rho}}_{t|T+k}\Big) \tag{37}$$

where

$$\ddot{\bar{\theta}}_{t|T+k}=\Bigg[\Big(\ddot{\bar{\rho}}_{t|T+k}^{-1}\cdot\bar{\theta}\Big)+\sum_{t=1}^{T}\Big((c_{it}^{\bar{m}}|\dot{\chi})'\cdot(\hat{\sigma}_u^2)^{-1}\cdot(c_{it}^{\bar{m}}|\dot{\chi})\Big)\hat{\theta}_t\Bigg] \tag{38}$$

$$\ddot{\bar{\rho}}_{t|T+k}=\Big[I_h-\Big(\bar{\rho}\cdot\ddot{\bar{\rho}}_{T+k|t}^{-1}\Big)\Big]\cdot\bar{\rho} \tag{39}$$

with

$$\hat{\theta}_t=\Big[(c_{it}^{\bar{m}}|\dot{\chi})'\cdot(\hat{\sigma}_u^2)^{-1}\cdot(c_{it}^{\bar{m}}|\dot{\chi})\Big]^{-1}\cdot\Big[(c_{it}^{\bar{m}}|\dot{\chi})'\cdot(\hat{\sigma}_u^2)^{-1}\cdot y_{it}\Big] \tag{40}$$

The equations (39) and (40) refer to the variance-covariance matrix of the conditional distribution of $\ddot{\bar{\theta}}_{t|T+k}$ and the GMM estimator, respectively. By rearranging the terms, equation (38) can be rewritten as

$$\ddot{\bar{\theta}}_{t|T+k}=\Bigg[\Big(\ddot{\bar{\rho}}_{t|T+k}^{-1}\cdot\bar{\theta}\Big)+\Bigg(\sum_{t=1}^{T}(c_{it}^{\bar{m}}|\dot{\chi})'\cdot(\hat{\sigma}_u^2)^{-1}\cdot y_{it}\Bigg)\Bigg] \tag{41}$$

where $\ddot{\bar{\theta}}_{t|T+k}$ and $\ddot{\bar{\rho}}_{t|T+k}$ denote the smoothed k-period-ahead forecasts of θ_t and of the variance–covariance matrix of the forecast error, respectively.

The above output of the Kalman filter is used to generate a random trajectory for $\left\{\theta_t\right\}_{t=1}^{T}$ by using the backward recursion starting with a draw of $\{\theta_t\}$ from $N\left(\ddot{\theta}_{T|T}, \ddot{\rho}_{T|T}\right)$[22]. Given (37), the other posterior distributions can be defined as:

$$\pi(\hat{\mu}_i|\hat{y}_{i,T+k}, \hat{\theta}_t) \sim N\left(\tilde{\delta}_{\mu_i}, \tilde{\Psi}_{\mu_i}\right) \tag{42}$$

$$\pi(\hat{y}_{i0}|\hat{\mu}_i^{\bar{m}}) = N(0, \bar{\kappa}) \tag{43}$$

$$\pi(\dot{\chi}) = \tilde{w}_{|\chi|} \cdot \binom{\xi^*}{|\chi|}^{-1} \tag{44}$$

$$\pi(\hat{\sigma}_u^2|\hat{y}_{i,T+k}) = IG\left(\frac{\ddot{\omega}}{2}, \frac{\tilde{\nu}}{2}\right) \tag{45}$$

Here, some considerations are in order.

In equation (42), $\tilde{\delta}_{\mu_i} \sim N\left(0, \bar{\zeta}\right)$ and $\tilde{\Psi}_{\mu_i} \sim IG\left(\bar{\varphi}/2, \bar{\varepsilon}/2\right)$, where $\bar{\zeta} = \zeta + (u_{it}^{'}u_{it})$, $\bar{\varphi} = \varphi \cdot \dot{\chi}$, and $\bar{\varepsilon} = \varepsilon \cdot \dot{\chi}$, with (ζ, ε) denoting the arbitrary scale parameters (sufficiently small) and φ referring to the arbitrary degree of freedom (chosen to be close to zero). In this analysis, $\tilde{\Psi}_{\mu_i}$ is obtained by using the (proposal) joint posterior density ($\pi^{\bar{m}}$) sampled via EM algorithm, $(\zeta, \varepsilon) \cong 0.001$, and $\varphi \cong 0.1$.

In equation (43), $\bar{\kappa} = \kappa + \mathbb{V}_{\theta,\phi,\pi(\bar{m})}^{(\mathcal{Y}_i,\mu_i)}[\mu_i]$, with κ and $\mathbb{V}_{\theta,\phi,\pi(\bar{m})}^{(\mathcal{Y}_i,\mu_i)}[\cdot]$ denoting the arbitrary scale parameter and the posterior predictive variance of μ_i, respectively. In this analysis, $\kappa \cong 1.0$ and $\mathbb{V}_{\theta,\phi,\pi(\bar{m})}^{(\mathcal{Y}_i,\mu_i)}[\mu_i]$ is obtained according to the sample size $|\chi|$ as described in (33).

In equation (44), $\tilde{w}_{|\chi|}$ refers to the model posterior choice according to the sum of the PMPs obtained in the second stage with respect to model size $|\chi|$, with $\tilde{w}_{|\chi|} = max^*(NT, |\chi|)$ accounting for the non-0 components of $\dot{\chi}$.

In equation (45), $\ddot{\omega} = \bar{\omega} + \hat{\hat{\omega}}$ and $\tilde{\nu} = \nu + \hat{\nu}$, with $\bar{\omega}$ and ν denoting the arbitrary degrees of freedom (sufficiently small) and the arbitrary scale parameter, respectively, $\hat{\hat{\omega}} = \left(\sum_{t=1}^{T} log(\tau_t)/t\right) + log\left(\sum_{t=1}^{T}(1/\tau_t)\right) - log(t)$ and $\hat{\nu} = (t \cdot \hat{\hat{\omega}})/\left(\sum_{t=1}^{T}(1/\tau_t)\right)$ referring to the Maximum Likelihood Estimates (MLEs). In this analysis, $\tau_t = \{\tau_1, \ldots, \tau_T\}$ is the random sample from the data $\{0, T\}$, $\bar{\omega} \cong 0.1 \cdot \dot{\chi}$, $\nu \cong 0.001$, and $\hat{\nu}$ is obtained by numerically computing $\hat{\hat{\omega}}$.

Finally, the last two hyperparameters to be defined in the vector ϖ are $\bar{\theta} = \hat{\theta}_0$, with $\hat{\theta}_0$ denoting the GMM estimators of equation (2) related to the posteriors $\hat{y}_{i0}$ in (43), and $\bar{\rho} = I_{\xi^*}$.

Let the stationarity and moment conditions in (3)-(5) hold, all posterior estimates ensure exact identification and unbiasedness of the time-varying coefficient vectors β_l and γ_l in a context of correlated random effects.

[22]See, for instance, Carro (2007).

4.4 Theoretical Properties

The proof of the ratio-optimality result and posterior consistency are as follow.

Assumption 4.1 (Identification: General Model). *Consider the DPD model in (2):*

1. *Model Setup*

   a. *$(c_{i0}, \mu_i, \sigma_u^2)$ are i.i.d. across i.*

   b. *For all t, conditional on (y_{it}, c_{it-l}), $y_{i,t}^c$ is independent of $\phi = (\mu_i, \sigma_u^2)$.*

   c. *$x_{i,0:T}$ is independent of $\phi = (\mu_i, \sigma_u^2)$.*

   d. *Conditioning on (c_{i0}, μ_i) and σ_u^2, they are independent of each other.*

   e. *Let $u_{it} \sim N(0, \sigma_u^2)$ is i.i.d. across i and independent of (c_{it-l}, μ_i).*

2. *Indentification*

   a. *The characteristic functions for $\mu_i | c_{i0}$ and $\sigma_u^2 | c_{i0}$ do not steadily disappearing altogether into high-shrinkage processing method.*

   b. *Let $v_{it} = \mu_i + u_{it}$ be the composite error at time t, the sequence $\{v_{it} : t = 1, 2, \ldots, T\}$ is almost certainly serially correlated, and definitely is if $\{u_{it}\}$ is serially uncorrelated.*

   c. *With the panel setup in (2) – with large N and sufficiently large T – x_{it} includes interactions of variables with time periods dummies and general non-linear functions and interactions, so the model is quite flexible.*

   d. *For the CRE approach, each kind of covariates in c_{it} is separated out and the heterogeneous factors in μ_i are correlated with them.*

Assumption 4.2 (Identification: Unbalanced Panel). *For all i:*

1. *Indentification*

   a. *c_{i0} is observed.*

   b. *(y_{iT}, z_{iT}, x_{iT}) are observed.*

   c. *With the panel setup in (2) – with large N and sufficiently large T – x_{it} includes interactions of variables with time periods dummies and general non-linear functions and interactions, so the model is quite flexible.*

   d. *For the CRE approach, each kind of covariates in c_{it-l} is separated out and the heterogeneous factors in μ_i are correlated with them.*

2. `Sequential Exogeneity (Conditional on the Unobserved Effects)`

   a. $E(y_{it}|c_{it}, c_{it-1}, \ldots, c_{i1}, \mu_i) = E(y_{it}|c_{it}, \mu_i) = \theta c_{it} + \mu_i$ *for* $i = 1, 2, \ldots, N$.

   b. *Sequential exogeneity is a middle ground between contemporaneous and strict exogeneity. It allows lagged dependent variables and other variables that change in reaction to past shocks.*

   c. *Because including contemporaneous exogeneity, standard kinds of endogeneity – where some elements of* c_{it} *are correlated with* u_{it} *– are ruled out such as measurement error, simultaneity, and time-varying omitted variables.*

   d. *Sequential exogeneity is less restrictive than strict exogeneity imposing restrictions on economic conditions.*

3. `Model Setup`

   a. *The term "correlated random effects" is used to denote situations where one models the relationship between* $\{\mu_i\}$ *and* $\{c_{it}\}$.

   b. *The CRE approach allows to unify fixed and random effects estimation approaches.*

   c. *With the CRE approach, time-constant variables can be included within the system.*

   d. *GMM estimators can be used to consistently estimate all time-varying parameters in* θ.

   e. $\hat{\theta}_{GMM} \xrightarrow{d} N\Big(\theta, \frac{1}{N}V_\theta\Big)$ *where* V_θ *denotes the covariance matrix estimated via posteriors on* $\{\bar{\theta}, \bar{\rho}\}$.

Assumption 4.3 (Identification: Conjugate Proper Informative Priors)**.** *Let* $\mathbb{E}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi(\cdot)}$ *be the expectations over the observed trajectories* ($\mathcal{Y}_i$) *and the unobserved heterogeneity* (μ_i)*, the CIPM priors in (16) - (20) can be rewritten as:*

i.

$$\theta_t|\theta_{t-l}, \mathbb{E}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi(\cdot)}[y_{i,T+k}] \sim N\Big(\bar{\theta}, \bar{\rho}\Big)$$

ii.

$$\mu_i|\mathbb{E}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi(\cdot)}[y_{i,T+k}] \sim N\Big(\delta_{\mu_i}, \Psi_{\mu_i}\Big) \quad \texttt{with} \quad \delta_{\mu_i} \sim N\Big(0, \zeta\Big) \quad \texttt{and} \quad \Psi_{\mu_i} \sim IG\Big(\frac{\varphi}{2}, \frac{\varepsilon}{2}\Big)$$

iii.

$$y_{i0}|\mathbb{E}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi(\cdot)}[y_{i,T+k}] = N(0, \kappa)$$

iv.

$$\pi(\chi) = w_{|\chi|} \cdot \binom{h}{|\chi|}^{-1}$$

v.

$$\sigma^2_u|\mathbb{E}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi(\cdot)}[y_{i,T+k}] \sim IG\Big(\frac{\bar{\omega}}{2}, \frac{\nu}{2}\Big)$$

Statement 4.3.1 (Posterior Distributions). *Under Assumption (4.3), all posterior distributions in (41)-(45) hold and are estimable through MCMC algorithms and implementations.*

Statement 4.3.2 (Moment Conditions and GMM Estimator). *Let all moment conditions in (3)-(5) hold for all i. Then, under Assumption (4.3) and Statement (4.3.1), the GMM estimator is consistent and equals to:*

$$\hat{\theta}_t = \left[\left(c_{it}^{\bar{m}} | \dot{\chi} \right)' \cdot \hat{\sigma}_u^2 \cdot \left(c_{it}^{\bar{m}} | \dot{\chi} \right) \right]^{-1} \cdot \left[\left(c_{it}^{\bar{m}} | \dot{\chi} \right)' \cdot \hat{\sigma}_u^2 \cdot y_{it} \right]$$

Theorem 4.4 (Correlated Random Coefficients: Cross-sectional Homoskedasticity). *Let the Tweedie's formula in (28) hold.*

1. *The proof of the Tweedie correction for (28) – when dealing with correlated random coefficients – builds on a finite mixture approximation of multivariate distributions as defined in (29).*
2. *With correlated random coefficients homoskedastic case, one would work with the following space:*

$$\mathcal{Q}^{\bar{m}} = \left\{ \bar{m} \subset \mathcal{E} : \iint \parallel \mu_i \parallel_2^2 \cdot \pi^{\bar{m}} \cdot q_{i0}^*(c_{i0}) \; d\mu_i \; dc_{i0} \leq M_{\xi^*} \right\} \quad \textit{for} \quad M_{\xi^*} > 0$$

   where $\pi^{\bar{m}}$ denotes the (proposal) joint density distribution to draw posteriors for $\hat{\mu}_i$ through Metropolis-Hastings algorithm, and q_{i0}^* denotes the true marginal density of c_{i0}.

3. *The space for common parameters $\upsilon = (\theta, \sigma_u^2)$ is $\Upsilon = \mathbb{R}_{|\chi|} \cdot \sigma_u^2$.*
4. *The conditional individual–specific likelihood function is described as:*

$$\mathcal{L}(y_{it}|\upsilon, \bar{m}) = \pi\Big(c_{it}|c_{it-l}, c_{i,0:T}\Big) \iint \prod_t \pi\Big(y_{it}; \theta' c_{it-l} + \mu_i, \sigma_u^2\Big) \cdot \pi^{\bar{m}} \cdot q_{i0}^*(c_{i0}) \; d\mu_i \; dc_{i0}$$

Theorem 4.5 (Posterior Consistency: Correlated Random Coefficients). *Given the DPB-CRE in (2):*

1. `Model` $\rightarrow$ *Assumptions (4.1) and (4.2).*
2. `Covariates` $\rightarrow$ $\left(y_{i,0}^o, y_{i,0}^c, z_{i,0}^s, z_{i,0}^p, x_{i,0:T}\right)$ *satisfy Assumptions (4.1) and (4.2).*
3. `Common Parameters`

   a. *υ is unknown and estimable.*

   b. *The domain of σ_u^2 is finite and estimable (homoskedastic case).*

Thus, the posterior would be highly consistent at υ given $\pi^{\bar{m}}$ with regimes $\bar{m}$.

Theorem 4.6 (Ratio-Optimality: Tweedie Correction)**.** *Suppose that Assumptions (4.1) - (4.3) hold: then, in the DPB-CRE in (2), the EB predictor ($\hat{y}_{i,T+k}$) defined in (35) achieves ϑ_0-ratio-optimality uniformly for priors $\pi^{\bar{m}} \subset \mathcal{E}$, with $\vartheta_0 \geq 0$.*

Statement 4.6.1 (Properties of the Common Parameters Estimation)**.** *The estimators of the common parameters (υ) have the following properties:*

i

$$\mathbb{E}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi^{\bar{m}}}[\hat{\sigma}_u]^4 = o_{u.\pi^{\bar{m}}}(N^+)$$

ii

$$\mathbb{E}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi^{\bar{m}}}\left[|\sqrt{N}(\hat{\theta}-\theta)|^4\right] = o_{u.\pi^{\bar{m}}}(N^+)$$

iii

$$\mathbb{E}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi^{\bar{m}}}\left[|\sqrt{N}(\hat{\sigma}^2_u-\sigma^2_u)|^2\right] = o_{u.\pi^{\bar{m}}}(N^+)$$

iv

$$N\iint \left[|\chi|\cdot\pi^*_\xi\left(\hat{\mu}^{\bar{m}}_i, y^{\bar{m}}_{i0} \mid c^{\bar{m}}_{i0}\right)\right]^2 \cdot \left[|\chi|\cdot\pi_\xi\left(\hat{\mu}_i, y_{i0} \mid c_{i0}\right)\right] d\hat{\mu}\; dy_{i0} = o_{u.\pi^{\bar{m}}}(N^+) \quad \textit{with} \quad \bar{m} < \infty$$

where $\pi^{\bar{m}} = |\chi|\cdot\pi^*_\xi\left(\hat{\mu}^{\bar{m}}_i, y^{\bar{m}}_{i0} \mid c^{\bar{m}}_{i0}\right)$ stands for the (designed) joint density distribution – under probability ($\mathfrak{p}_{\bar{m}}$) – to get draw samples from posteriors of the empirical distribution of $\hat{\mu}_i$.

v

$$\mathfrak{p}_{\bar{m}} = \frac{\pi\left(\hat{\mu}^{\bar{m}}_i, y^{\bar{m}}_{i0} \mid \hat{\mu}^{\bar{m}-l}_i, \mathcal{Y}_i, \left\{\theta_t\right\}_{t=1}^{T}, c^{\bar{m}}_{i0}\right)\cdot\pi^{\bar{m}-l}}{\pi\left(\hat{\mu}^{\bar{m}-l}_i, y^{\bar{m}}_{i0} \mid \hat{\mu}^{\bar{m}-l}_i, \mathcal{Y}_i, \left\{\theta_t\right\}_{t=1}^{T}, c^{\bar{m}}_{i0}\right)\cdot\pi^{\bar{m}}} \leq 1$$

where the further $\mathfrak{p}_{\bar{m}}$ will be close to zero, the lower the PMP and then the lower the possibility that individual outcome values would perform well the data in the cross-sectional information subset $\mathcal{E}$. Whether $\mathfrak{p}_{\bar{m}} = 1$, the associated model solution (M_{ξ^*}) would contain best estimates for the μ_i's sufficient statistic and individual outcome values in (28). Thus, one would expect to find the final model solution with higher lBF.

Statement 4.6.2 (Posterior Mean Functions in DPB-CRE with Empirical Bayes Approach)**.** *The Theorem (4.6) can be proved by showing that the below inequality holds:*

$$\limsup_{N\to\infty}\ \sup_{\pi^{\bar{m}}\subset\mathcal{E}} \frac{N\mathbb{E}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi^{\bar{m}}}\left[\left(\hat{y}_{i,T+k}-(\mu_i+\theta y_{it})\right)^2\right]}{N\mathbb{E}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi^{\bar{m}}}\left[\left(\mu_i-\mathbb{E}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi^{\bar{m}}}[\mu_i]\right)^2\right]+N^{\vartheta_0}} \leq 1$$

Thus, according to a generalization of the Brown and Greenshtein (2009)'s insight[23]*, the sufficient condition (4.6.2) is proved by decomposing the discrepancy between the predictor $\hat{y}_{i,T+k}$ and the unknown parameters $(\mu_i+\theta y_{it})$ into three terms:*

i

$$N\mathbb{E}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi^{\bar{m}}}[\hat{y}_{i,T+k}]^2 = o_{u.\pi^{\bar{m}}}(N^{\vartheta_0})$$

It displays the difference between the posterior mean of μ_i according to the Tweedie correction defined in (29) and the (proposal) joint density distribution $\pi^{\bar{m}}$ with probability $\mathfrak{p}_{\bar{m}}$ in (30).

ii

$$\limsup_{N\to\infty}\ \sup_{\pi^{\bar{m}}\subset\mathcal{E}} \frac{N\mathbb{E}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi^{\bar{m}}}\left[\mu_i+\theta y_{it})\right]^2}{N\mathbb{E}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi^{\bar{m}}}\left[\left(\mu_i-\mathbb{E}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi^{\bar{m}}}[\mu_i]\right)^2\right]+N^{\vartheta_0}} \leq 1$$

It displays the difference between the posterior density estimated in (29) and the (proposal) joint density distribution $\pi^{\bar{m}}$ with probability $\mathfrak{p}_{\bar{m}}$ in (30).

iii

$$N\mathbb{E}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi^{\bar{m}}}[\pi^{\bar{m}}] = o_{u.\pi^{\bar{m}}}(N^{+})$$

It displays the structural model uncertainty dealt with replacing the common parameters (θ) by estimates.

The results (i), (ii), and (iii) are relatively straightforward under Statement (4.6.1) and can be handled according to Statements (4.3.1) and (4.3.2).

Theorem 4.7 (Finite Mixtures of Multivariate Distributions)**.** *Suppose that Assumptions (4.1)-(4.3) hold: then, the density $\pi\Big(\hat{\mu}_i(\theta), y_{i0}\Big)$ can be estiamted using FMM distributions as defined in (29).*

Statement 4.7.1 (Mixture Components and Model Classes)**.** *The finite mixture of multivariate distributions in (29) is able to approximate a large set of distributions as the number of mixture (potential) combination of predictors ($|\chi|$) – set into $\mathcal{E}$ – increases. In this study, I use finite mixtures of multivariate normal-inverse-gamma distributions dealing with the common parameters υ (Theorem (4.4)). According to posterior distributions (41)-(45) – under Assumption (4.3) – the finite mixtures of multivariate distributions will be:*

[23]See, for instance, Liu et al. (2020) using the insight of Brown and Greenshtein (2009) to prove ratio-optimality by replacing the μ_i's distribution with a kernel density estimator.

$$\pi^*_{mix}\Big(\hat{\mu}_i^{\bar{m}}, y_{i0}^{\bar{m}} \mid |\chi|, c_{i0}^{\bar{m}}\Big) = |\chi| \cdot \pi^*_{\xi}\Big(\hat{\mu}_i^{\bar{m}}, y_{i0}^{\bar{m}} \mid c_{i0}^{\bar{m}}\Big) \quad \textit{with} \quad |\chi| > 0 \quad \textit{and} \quad \mathfrak{p}_{\bar{m}} \leq 1$$

Theorem 4.8 (Normal-Inverse-Gamma-Distribution)**.** *It is the conjugate prior of a normal distribution with unknown mean and variance. For instance, suppose that the distribution of the (non-)homogeneous parameters θ would be affected by heterogeneous effects ϕ_i. According to CIPM in (15):*

$$\theta|\sigma_u^2 \sim N\Big(\bar{\theta}, \bar{\rho} \otimes \sigma_u^2\Big) \quad \textit{and} \quad \sigma_u^2|\bar{\omega}, \nu \sim IG\Big(\frac{\bar{\omega}}{2}, \frac{\nu}{2}\Big)$$

Then,

$$\Big(\theta, \sigma_u^2\Big) \sim NIG\Big(\bar{\theta}, \bar{\rho}, \dot{\bar{\omega}}, \dot{\nu}\Big)$$

with $\dot{\bar{\omega}} = \frac{\bar{\omega}}{2}$ and $\dot{\nu} = \frac{\nu}{2}$.

Concerning $\Big(\theta, \mu_i\Big)$:

$$\theta|\mu_i \sim N\Big(\bar{\theta}, \bar{\rho} \otimes \Psi_{\mu_i}\Big) \quad \textit{and} \quad \mu_i|\delta_{\mu_i}, \Psi_{\mu_i} \sim IG\Big(\delta_{\mu_i}, \Psi_{\mu_i}\Big)$$

Then,

$$\Big(\theta, \mu_i\Big) \sim NIG\Big(\bar{\theta}, \bar{\rho}, \delta_{\mu_i}, \Psi_{\mu_i}\Big)$$

Theorem 4.9 (Time-varying Parameters and GMM Estimators)**.** *Let the stationarity and moment conditions in (3)-(5) hold in the system, then the time-series regressions are valid (or computational) and GMM estimators are feasible.*

5 Empirical Evidence: Socio–demographic Factors and Policy Tools during the Coronavirus Outbreak

5.1 Literature Review and Discussion

The COVID-19 pandemic crisis, started in Wuhan (China) in December 2019, is a major global crisis with far-reaching implications in terms of health and economics (see, e.g., Zaki et al. (2020), Sorokowski et al. (2020), and She et al. (2020)). Due to the lunar year celebration in China, the huge movement of people from and between the Asian region and other parts of the world have increased the geographical spreading of the COVID-19 virus, diffusing to more than 215 countries and regions (see, e.g., Boulos et al. (2020), Zheng (2020), and Gössling et al. (2020)). Then, the pandemic situation has been aggravated by urbanization because of countries with higher population density tend to have higher risk of contamination (see, for instance, Lindahl and Grace (2015)). On March 11, 2020, the World Health Organization (WHO) has officially declared the COVID-19 outbreak a global pandemic, where the main distressing social consequences of the pandemic have been prejudice and anxiety, increasing psychological distancing from the nations most affected (see, e.g., Karwowski et al. (2020) and Sorokowski et al. (2020)). Moreover, the spread of coronavirus has highlighted the importance of dealing with heterogeneity, commonality, and interdependency among countries and sectors. More precisely, even if the virus has rapidly moved across borders because of business links and other existing relationships, perceptions about the crisis and social behaviours have been addressed with lags according to the observed experiences abroad. Indeed, most countries have had the opportunity of learning from others about social adjustments that have been more or less effective in containing the disease. At the present time, after initial outbreaks have been supressed for now in China and South Korea, the virus continues to spread at an alarming rate throughout Europe.

All governments have been overflowed by the widespread pandemic and then forced to take hard measures to reduce the impact of the disease spread and avoid a complete health system collapse which conversely would have resulted in a more negative valuation of the government's policy response. Globally, there have been implemented social distancing restrictions – such as closure of schools, airports, borders, restaurants and shopping malls – and, in the most severe cases, lockdowns prohibiting all citizens from leaving their homes. Such measures have subsequently led to a consistent economic downturn with stock markets plummeted, international trade slowed down, bankrupt businesses, and people unemployed (see, e.g., Temsah et al. (2020), Mikolai et al. (2020), and Nicola et al. (2020)). Although the implemented restrictions have significantly challenged the expected shock from the pandemic, the extent of the disease spread among countries has highly varied from one economy to another.

Several frequentist statistical, dynamic and multicountry approaches have been proposed to identify and analyze socioeconomic effects and policy measures during the outbreak of the COVID-19 (see, e.g., Tuite and Fisman (2020), Wu et al. (2020), Hassan et al. (2020), Laskowski et al. (2011), and Bernanke (2020) for some relevant discussions). More precisely, standard statistical inference and multivariate regression analysis have been performed to evaluate the effects of a pool of variables – such as economic status, population density, the median age of the population, and urbanization pattern – on the spread of the COVID-19 among countries. These studies achieve four main results: (*i*) countries' population density does not have any positive significant relations with COVID-19 outcomes; (*ii*) high-income countries have featured larger fiscal policies and unconventional monetary policy tools than lower-income countries; (*iii*) lower-middle-income and upper-middle-income countries are less likely to have an increased recovery rate compared to high-income countries; and (*iv*) countries with higher economic–financial statistics have implemented better workplaces for more income, better quality of life, and better facilities to reduce the chance of contamination. However, frequentist approaches suffer from well-known downward bias and related asymmetric distribution. Moreover, they do not deal with unobserved heterogeneity and misspecified dynamics among countries, and thus cannot capture significant interactions between potential determinants.

In Bayesian statistics, these disadvantages can be accounted for. However, there are limited research results on the global risk factors of COVID-19 transmission and patterns of spread. In a recent study, Stojkoski et al. (2020) have performed Bayesian Model Averaging (BMA) techniques and country level data to investigate potential determinants and policy tools in explaining the COVID-19 pandemic outcomes. Their analysis suggests three main findings: (*i*) two determinants strongly related to the coronavirus cases are the population size and the government health expenditure; (*ii*) more populated economies show greater resistance to being infected by the virus, while countries with larger government expenditure display greater susceptibility to the virus infection; and (*iii*) there is no determinant strongly related to the coronavirus deaths per million population. Nevertheless, in BMA and BMS, some open related features need to be accounted for such as endogeneity issues (because of unobserved heterogeneity and omitted factors), overfitting, and structural model uncertainty (when one or more forms of misspecification matter).

5.2 Data Description and Results

The DPB-CRE in (2) contains 22 country-specific models, including 9 advanced economies[24], 7 emerging economies[25], and 6 non European countries[26]. Moreover, all advanced countries – except for *SVN* – refer to Western Europe (WE) economies and all emerging countries – except for *GRC* – refer to Central-Eastern Europe (CEE) economies, respectively. All European countries are Eurozone members, with the exception of *CZE* and *POL*, and thus interdependencies and inter-country linkages can be investigated in depth. The estimation sample is expressed in years and covers the period from $1990 - 2020$, and all data comes from World Bank database. Given the hierarchical structural conformation of the model and a sufficiently large number of years describing economic–financial and policy issues, it is able to capture: (*i*) endogeneity issues;

[24]Austria (*AUT*), Finland (*FIN*), France (*FRA*), Germany (*DEU*), Ireland (*IRL*), Italy (*ITA*), Netherlands (*NLD*), Slovenia (*SVN*), and Spain (*ESP*).

[25]Czech Republic (*CZE*), Poland (*POL*), Slovak Republic (*SVK*), Estonia (*EST*), Latvia (*LVA*), Lithuania (*LTU*), and Greece (*GRC*).

[26]United States (*USA*), China (*CHN*), Korea (*KOR*), Japan (*JPN*), United Kingdom (*GBR*), and Chile (*CHL*).

(*ii*) interdependency, commonality, and homogeneity; (*iii*) relevant monetary and fiscal policy interactions, and contagion measures; and (*iv*) misspecified dynamics.

Given the DPB-CRE in (2), I use the conditioning set c_{it} in (6) to identify potential determinants during the current COVID-19 pandemic crisis and then perform future policy strategies to prevent the emergence of epidemics on the global economy. The decomposed vectors of the observed time-varying endogenous variables ($y_{i,t-l}$, $z_{i,t-l}$) are: (*i*) $y^{o'}_{i,t-l}$ denoting lagged outcomes to capture the persistence; (*ii*) $y^{c'}_{i,t-l}$ indicating general economic conditions; (*iii*) $z^{s'}_{i,t-l}$ indicating socio–demographic statistics (including socioeconomic factors); and (*iv*) $z^{p'}_{i,t-l}$ denoting economic–financial variables (including fiscal and monetary tools). Finally, the strictly exogenous factors x_{it} contain dummy variables to test the presence of structural breaks[27]. Even if $y_{i,t-l}$ and $z_{i,t-l}$ are described by common coefficients (β and γ), they could have non-homogeneous effects on the outcomes (y_{it}). For example, high-income countries could have better recovery rate than the lower-income ones or countries with better economic–financial status would be more likely to implement better workplaces and health services.

The dataset contains 92 observable variables dealing with all potential determinants and policy tools described through the vectors $y_{i,t-l}$ and $z_{i,t-l}$. In this study, I split them in four groups.

(i.) Economic Status:

The group refers to 41 determinants combining information on education (total enrolment rates at the primary, secondary, and tertiary level), income (through Gini index[28] measures as annual % and current US$), economic development (GDP per capita as current US$), labour market (labour force participation rate at national and education level), wage (national level), salaried workers (national level), employment/unemployment rate (national and working level), trade (exports and imports of goods and services as % of GDP and annual % growth), and foreign direct investments.

(ii.) Healthcare Statistics:

The group accounts for 11 determinants combining information on health coverage (capital investments[29] and current expenditures[30] as % of GDP), public expenditures on health from domestic sources (transfers, internal grants and transfers, subsidies to voluntary health insurance beneficiaries, and non-profit institutions serving households), domestic health coverage (as % of GDP and general government expenditures), and death race (per 1,000 people occurred during each year).

(iii.) Demographic and Environment Statistics:

The group accounts for 28 determinants combining information on fertility rate (per 1,000 women during each year), age dependency ratio (per 100 working-age population), population (as % of total population and annual % growth), rural and urban population (as % of total population and annual

[27] In this study, (potential) structural breaks are assessed using the Chow test.
[28] The Gini index is a measure of statistical dispersion representing the income or wealth inequality.
[29] Capital health investments also include health infrastructure (e.g., buildings, machinery, and Information Technology) and stocks of vaccines for outbreaks.
[30] Current health expenditures also include healthcare goods and services consumed per year.

% growth), tobacco use and alcohol consumption per capita (as % of adults), net energy imports (as % of energy use), sources of electricity[31], and CO2 emissions from electricity production and use of natural gas.

(iv.) Economic-financial Issues:

The group refers to 12 determinants dealing with real–financial economy (describing fiscal policy measures) and financial markets (describing monetary policy tools)[32]. The former includes – for example – GDP growth per capita, general government final consumption expenditure (as % of GDP), and gross fixed capital formation (as % of GDP) for the real dimension; and current account balance (as % of GDP), public debt (as % of GDP), and long term interest rate for the financial dimension. Lending markets account for inflation (consumer price index as annual %), bank leverage (as loan to deposit ratio), domestic credit (provided by financial sector as % of GDP), and net transactions in financial assets and liabilities (as lending and borrowing ratio). The outcomes of interest corresponds to GDP per capita in PPP[33] (hereafter, 'productivity').

By looking into which (potential) candidate predictors are included with higher frequency in the final subset better fitting the data, I run the MPROB procedure described in Section 3.2. In the first stage, I find that 53 *best* covariates[34] better fit the data with PIPs[35] $\geq \tau$ in (11) and $\chi = 1$. Thus, I obtain 2^{53} *best* model solutions ($M_j \subset \mathcal{S}$). Because of the curse of dimensionality, I further shrink the data performing the second stage implicit in the MPROB procedure. Overall, 31 *top best* covariates are found, obtaining 2^{31} *top best* model solutions ($M_\xi \subset \mathcal{E}$ in (12)) with $\dot{\chi} = 1$ (Table 1). Here, some preliminary results can be addressed. (*i*) Most of model uncertainty and overfitting are avoided: indeed, the CPS[36] tends to be close to 0 – such as for predictors $(7, 25)$ – and 1 – such as for predictors $(2, 9, 11, 26, 31)$. (*ii*) Uncertain effects tend to persist in predictors $(3, 5, 10, 16, 20, 22)$: thus, they should be interpreted with care. (*iii*) Socioeconomic factors (including healthcare, demographic, and environment statistics) matter more than economic status because of the sudden outbreak of the epidemiology (16/39 factors for socioeconomic–demographic statistics compared to 8/41 factors for economic status). (*iv*) The main policy tools correspond to some of the core variables of real and financial business cycles affecting the spreading and transmission of spillover effects (such as current account balance, gross fixed capital formation, credit, and inflation rate). And (*v*) All predictors with PIPs $\geq \dot{\tau}$ (in bold) will correspond to the ones to be accounted for the final solution.

Nevertheless, although the intense shrinkage in the parameter space, the final solution would still require

[31] Electricity production is expressed in % including coal, hydroelectric, natural gas, nuclear, and oil sources.

[32] The analysis focuses on recent studies concerning implications and interactions between fiscal and monetary policy in advanced and emerging economies (see, for instance, Pacifico (2019) and Pacifico (2020a,b)).

[33] The acronym PPP stands for Purchasing Power Parity and is used to measure prices at different locations controlling for cost differences.

[34] More precisely, 19 predictors refer to `Economic Status`, 8 predictors account for `Healthcare Statistics`, 16 predictors account for `Demographic and Environment Statistics`, and 10 predictors refer to `Economic-financial Issues`.

[35] The Posterior Inclusion Probabilities (PIPs) correspond to the sum of the PMPs in (9) for all (potential) model solutions (or combination of predictors) wherein a covariate – in the conditioning set (c_{it}) – has been included.

[36] The Conditional Posterior Sign (CPS) is defined to deal with the sign certainty, taking values close to 1 or 0 if a covariate in c_{it} has a positive or negative effect on the outcomes of interest, respectively.

Table 1: *Top Best* Candidate Predictors – MPROB (`second stage`)

Idx.	Predictor	Label	Unit	PIP (%)	CPS
	Economic Status				
1	Current Education Expenditure, Secondary	`edusec`	total exp. (%)	0.43	0.63
2	Employers, Total	`emto`	total emp. (%)	**46.75**	1.00
3	Employment to Population Ratio, 15+	`empo`	total pop. (%)	0.17	0.33
4	Foreign Direct Investment, Net Inflows	`fdinet`	% GDP	**16.41**	0.96
5	Labor Force Partecipation Rate, 15+	`labpar`	total pop. (%)	0.22	0.27
6	Labor Force, Total	`labtot`	logarithm (thousands)	**33.65**	0.68
7	Unemployment Change	`unem`	total labor force (%)	**65.51**	0.00
8	Wage and Salaried Workers	`wage`	total emp. (%)	**27.40**	0.91
	Healthcare Statistics				
9	Capital Health Expenditure	`cahe`	% GDP	**31.56**	1.00
10	Current Health Expenditure	`cuhe`	% GDP	**44.02**	0.37
11	Dom. Gen. Gov. Health Expenditure	`gghe`	% GDP	**41.04**	0.95
12	Dom. Gen. Gov. Health Expenditure	`hegg`	% gen. gov. exp.	**28.13**	1.00
13	Current Tobacco Use	`tobuse`	% adults (15+)	**17.37**	0.61
14	Alcohol Consumption per Capita	`alcuse`	logarithm (adults, 15+)	0.36	0.33
	Demographic and Environment Statistics				
15	CO2 Emissions, Total	`co2tot`	total (%)	**23.06**	0.16
16	Age Dependency Ratio	`arat`	working-age pop. (%)	**48.12**	0.44
17	Fertility Rate, Total	`frat`	births per woman	**35.43**	0.10
18	Death Race, Crude	`death`	per 1,000 people	0.15	0.06
19	Energy Imports, Net	`eneim`	energy use (%)	**28.31**	0.71
20	Population, Total	`pop`	logarithm (thousands)	0.23	0.47
21	Rural Population	`rural`	total pop. (%)	0.18	0.35
22	Urban Population	`urban`	total pop. (%)	**21.33**	0.51
23	School Enrollment, Secondary	`school`	total pop. (% net)	0.36	0.68
24	Human Capital Index	`hci`	working-age pop. [0-1]	0.32	0.81
	Economic–financial Issues				
25	Central Government Debt, Total	`debt`	% GDP	**37.87**	0.00
26	Current Account Balance	`cab`	% GDP	**67.31**	1.00
27	Domestic Credit, Financial Sector	`crefin`	% GDP	0.41	0.83
28	Gen. Gov. Final Consumption Exp.	`ggfce`	% GDP	0.24	0.75
29	Gross Fixed Capital Formation	`gfcf`	% GDP	**61.50**	0.92
30	Inflation, Consumer Prices	`inf`	% GDP	**63.24**	0.04
31	GDP Growth per Capita	`gdpg`	annual %	**74.45**	1.00
-	GDP per capita	`gdp`	PPP	-	-

The Table is so split: the first column denotes the predictor number; the second and the third column describe the predictors and the corresponding labels; the fourth column refers to the measurement unit; and the last two columns displays the PIPs (in %) and the CPS, respectively. The last row refers to the outcomes of interest. All contractions stand for: *exp.*, 'expenditure'; *emp.*, 'employment'; *pop.*, 'population'; *dom.*, 'domestic'; *gen. gov.*, 'general government'; and *int.*, international. All data refer to World Bank database.

more effort: indeed, there are 20 (potential) *top best* predictors better fitting the data. Thus, I run the third and last stage implicit in the MPROB procedure testing for panel Granger (Non-)Causality among all selected predictors. In Table 2, I display the only covariates with p-value $< \dot{\tau}$ and then included in the final submodels $M_{\xi^*} \subset \mathcal{E}$. All socioeconomic factors, including demographic and environment statistics, matter as much as economic–financial variables affecting outcomes (y_{it}) and then the need to be accounted for facing different triggering events (e.g., global financial and pandemic crises), but with some common features (see Section 5.3).

Finally, the final model solution better performing the data – with $lBF = 13.49$ according to (13) –

consists of 10 final *top best* covariates so split: predictors $(2,7)$ for $y^{c'}_{i,t-l}$; predictors $(10,11,16,17)$ for $z^{s'}_{i,t-l}$; and predictors $(26,29,30,31)$ for $z^{p'}_{i,t-l}$). All their available lags – including lagged outcomes ($y^{o'}_{i,t-l}$) and all predictors in $M_{\xi} \subset \mathcal{E}$ with PIPs $\geq \dot{\tau}$ (Table 1, in bold) – are then included as *external* instruments in the estimating procedure. In the estimation method, I also include two time-invariant effects (x_{1t} and x_{2t}) denoting the presence of structural breaks in 2018 (due to the global financial crisis) and in 2020 (due to the COVID-19 pandemic).

Table 2: Granger (Non-)Causality Test – MPROB (`third stage`)

`From` c^*_{it} `to` y_{it}	`emto`	`unem`	`cuhe`	`gghe`	`arat`	`frat`	`cab`	`gfcf`	`inf`	`gdpg`
Z-tilde	5.40 (0.00)	4.56 (0.00)	5.85 (0.00)	4.08 (0.00)	5.17 (0.00)	3.66 (0.00)	5.39 (0.00)	4.04 (0.00)	3.42 (0.00)	5.98 (0.00)
`From` y_{it} `to` c^*_{it}	`emto`	`unem`	`cuhe`	`gghe`	`arat`	`frat`	`cab`	`gfcf`	`inf`	`gdpg`
Z-tilde	7.59 (0.00)	3.45 (0.00)	3.10 (0.00)	2.50 (0.01)	4.21 (0.00)	3.23 (0.00)	5.03 (0.00)	2.48 (0.01)	6.05 (0.00)	3.44 (0.00)

The Table displays all Z-tilde test statistics and p-values (in parenthesis) on the panel Granger (Non-)Causality test. Here, c^*_{it} stands for the *top best* final candidate predictors in $M_{\xi^*} \subset \mathcal{E}$ with higher lBF (equation (13)).

In Table 3, I display the main estimation outputs – including diagnostic tests – highlighting the performance of the DPB-CRE model. Here, some considerations are in order. (*i*) The best optimal lag chosen according to Arellano (2003)'s test is 3. (*ii*) All estimates are consistent and valid, showing no autocorrelation among residuals and highly strong linear dependencies; thus, endogeneity issues and model misspecification problems need to be accounted for. (*iii*) The posterior predictive variance of μ_i reenters in the range displayed in (33), dealing with high dimensional data carefully ($\mathbb{V}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi(\bar{m})}[\mu_i] = 0.74$ with $\xi^* > 10$). (*iv*) In Table 2, highly strong causal links confirms the presence of heterogeneity among a restricted subgroup of units (alternative hypothesis in the Granger (Non-)Causality test). (*v*) The Posterior Model Size Distribution (PMSD) is close to 10 and then to the *top best* candidate predictors better explaining the data ($\dot{\chi}$). As emphasized in this analysis, let $\dot{\chi}$ be 20 (Table 1), there would be some covariates not fitting the data well (absence of the covariate in $M_{\xi^*} \subset \mathcal{E}$), and thus they should be not considered in the final solution (Table 2). And (*iv*) the latter is robust dealing with most of the explained variability of the outcomes of interest ($R^2_{adj.} = 0.78$). To prove it, I run two different DPD models accounting for the only time- and country-specific effects, obtaining a robustness equals to 0.98 and 0.97, respectively. The higly large $R^2_{adj.}$ would highlight the presence of heterogenous effects strongly affecting the data and then the need to be dealt with.

5.3 Individual–specific Forecasting and Future Policy-relevant Strategies

Dynamic analyses have been conducted via accurate MCMC algorithms and implementations. I use $1,000$ until $5,000$ draws and find that convergence is obtained at about $1,000$ draws[37]. The total number of draws has been $2,000 + 3,000 = 5,000$, which corresponds to the sum of the final number of draws to discard

[37]The convergence has been found by amounting to about 1.2 draws per regression parameter.

Table 3: Estimation Outputs and Diagnostic Tests

Main Statistics	Results
$AR(l^*)$	3
ξ^*	10
LGB_s	0.00
LGB_r	0.91
$\mathbb{V}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi(\bar{m})}[\mu_i]$	0.74
$PMSD$	9.92
lBF	13.49
$R^2_{adj.}$	0.78

The Table shows the main estimation outputs and diagnostic tests of the DPB-CRE model in (2). Here, l^* denotes the optimal lag; LGB_s and LGB_r stand for Ljung-Box test statistics of the series and residuals in terms of p-values, respectively; $\mathbb{V}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi(\bar{m})}[\mu_i]$ accounts for the estimated posterior predictive variance for μ_i; PMSD refers to the Posterior Model Size Distribution; lBF refers to the log Bayes Factor in (13); and $R^2_{adj.}$ denotes the adjusted R^2 measuring the robustness.

and save, respectively. A total of $1,000$ retained replications has been used to conduct posterior inference at each t. The outcomes absorb the conditional forecasts computed for a time frame of 2 years in order to also investigate interesting findings concerning the impact of an ongoing pandemic crisis on the global economy, or the evolution of the productivity in the next years. The natural conjugate prior refers to three subsamples: (i) 2007–2009 to evaluate the impact of the Great Recession; (ii) 2010–2018 to adress how fiscal consolidation periods affected the dynamics of the productivity among countries; and (iii) 2019–2020 to assess the impact of the current pandemic crisis on the global economy.

Without restrictions, the estimation sample amounts to 682 regression parameters: every estimates of the DPB-CRE in (2) account for 22 country indices and 31 time periods. Let the Assumptions (4.1)-(4.3) hold and the hyperparameters in $\tilde{\omega}$ be all known and estimable, posterior distributions are computed according to equations (37)-(39) for $\theta_t|\theta_{t-l}, \hat{y}_{i,T+k}$, (42)-(43) for moment distributions in $\mu_i|\hat{y}_{i,T+k}$ given initial values $(\hat{y}_{i0}|\hat{\mu}_i^{\bar{m}})$, (44) for the final *top best* parameter space, and (45) for $u_t|\hat{y}_{i,T+k}$. All data are expressed in standard deviations.

In Figure 1, (empirical) forecasts for outcomes $\hat{y}_{i,T+k}$ – with individial-specific (μ_i) and time-fixed (α) effects – are drawn for advanced (top plot) and emerging (bottom plot) economies. They aim to investigate potential determinants performing suitable policy strategies because of 'dramatic' structural breaks (e.g., recessions, epidemics) on the global economy. By construction, incidental parameters, endogeneity issues, and structural model uncertainty are dealt with. In addition, cross-unit interdependencies and commonality, dynamic feedback, and causal interactions are also addressed. The yellow and red curves denote the 95%

confidence bands, and the blue and purple curves denote the conditional[38] and unconditional[39] projections of outcomes $\hat{y}_{i,T+k}$ for each N country indices and T time periods.

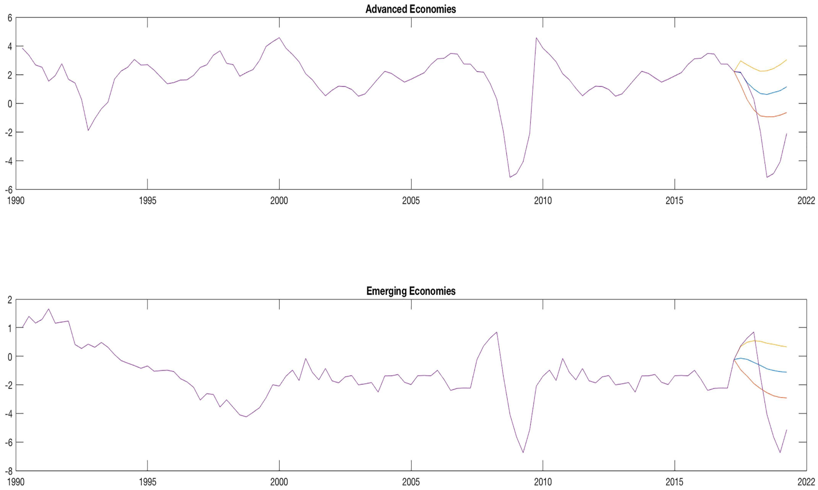

Figure 1: The plot draws (empirical) forecasts for outcomes $\hat{y}_{i,T+k}$ given individual–specific (μ_i) and time-fixed (α) effects given a pool of socioeconomic–demographic, real–financial, and policy determinants. All time-varying parameters are posterior means by semiparametric estimating procedure. They correspond to conditional (blue line) and unconditional (purple line) projections of each supposed variable assessed in (2).

From a modelling perspective, four main findings are addressed. (i) Even if there has been evidence of significant co-movements and interdependencies among countries, consistent heterogeneities matter in both the spreading and the intensity of countries' dynamics. Thus, the need for forecasters and policymakers to account for heterogeneous effects (correlated random coefficients) when formulating policy strategies and forecasting in multivariate dynamic panel data. (ii) Conditional projections lie in the confidence interval; conversely, unconditional projections tend to diverge over T. Thus, when studying macroeconomic–financial and socioeconomic–demographic issues, cross-unit lagged interdependencies – along with dynamic feedback and interactions – have to be dealt with. (iii) Outward countries' responses in advanced economies (net senders) emphasize consistent economic–institutional implications, while emerging economies show inward responses (net receivers) due to stringent interlinkages (e.g., capital flows and trade exposures) with Western European countries. Thus, although recent dynamics would suggest significant improvements in fiscal sustainability (e.g., during post-crisis periods), the risk of a cascade of policy errors, adverse political economy incentives, and divergence in financial integration are relevant issues for an early and coordinated fiscal consolidation. The results highlight a great caution in efforts to fine-tune the economy via fiscal structures

[38] Generally, the conditional projection in forecasting models is the one that the model would have obtained over the same period conditionally on the actual path of unexpected dynamics for that period (μ_i dependent on y_{i0}).

[39] Generally, the unconditional projection in forecasting models is the one that the model would obtain for output growth for that period only on the basis of historical information, and it is consistent with a model-based forecast path for the other variables (μ_i independent of y_{i0}).

and boost productivity to potential growth via accurate structural reforms and policy adjustments. In that context, the *USA*, *JPN*, and *CHN* variables would act like persistent net senders in driving the transmission of international financial shocks and net receiver in allowing shocks to spill over into real economy among European and non-European countries. And (*iv*) the highly strong significance in socioeconomic–demographic factors show that healthcare system capacity and cost-related indices would affect government's strategy and measures (Table 1). For instance, they could delay well-timed measures because of generating an overconfidence in the government's capacity to fight an unexpected outbreak. However, a hint of boosting productivity to potential growth can be observed among countries in the next years, mainly among advanced economies. Thus, in the current pandemic crisis – a game against nature with incomplete information, increased knowledge, and reduced uncertainty on other countries' policy responses and epidemic development – health services and expected economic costs from hard measures managing the recession have increased the agility of the country's policy actions.

From a policy perspective, the estimation sample considered in this study deals with two 'severe' global economic crisis causing a recession/depression in most countries: global financial crisis and COVID-19 pandemic. Even if they are distinct between them, potential common linkages matter such as sovereign debt accumulation and possible cuts in public health spending, the slowing of economic growth and labour mobility, 'bank zombie'[40], and difficulties for designing and implementing economic support policies. According to Figure 1, four main results can be highlighted. (*i*) Empirical forecasts show that most European emerging economies are strongly exposed to financial interlinkages and then highly dependent on other European countries (e.g., Western European countries). Nevertheless, the presence of persistent heterogeneities among countries' responses emphasize the need for accelerating financial development in developing countries, stimulating domestic resource mobilization, and supporting consistent reforms of the international financial system in order to boost investment and growth. (*ii*) Even if several measures have already been taken at the international and EU levels, most countries have been limited to use monetary and fiscal tools effectively due to stringent economic–institutional interdependencies, and then they have not been able to deploy conventional consolidation measures during triggering events. Moreover, most countries have failed to control the extent of COVID-19 due to people's attitudes of denial and misunderstanding of social distancing for the control of the outbreak. Thus, in a context of radical uncertainty and heterogeneous territorial effects, appropriate policy measures need to be addressed at the local level rather than globally. (*iii*) Heterogeneity among countries' responses would cover different reasons. It could reflect the fact that larger government health spending (socioeconomic factors) implies a more developed economy, which in turn suggests an older population and better social welfare (demographic and environment statistics). In addition, larger government health expenditure could matter due to inflated costs (economic–financial issues) and then not necessarily reflecting the quality of the public health-care system. Also, countries with lower health-care expenditure and/or weaker public health-care system could be aware of their deficiencies

[40]Generally speaking, a 'zombie bank' is an insolvent financial institution able to continue operating because of the government's support.

and then act more aggressively and early in the outbreaks (such as most of the Central and Eastern European countries). And (*iv*) most policy adjustments – applied by governments during a recession – have followed distinct national rather than consensual international standards (such as in the current outbreak and previous pandemic crises). Overall, all policy tools should be implemented by closely monitoring the evolution of the economic status for each country and coordinating country-specific European and international measures. Then, a participatory government is needed for ensuring more resilient and robust health systems and improving public health outcomes so as to safeguard against an ongoing pandemic crisis on the global economy.

6 Relative Regrets for Tweedie Correction via MCMC-based Experiments

In this example, the performance of the estimation method is investigated by summarizing the regrets for Tweedie correction in (29) – relative to the posterior predictive variance of μ_i for optimal point forecasts as specified in Theorems (4.6) and (4.7) through MCMC-based simulations. More precisely, according to (33), I consider three sequences of $\left(N, \xi^*, \mathbb{V}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi(\bar{m})}[\mu_i]\right)$ with correlated random coefficients homoskedastic case to evaluate different improvements in the forecasting performance: (*i*) $(10000, 15, 1.0)$, heterogeneity with high dimension; (*ii*) $(10000, 10, 0.5)$, sufficient-homogeneity with moderate dimension; (*iii*) $(10000, 5, 0.0)$, near-homogeneity with small dimension (Table 4). I suppose a basic DPD model with $\alpha = \gamma = 0$, homoskedastic variance $\sigma^2 = 1$, and regimes $\bar{m} = 1$ (just one common individual–specific effect across units).

Table 4: MCMC-based Designs

Law of Motion	$y_{it} = \mu_i + \beta y_{i,t-1} + u_{it}$	where	$\beta = 0.5$	,	$u_{it} \sim i.i.d.N(0,1)$
Initial Observations		$y_{i0} \sim N(0,1)$			
Correlated Random Effects	$\mu_i \mid y_{i0} \sim N(0, \Psi_{\mu_i})$	where	$\Psi_{\mu_i} \sim IG\left(\frac{0.1}{2}, \frac{0.01}{2}\right)$		

The Table shows the three sequences of $\left(N, \xi^*, \mathbb{V}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi(\bar{m})}[\mu_i]\right)$ with correlated random coefficients homoskedastic case conducted in the simulated example through MCMC-based simulations according to (33).

In this analysis, I include two additional empirical Bayes estimators dealing with alternative Tweedie corrections: Kernel Density (KD) estimator and NonParametric Maximum Likelihood (NPML) estimation. Here, some considerations are in order. The Liu et al. (2020)'s analysis develops a KD estimator dealing with the problem of forecasting a collection of short time-series processes through the cross-sectional information in a dynamic panel data. The authors construct a nonparametric kernel estimate of the Tweedie correction, showing its asymptotic equivalence to the risk of an empirical predictor treating the CREs' distribution as known. Concerning the NPML estimation, the EB estimator is constructed by specifying appropriate bounds for the domain of CREs and then partitioning them into a predetermined set of bins (see, for instance, Gu and Koenker (2017b).).

Table 5 provides the relative regrets for FMM distributions and alternative Tweedie corrections for all of the three supposed MCMC-based designs. The best choice of ϑ_0 improving the forecasting performance in terms of ratio-optimality was set 0.5 (middle point in an arbitrary range [0.1 - 0.9]). The findings highlight the usefulness of the MPROB procedure for dramatically shrinking the model size with high dimensional data, and the performance of the FMM-based Empirical Bayes estimator (Tweedie correction) for performing better forecasts. For instance, lower posterior predictive variances of μ_i are associated to less relative regrets. Compared to KD and NPLM estimates, FMM with MPROB procedure shows lower regrets. Thus, the DPB-CRE model would be able to perform better accurate forecasts dealing with (potential) semiparametric problems when addressing heterogeneous effects in dynamic panel setups (through CR coefficients) and investigating dynamic feedback and interactions in high dimensional time-varying data (through MPROB procedure). Furthermore, the experiment was replicated accounting for highly larger sample size (e.g., $N = 100,000$) and lower ratio-optimality (e.g., $\vartheta_0 \cong 0.1$). The results show that the relative regrets are negatively correlated with the number of cross-sectional units N, and less ratio-optimality – even if reduces computational costs – would suffer to higher associated regrets.

Table 5: Relative Regrets for Tweedie Corrections for MCMC-based Designs

	Design I	Design II	Design III
N	10000	10000	10000
$\mathbb{V}^{\mathcal{Y}_i,\mu_i}_{\theta,\phi,\pi^{\bar{m}}}[\mu_i]$	1.0	0.5	0.0
ξ^*	15	10	5
N_{sim}	10000	10000	10000
KD	0.026	0.051	0.074
FMM	0.014	0.010	0.007
NPML	0.021	0.019	0.013

Relative regrets for Tweedie corrections for all of the three supposed MCMC-based designs are provided. The regret is standardized by the average posterior predictive variance of μ_i, with $\vartheta_0 = 0.5$.

All results in Table 5 find confirmation in Figure 2. More precisely, lower posterior predictive variances of μ_i are associated to less MSE and then better accuracy forecasts (associated with less relative regrets). Moreover, the (designed) joint density distribution of $\pi^{\bar{m}}$ – depicting posterior draw samples of the empirical distribution of $\hat{\mu}_i$ (Statement (4.6.1)) – asymptotically converges to a Normal and then the FMM-based Tweedie Correction – in Theorem (4.6) – approaches linear distribution function. Furthermore, in the second and third designs, the empirical realizations of $\hat{\mu}_i$ are greater and lie in the distribution highlighting lower MSEs and less sampling variance in the estimated posterior means.

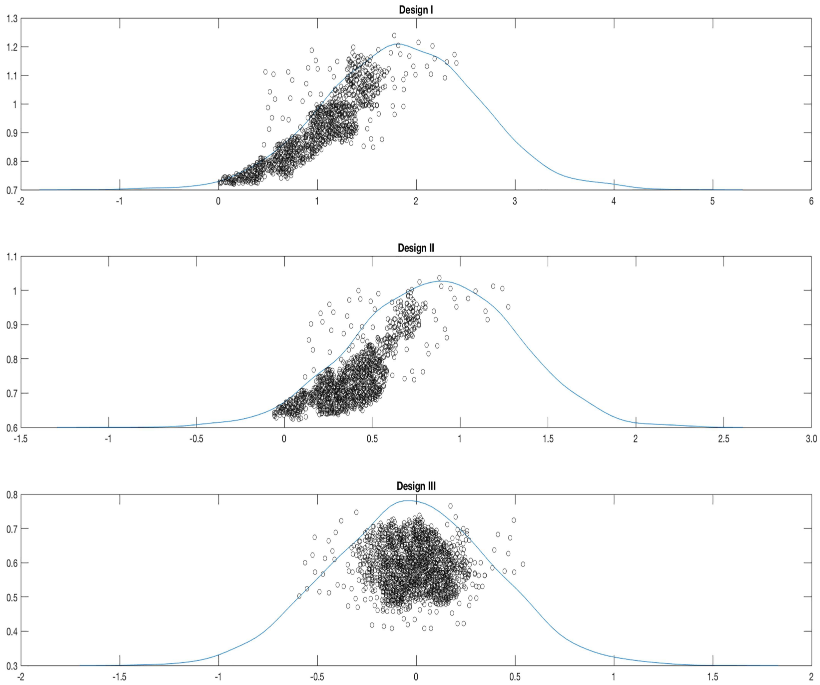

Figure 2: The panels show the MSE associated to all of the three supposed MCMC-based designs. The solid lines display the posterior draw samples of the empirical distribution of $\hat{\mu}_i$ according to the (designed) joint density distribution $\pi^{\tilde{m}}$ in Statement (4.6.1), and the FMM-based Tweedie Correction in Theorem (4.6).

7 Concluding remarks

This study aims to construct and develop a methodology to improve the recent literature on DPD models when dealing with (i) individual–specific forecasts, (ii) Bayesian analyses with parametric priors on heterogeneous parameters, (iii) ratio-optimality and posterior consistency in dynamic panel setups, (iv) empirical Bayes estimator and alternative Tweedie corrections, and (v) the curse of dimensionality when estimating time-varying data. The contributions of this paper are fourfold. First, I develop a hierarchical structural Bayesian approach to deal with potential features in real-world data such as non-linearity, incidental parameters, endogeneity issues, and misspecified dynamics. Second, I use a MPROB procedure to obtain a final subset containing *top best* candidate predictors better fitting the data dealing with high dimensional panel setups. Third, MCMC algorithms and implementations are addressed to set potential determinants performing future policy strategies to prevent the emergence of epidemics on the global economy. Fourth, all available lags of the *top best* candidate predictors are included as *external* instruments, and unobserved individual heterogeneities are treated as random variables and then possibly correlated with the outcomes.

An empirical application on a pool of advanced and emerging economies is assessed describing the functioning and the performance of the methodology. The estimation sample refers to the period $1990-2020$, covering a sufficiently large sample to address potential causal links and interdependencies between outcomes and a set of time-varying factors, including heterogeneous individual-specific and time-fixed effects. A simulated experiment – compared to related works – is also addressed to highlight the performance of the estimating procedure through Monte Carlo simulations.

Future works would implement the methodology proposed in this study accounting for dynamic panel data with cross-sectional heterogeneity and time-varying heteroskedasticity. The latter is useful to disentangle the uncertainty generated from either unknown cross-sectional heterogeneity (μ_i) or independent shocks (u_{it}), i.e. a deconvolution problem. In that context, the semiparametric Bayesian approach can be implemented to perform density forecasts. Methodologically, for correlated random coefficients, the algorithm proposed in (29) can be extended to a nonparametric conditional density estimation problem using stick-breaking process mixture priors.

Compliance with Ethical Standards

Funding: No fund was used to conduct this study.

Conflict of Interest: The author declares no conflict of interest.

References

Alvarez, J. and Arellano, M. (2003). The time series and cross-section asymptotics of dynamic panel data estimators. *Econometrica*, 71(4):1121–1159.

Anderson, T. W. and Hsiao, C. (1981). Estimation of dynamic models with error components. *Journal of the American Statistical Association*, 76(375):598–606.

Arellano, M. (2003). Panel data econometrics. *Oxford University Press, New York.*

Arellano, M. and 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.

Arellano, M. and Bonhomme, S. (2011). Nonlinear panel data analysis. *Annual Review of Economics*, 3:395–424.

Arellano, M. and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1):29–51.

Arellano, M. and Hahn, J. (2007). Understanding bias in nonlinear panel models: Some recent developments. *Advances in Economics and Econonometrics: Theory and Applications, Ninth World Congress*, ed. by W. N. R. Blundell, and T. Persson. Cambridge University Press, Cambridge.

Arellano, M. and Hahn, J. (2016). A likelihood-based approximate solution to the incidental parameter problem in dynamic nonlinear models with multiple effects. *Global Economic Review*, 45(3):251–274.

Arellano, M. and Honore, B. (2001). Panel data models: Some recent developments. *Handbook of Econometrics*, ed. by J. Heckman, and E. Leamer, 5:3229–3296.

Bernanke, B. S. (2020). The new tools of monetary policy. *American Economic Review*, 110(4):943–983.

Bester, C. A. and Hansen, C. (2009). A penalty function approach to bias reduction in non-linear panel models with fixed effects. *Journal of Business and Economic Statistics*, 27(2):131–148.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115–143.

Boulos, M., Kamel, N., and Geraghty, E. M. (2020). Geographical tracking and mapping of coronavirus disease covid-19/severe acute respiratory syndrome coronavirus 2 (sars-cov-2) epidemic and associated events around the world: How 21st century gis technologies are supporting the global fight against outbreaks and epidemics. *International Journal of Health Geographics*, 19(8):1–12.

Brown, L. D. and Greenshtein, E. (2009). Nonparametric empirical bayes and compound decision approaches to estimation of a high-dimensional vector of normal means. *The Annals of Statistics*, 37:1685–1704.

Carro, J. (2007). Estimating dynamic panel data discrete choice models with fixed effects. *Journal of Econometrics*, 140(2):503–528.

Chamberlain, G. (1984). Panel data. *Handbook of Econometrics*, ed. by Z. Griliches, and M. D. Intriligator, 2:3847–4605.

Chamberlain, G. (2010). Binary response models for panel data: Identification and information. *Econometrica*, 78:159–168.

Chamberlain, G. and Hirano, K. (1999). Predictive distributions based on longitudinal earnings data. *Annales d'Economie et de Statistique*, 55-56:211–242.

Compiani, G. and Kitamura, Y. (2016). Using mixtures in econometric models: a brief review and some new results. *The Econometrics Journal*, 19(3):C95–C127.

Dumitrescu, E. and Hurlin, C. (2012). Testing for granger non-causality in heterogeneous panels. *Economic Modelling*, 29(4):1450–1460.

Efron, B. (2011). Tweedie's formula and selection bias. *Journal of the American Statistical Association*, 106(496):1602–1614.

Fernandez-Val, I. (2009). Fixed effects estimation of structural parameters and marginal effects in panel probit models. *Journal of Econometrics*, 150(1):71–85.

Gelfand, A. E. and Dey, D. K. (1994). Bayesian model choice: Asymptotics and exact calculations. *Journal of the Royal Statistical Society: Series B*, 56(3):501–514.

George, E. I. and Foster, D. P. (2000). Calibration and empirical bayes variable selection. *Biometrika*, 87:731–747.

Gössling, S., Scott, D., and Hall, C. M. (2020). Pandemics, tourism and global change: A rapid assessment of covid-19. *Journal of Sustainable Tourism*, 29(1):1–20.

Gu, J. and Koenker, R. (2017a). Empirical bayesball remixed: Empirical bayes methods for longitudinal data. *Journal of Applied Economics*, 32(3):575–599.

Gu, J. and Koenker, R. (2017b). Unobserved heterogeneity in income dynamics: An empirical bayes methods for longitudinal data. *Journal of Business and Economic Statistics*, 35(1):1–16.

Hahn, J. and Kuersteiner, G. (2011). Bias reduction for dynamic nonlinear panel models with fixed effects. *Econometric Theory*, 72:1295–1319.

Hahn, J. and Newey, W. (2004). Jackknife and analytical bias reduction for nonlinear panel models. *Econometrica*, 72:1295–1319.

Hassan, M. M., Kalam, M. A., Shano, S., K., N. M. R., Rahman, M. K., Khan, S. A., and Islam, A. (2020). Assessment of epidemiological determinants of covid-19 pandemic related to social and economic factors globally. *Journal of Risk and Financial Management*, 13(194):1–14.

Hirano, K. (2002). Semiparametric bayesian inference in autoregressive panel data models. *Econometrica*, 70(2):781–799.

Jacquier, E., Polson, N., and Rossi, P. (1994). Bayesian analysis of stochastic volatility. *Journal of Business and Economic Statistics*, 12:371–417.

Jiang, W. and Zhang, C.-H. (2009). General maximum likelihood empirical bayes estimation of normal means. *The Annals of Statistics*, 37(4):1647–1684.

Karwowski, M., Kowal, M., Groyecka, A., Białek, M., Lebuda, I., Sorokowska, A., and Sorokowski, P. (2020). When in danger, turn right: Does covid-19 threat promote social conservatism and right-wing presidential candidates. *International Society for Human Ethology*, 35:37–48.

Kiefer, J. and Wolfowitz, J. (1956). Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. *The Annals of Mathematical Statistics*, 27(4):887–906.

Lancaster, T. (2002). Orthogonal parameters and panel data. *The Review of Economic Studies*, 69(3):647–666.

Laskowski, M., Mostaço-Guidolin, L. C., Greer, A. L., Wu, J., and Moghadas, S. M. (2011). The impact of demographic variables on disease spread: Influenza in remote communities. *Scientific Reports*, 1(105):1–7.

Levine, R. A. and Casella, G. (2014). Implementations of the monte carlo em algorithm. *Journal of Computational and Graphical Statistics*, 10(3):422–439.

Lindahl, J. F. and Grace, D. (2015). The consequences of human actions on risks for infectious diseases: A review. *Infection Ecology & Epidemiology*, 5(1):1–11.

Liu, L. (2018). Density forecasts in panel data models: A semiparametric bayesian perspective. *Working Paper*, Cornell University:1–32. Available at https://arxiv.org/abs/1805.04178.

Liu, L., Moon, H. R., and Schorfheide, F. (2019). Forecasting with a panel tobit model. *NBER Working Papers*, 26569.

Liu, L., Moon, H. R., and Schorfheide, F. (2020). Forecasting with dynamic panel data models. *Econometrica*, 88(1):171–201.

Mikolai, J., Keenan, K., and Kulu, H. (2020). Intersecting household-level health and socio-economic vulnerabilities and the covid-19 crisis: An analysis from the uk. *SSM - Population Health*, 12:1–9.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49(6):1417–1426.

Nicola, M., Alsafi, Z., Sohrabi, C., Kerwan, A., Al-Jabir, A., Iosifidis, C., Agha, M., and Agha, R. (2020). The socio-economic implications of the coronavirus pandemic (covid-19): A review. *International Journal of Surgery*, 78(6):185–193.

Pacifico, A. (2019). International co-movements and business cycles synchronization across advanced economies: A spbvar evidence. *International Journal of Statistics and Probability*, 8(4):68–85.

Pacifico, A. (2020a). Fiscal implications, misspecified dynamics, and international spillover effects across europe: A time-varying multicountry analysis. *International Journal of Statistics and Economics*, 21(2):18–40.

Pacifico, A. (2020b). International macroeconomic-financial linkages and policy interactions in time-varying multicountry panel setups: An application to emerging economies. *EconModels, Journal of Policy Modeling (JPM) Database*, pages 1–31. Available at http://www.econmodels.com.

Pacifico, A. (2020c). Robust open bayesian analysis: Overfitting, model uncertainty, and endogeneity issues in multiple regression models. *Econometric Reviews*, 40(2):148–176.

Pacifico, A. (2021). Structural panel bayesian var with multivariate time-varying volatility to jointly deal with structural changes, policy regime shifts, and endogeneity issues. *Forthcoming in Econometrics.* Available at https://mpra.ub.uni-muenchen.de/104292/.

Robbins, H. (1964). The empirical bayes approach to statistical decision problems. *The Annals of Mathematical Statistics*, 35:1–20.

Robert, C. (1994). The bayesian choice. *New York: Springer Verlag.*

Scott, J. G. and Berger, J. O. (2010). Bayes and empirical-bayes multiplicity adjustment in problem. *The Annals of Statistics*, 38:2587–2619.

She, J., Jiang, J., Ye, L., Hu, L., Bai, C., and Song, Y. (2020). 2019 novel coronavirus of pneumonia in wuhan, china: Emerging attack and management strategies. *Clinical and Translational Medicine*, 9:1–7.

Sorokowski, P., Groyecka, A., Kowal, M., Sorokowska, A., Białek, M., Lebuda, I., Dobrowolska, M., Zdybek, P., and Karwowski, M. (2020). Can information about pandemics increase negative attitudes toward foreign groups? a case of covid-19 outbreak. *Sustainability*, 12(12):1–10.

Stojkoski, V., Utkovski, Z., Jolakoski, P., Tevdovski, D., and Kocarev, L. (2020). The socio-economic determinants of the coronavirus disease (covid-19) pandemic. *Cornell University, Working Paper, 2020*, pages 1–30.

Temsah, M.-H., Al-Sohime, F., Alamro, N., Al-Eyadhy, A., Al-Hasan, K., Jamal, A., Al-Kamlouth, I., Aljamaan, F., Al-Amri, M., Barry, M., Al-Subaie, S., and Somily, A. M. (2020). The psychological impact of covid-19 pandemic on health care workers in a mers-cov endemic country. *Journal of Infection and Public Health*, 13(6):877–882.

Tuite, A. R. and Fisman, D. N. (2020). Reporting, epidemic growth, and reproduction numbers for the 2019 novel coronavirus (2019-ncov) epidemic. *Annals of Internal Medicine.*

Wu, J. T., Leung, K., and Leung, G. M. (2020). Nowcasting and forecasting the potential domestic and international spread of the 2019-ncov outbreak originating in wuhan, china: A modelling study. *The Lancet*, 395(10225):689–697.

Zaki, N., Hany, A., and Sahar, I. (2020). Association of hypertension, diabetes, stroke, cancer, kidney disease, and high-cholesterol with covid-19 disease severity and fatality: A systematic review. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 14:1133–1142.

Zheng, J. (2020). Sars-cov-2: An emerging coronavirus that causes a global threat. *International Journal of Biological Sciences*, 16(10):1678–1685.