Spatial disaggregation of time series

A. Tobar[1], A. Mir[1], R. Alberich[1], I. Garcia[1], M. Miró[1], D. Miranda[1], NA. Cruz[1],

Abstract

Vamos a escribir

Keywords:
Spatial models, Spatial autoregressive model, heteroscedasticity, Spatial regression, Maximun likelihood estimation

1. Introduction

Modeling spatio-temporal data has gained increasing relevance in econometrics and regional science, due to the simultaneous presence of spatial spillovers and temporal dynamics in key macroeconomic indicators. Autoregressive spatial models (SAR) have long been used to capture spatial autocorrelation among neighboring regions, improving inference and prediction by explicitly accounting for spatial dependencies (LeSage and Pace, 2015; Cressie and Wikle, 2011). Recent developments have extended these models to incorporate temporal structure, yielding rich spatio-temporal frameworks that model both forms of dependence jointly (Gelfand et al., 2019; Hsu et al., 2021).

A particularly relevant application of this framework is the temporal disaggregation of macroeconomic aggregates, such as GDP, across subnational units. In many countries, national statistical agencies publish annual regional GDP figures, while only national totals are available quarterly. For policy evaluation, business cycle analysis, and regional monitoring, there is a strong demand for consistent quarterly estimates at the regional level. Benchmarking techniques, such as the Chow–Lin method (Chow and Lin, 1971) and its Eurostat-endorsed variants (Eurostat, 2013), have been extensively applied for temporal disaggregation using high-frequency auxiliary indicators. However, these approaches do not incorporate the spatial dependence between regions, potentially missing important information.

Recent studies have started to explore the integration of spatial elements in disaggregation and forecasting, including machine learning and autoregressive alternatives (Cuevas et al., 2015; Han and Howe, 2024), but often lack a fully econometric or interpretable structure. In this context, it is desirable to develop models that retain the inferential transparency of classical methods while accounting for spatial dependence and irregular availability of regional data.

This paper contributes to the literature by proposing a new methodology for the spatio-temporal disaggregation of economic aggregates, combining spatial autoregressive modeling, benchmarking constraints, and covariate information. Before presenting our approach, it is important to highlight several features that make it both theoretically and practically appealing. First, the method is supported by rigorous asymptotic theory: we establish global identifiability and asymptotic normality of the joint estimator under general conditions that do not require Gaussianity. This makes the model robust to the non-normal and heteroskedastic nature of economic residuals. Second, we introduce a novel anchoring mechanism that incorporates partially observed regional values as linear constraints. This step mimics real-world situations where sporadic regional data (e.g., for capital cities or large communities) are available and helps reduce the variance of predictions. Third, we conduct an extensive simulation study encompassing 74,088 parameter configurations, covering a wide range of spatial autocorrelations, covariate informativeness, and grid sizes. These simulations confirm the robustness and stability of our approach across diverse scenarios. Finally, we demonstrate the practical relevance of our method with a case study on the disaggregation of Spanish GDP across the 17 autonomous communities from 2002 to 2023, using a rich set of auxiliary indicators and principal component analysis for dimensionality reduction.

In doing so, our methodology extends the classical Chow–Lin approach to the spatial domain, allowing for flexible and accurate estimation of quarterly regional GDP consistent with national aggregates, while accounting for both spatial spillovers and the availability of partial observations.

2. Spatial autoregressive model

Let Y_t be a time series measured at time $t = 1, \ldots, T$ over the spatial unit $\mathcal{R}$, which is subdivided in n spatial regions $\{\mathcal{R}_1, \ldots, \mathcal{R}_n\}$; let $\mathbf{z}_{it}$ a set of k known explicative variables measured at time t in spatial region i. Let

Y_{it} the ith unknown time serie disagregated at time t over the region i with the restriction defined as:

$$\sum_{i=1}^{n} Y_{it} = Y_t, \qquad \forall t = 1, \ldots, T \tag{1}$$

The objective is obtain a estimated values for Y_{it} using the information allocate both in $\mathbf{z}_{it}$ and $\mathbf{Y}_a = (Y_1, \ldots, Y_T)$. A spatial autoregressive model is supposed for the vector $\mathbf{Y}_t = (Y_{1t}, \ldots, Y_{nt})$, the completed model is given by:

$$Y_{it} = \rho \sum_{j=1}^{n} w_{ij} Y_{jt} + \mathbf{z}_{it}^{\top} \boldsymbol{\beta} + u_{it} \tag{2}$$

where $\epsilon_t \sim N(0, \sigma^2)$, $|\rho| < 1$, and

$$\mathbf{Z} = \begin{bmatrix} \mathbf{z}_{11}^{\top} \\ \mathbf{z}_{21}^{\top} \\ \vdots \\ \mathbf{z}_{n1}^{\top} \\ \mathbf{z}_{12}^{\top} \\ \vdots \\ \mathbf{z}_{n2}^{\top} \\ \vdots \\ \mathbf{z}_{nT}^{\top} \end{bmatrix}_{nT \times k}, \mathbf{Y} = \begin{bmatrix} \mathbf{Y}_{11} \\ \mathbf{Y}_{21} \\ \vdots \\ \mathbf{Y}_{n1} \\ \mathbf{Y}_{12} \\ \vdots \\ \mathbf{Y}_{n2} \\ \vdots \\ \mathbf{Y}_{nT} \end{bmatrix}_{nT \times 1}, \mathbf{W} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix}, \mathbf{U} = \begin{bmatrix} \mathbf{u}_{11} \\ \mathbf{u}_{21} \\ \vdots \\ \mathbf{u}_{n1} \\ \mathbf{u}_{12} \\ \vdots \\ \mathbf{u}_{n2} \\ \vdots \\ \mathbf{u}_{nT} \end{bmatrix}_{nT \times 1} \tag{3}$$

For the vector $\mathbf{U}_i = (u_{i1}, \ldots, u_{iT})$, an $ARIMA(1,0,0)$, i.e. $\phi_1(B)u_{it} = \epsilon_{it}$, and $\epsilon_{it} \sim (0, \sigma^2)$ is proposed for modelling the unobservable time series, similar to the proposal of ?, Quilis (2018) and Bisio (2024). This assumption over $\mathbf{U}_i$ allow defined the covariance matrix of $\mathbf{U}_i$ as:

$$\boldsymbol{\Sigma}_U = Var(\mathbf{U}_i) = \left\{ \frac{\sigma^2 \phi_1^{|i-j|}}{1 - \phi_1^2} \right\}_{T \times T} \tag{4}$$

where ϕ is the autoregressive parameter of the $ARIMA(1,0,0)$ and σ^2 their associated variance. Then, the model defined in Equation (2) can be ex-

pressed as:

$$\mathbf{Y} = \rho(\mathbb{I}_T \otimes \mathbf{W})\mathbf{Y} + \mathbf{Z}\boldsymbol{\beta} + \mathbf{U} \tag{5}$$

the problem is that $\mathbf{Y}$ and $\mathbf{U}$ are unobservable by the characteristics of the measure. But $\mathbf{Z}$ and Y_t are observable and known, then, defined the matrix $\mathbf{C}$ as:

$$\mathbf{C} = \mathbb{I}_T \otimes \mathbf{1}_n^\top \tag{6}$$

such that $\mathbf{CY} = \mathbf{Y}_a$, then Equation (5) can be expressed as:

$$\mathbf{Y} = [\mathbf{I}_{nT} - \rho(\mathbf{I}_T \otimes \mathbf{W})]^{-1}\mathbf{Z}\boldsymbol{\beta} + [\mathbf{I}_{nT} - \rho(\mathbf{I}_T \otimes \mathbf{W})]^{-1}\mathbf{U}$$
$$\mathbf{CY} = \mathbf{C}[\mathbf{I}_{nT} - \rho(\mathbf{I}_T \otimes \mathbf{W})]^{-1}\mathbf{Z}\boldsymbol{\beta} + \mathbf{C}[\mathbf{I}_{nT} - \rho(\mathbf{I}_T \otimes \mathbf{W})]^{-1}\mathbf{U}$$
$$\mathbf{Y}_a = \mathbf{C}[\mathbf{I}_T \otimes (\mathbf{I}_n - \rho\mathbf{W})]^{-1}\mathbf{Z}\boldsymbol{\beta} + \mathbf{C}[\mathbf{I}_T \otimes (\mathbf{I}_n - \rho\mathbf{W})]^{-1}\mathbf{U}$$
$$\mathbf{Y}_a = \mathbf{CA}^{-1}\mathbf{Z}\boldsymbol{\beta} + \mathbf{CA}^{-1}\mathbf{U} \tag{7}$$

It is obtained that:

$$Cov(\mathbf{U})_{nT \times nT} = \boldsymbol{\Sigma}_U \otimes \mathbb{I}_n$$
$$Cov(\mathbf{CA}^{-1}\mathbf{U})_{T \times T} = \mathbf{CA}^{-1}(\boldsymbol{\Sigma}_U \otimes \mathbb{I}_n)(\mathbf{A}^{-1})^\top \mathbf{C}^\top = \boldsymbol{\Sigma}_Y = \mathbf{B} \tag{8}$$

therefore, the generalized least squares estimator for $\boldsymbol{\beta}$ is given by:

$$\hat{\boldsymbol{\beta}} = (\mathbf{Z}^\top (\mathbf{A}^{-1})^\top \mathbf{C}^\top \mathbf{B}^{-1}\mathbf{CA}^{-1}\mathbf{Z})^{-1}\mathbf{Z}^\top (\mathbf{A}^{-1})^\top \mathbf{C}^\top \mathbf{B}^{-1}\mathbf{Y}_a \tag{9}$$

It should be noted that the estimator $\hat{\boldsymbol{\beta}}$ depends on the values of the parameters ρ, σ^2, and ϕ_1. To estimate these parameters, the following quasi-likelihood maximization problem is solved:

$$(\hat{\rho}, \hat{\sigma}^2, \hat{\phi}_1) = \underset{(\rho, \phi_1, \sigma^2) \in \mathcal{U}}{\operatorname{argmax}} \left\{ -\frac{1}{2}\log|\boldsymbol{\Sigma}_Y| - \frac{1}{2}(\mathbf{Y}_a - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}_Y^{-1}(\mathbf{Y}_a - \boldsymbol{\mu}) \right\} \tag{10}$$

where

$$\boldsymbol{\mu} = \mathbf{CA}^{-1}\mathbf{Z}\boldsymbol{\beta} \quad \text{and} \quad \boldsymbol{\Sigma}_Y = \mathbf{CA}^{-1}(\boldsymbol{\Sigma}_U \otimes \mathbb{I}_n)(\mathbf{A}^{-1})^\top \mathbf{C}^\top.$$

In practice, this maximization in Equation (10) is carried out in R Core Team (2024) using the `optim` function with the "L-BFGS-B" algorithm, which allows the constraints on the parameter space $\mathcal{U}$ to be explicitly im-

posed. The L-BFGS-B algorithm is a limited-memory version of the Broyden–Fletcher–Goldfarb–Shanno (BFGS) quasi-Newton method, adapted for large-scale problems with simple box constraints (Byrd et al., 1995).

Based on this joint estimation approach, the following iterative procedure is proposed to obtain consistent estimators of $(\boldsymbol{\beta}, \rho, \sigma^2, \phi_1)$ and the disaggregated spatio-temporal series:

1. Estimate $(\hat{\rho}, \hat{\phi}_1, \hat{\sigma}^2)$ by maximizing the quasi-likelihood function using the `optim` function in `R` with the `"L-BFGS-B"` algorithm, which allows the box constraints $(\rho, \phi_1) \in (-1, 1)$ and $\sigma^2 > 0$ to be imposed explicitly (Byrd et al., 1995).

2. Given these estimates, obtain $\hat{\boldsymbol{\beta}}$ as:

$$\hat{\beta} = \left(\mathbf{Z}^\top \mathbf{A}^{-1} \mathbf{C}^\top (\mathbf{C}\mathbf{B}\mathbf{C}^\top)^{-1} \mathbf{C}\mathbf{A}^{-1} \mathbf{Z}\right)^{-1} \mathbf{Z}^\top \mathbf{A}^{-1} \mathbf{C}^\top (\mathbf{C}\mathbf{B}\mathbf{C}^\top)^{-1} \mathbf{Y}_a.$$

3. The disaggregated series is then estimated by:

$$\widehat{\boldsymbol{Y}} = \widehat{\mathbf{A}}^{-1} \mathbf{Z}\hat{\beta} + \widehat{\mathbf{B}}\, \mathbf{C}^\top \left(\mathbf{C}\widehat{\mathbf{B}}\, \mathbf{C}^\top\right)^{-1} \left(\mathbf{Y}_a - \mathbf{C}\widehat{\mathbf{A}}^{-1} \mathbf{Z}\hat{\beta}\right).$$

4. **Anchorage step.** When partial information is available for specific regions and time points, such information can be incorporated as additional linear constraints of the form

$$\mathbf{H}\boldsymbol{Y} = \mathbf{d},$$

where $\mathbf{H}$ is a known indicator matrix (with entries in $\{0, 1\}$) that selects the positions in $\boldsymbol{Y}$ for which values are known, and $\mathbf{d}$ is the corresponding vector of observed values. This anchorage does not require full annual coverage and may include only a few scattered observations across time and regions (e.g., benchmark values or administrative figures). The constraint is incorporated jointly with the aggregation constraint $\mathbf{C}\boldsymbol{Y} = \mathbf{Y}_a$ by stacking both into an extended constraint system:

$$\begin{bmatrix} \mathbf{C} \\ \mathbf{H} \end{bmatrix} \boldsymbol{Y} = \begin{bmatrix} \mathbf{Y}_a \\ \mathbf{d} \end{bmatrix}.$$

The same formula for $\widehat{\boldsymbol{Y}}$ applies, replacing $\mathbf{C}$ and $\mathbf{Y}_a$ with the extended

versions:

$$\widehat{\boldsymbol{Y}} = \widehat{\mathbf{A}}^{-1}\mathbf{Z}\hat{\boldsymbol{\beta}} + \widehat{\mathbf{B}}\,\widetilde{\mathbf{C}}^{\top}\left(\widetilde{\mathbf{C}}\,\widehat{\mathbf{B}}\,\widetilde{\mathbf{C}}^{\top}\right)^{-1}\left(\widetilde{\mathbf{Y}}_a - \widetilde{\mathbf{C}}\,\widehat{\mathbf{A}}^{-1}\mathbf{Z}\hat{\boldsymbol{\beta}}\right),$$

where

$$\widetilde{\mathbf{C}} = \begin{bmatrix}\mathbf{C} \\ \mathbf{H}\end{bmatrix}, \quad \widetilde{\mathbf{Y}}_a = \begin{bmatrix}\mathbf{Y}_a \\ \mathbf{d}\end{bmatrix}.$$

This augmentation reduces uncertainty in units with weak auxiliary signal and can substantially improve local accuracy.

The estimator $\widehat{\boldsymbol{Y}}$ corresponds to the best linear unbiased predictor (BLUP) under the aggregation constraint. Specifically, it arises from the conditional expectation $\mathbb{E}(\mathbf{Y}|\mathbf{Y}_a)$ when the latent disaggregated vector $\boldsymbol{Y}$ follows $\mathbb{E}(\boldsymbol{Y}) = \mathbf{A}^{-1}\mathbf{Z}\boldsymbol{\beta}$ and $\mathrm{Cov}(\boldsymbol{Y}) = \mathbf{B}$, and the observed aggregate satisfies $\mathbf{Y}_a = \mathbf{C}\boldsymbol{Y}$. Standard results on multivariate normal conditional distributions imply that (Harvey, 1990; Durbin and Koopman, 2012):

$$\mathbb{E}(\mathbf{Y}|\mathbf{Y}_a) = \tilde{\boldsymbol{Y}} + \mathrm{Cov}(\boldsymbol{Y}, \mathbf{Y}_a)\,\mathrm{Var}(\mathbf{Y}_a)^{-1}(\mathbf{Y}_a - \mathbb{E}(\mathbf{Y}_a)), \quad \text{with} \quad \tilde{\boldsymbol{Y}} = \mathbf{A}^{-1}\mathbf{Z}\boldsymbol{\beta}.$$

Substituting the model components yields the explicit form:

$$\widehat{\boldsymbol{Y}} = \widehat{\mathbf{A}}^{-1}\mathbf{Z}\hat{\boldsymbol{\beta}} + \widehat{\mathbf{B}}\,\mathbf{C}^{\top}\left(\mathbf{C}\widehat{\mathbf{B}}\,\mathbf{C}^{\top}\right)^{-1}\left(\mathbf{Y}_a - \mathbf{C}\widehat{\mathbf{A}}^{-1}\mathbf{Z}\hat{\boldsymbol{\beta}}\right).$$

By construction, this predictor ensures aggregation coherence since:

$$\mathbf{C}\widehat{\boldsymbol{Y}} = \mathbf{C}\tilde{\boldsymbol{Y}} + \mathbf{C}\mathbf{B}\mathbf{C}^{\top}(\mathbf{C}\mathbf{B}\mathbf{C}^{\top})^{-1}(\mathbf{Y}_a - \mathbf{C}\tilde{\boldsymbol{Y}}) = \mathbf{C}\tilde{\boldsymbol{Y}} + (\mathbf{Y}_a - \mathbf{C}\tilde{\boldsymbol{Y}}) = \mathbf{Y}_a.$$

Therefore, the predicted disaggregated series exactly sums to the observed total. This guarantees coherence between the disaggregated estimates and the aggregate data while minimizing the conditional prediction variance (Harvey, 1989; Durbin and Koopman, 2012).

Finally, basic regularity conditions are assumed to ensure consistency and asymptotic normality of the estimators, with additional conditions detailed in the following section.

Assumption 1. *The $\{\epsilon_{it}\}$, $i = 1, \ldots, n$ and $t = 1, \ldots, T$ are i.i.d with mean zero and variance σ^2. Its moment $\mathbb{E}(|\epsilon_{it}|^{4+\gamma})$ for some $\gamma > 0$ exits.*

Assumption 2. *The elements w_{ij} satisfies that $\sum_{j=1}^{n} w_{ij} = 1$, $i = 1, \ldots, n$. And $\exists j \neq j'$ such that:*

$$\sum_{i=1}^{n} w_{ij} \neq \sum_{i=1}^{n} w_{ij'}$$

Assumption 3. *$n > 2$ is fixed for all $t = 1, \ldots, T$, $|\rho| < 1$ and, $|\phi_1| < 1$*

Assumption 4. *The elements of $\mathbf{Z}$ are uniformly bounded constants for all n and T. The $\lim_{T \to \infty} \frac{1}{T} \mathbf{Z}^{\top} \mathbf{Z}$ exists and is nonsingular.*

Assumption 5. *The $\lim_{T \to \infty} \frac{1}{T} (\mathbf{Z}, \mathbf{GZ\beta})^{\top} (\mathbf{Z}, \mathbf{GZ\beta})$ exists and is nonsingular, where $\mathbf{G} = (\mathbb{I}_T \otimes \mathbf{W}) \mathbf{A}^{-1}$*

Theorem 1. *Under Assumptions 1-5, the parameter vector $\boldsymbol{\theta} = (\boldsymbol{\beta}^{\top}, \phi_1, \sigma^2, \rho)^{\top}$ is globally identifiable in Equation (2). Additionally, the estimator obtained in Equations (9), (10) satisfies that:*

$$\hat{\boldsymbol{\theta}} \xrightarrow[T \to \infty]{d} \mathbf{N}_{p+3}(\boldsymbol{\theta}, \boldsymbol{\Sigma}_{\theta}^{-1} + \boldsymbol{\Sigma}_{\theta}^{-1} \boldsymbol{\Delta} \boldsymbol{\Sigma}_{\theta}^{-1}) \tag{11}$$

where $\boldsymbol{\Sigma}_{\theta}$ and $\boldsymbol{\Delta}$ are defined in Equations (A.24) and (A.25), respectively. Additionaly, if $\epsilon_{it} \sim N(0, \sigma^2)$, then, $\boldsymbol{\Delta} = \mathbf{0}$

Proof. See Appendix A $\qquad \square$

3. Theoretical R^2 and RMSE

The unconditional variance of the disaggregated series is thus:

$$Var(\mathbf{Y}) = \boldsymbol{\Sigma}_Y = \mathbf{A}^{-1}(\boldsymbol{\Sigma}_U \otimes \mathbb{I}_n)(\mathbf{A}^{-1})^{\top}. \tag{12}$$

The proportion of variance explained by the covariates at the disaggregated level is given by:

$$R^2(\mathbf{Y}|\mathbf{Z}) = 1 - \frac{\text{trace}\left[Var(\mathbf{Y}|\mathbf{Z}, \mathbf{Y}_a)\right]}{\text{trace}\left[Var(\mathbf{Y})\right]}, \tag{13}$$

where the conditional variance is:

$$Var(\mathbf{Y}|\mathbf{Z}, \mathbf{Y}_a) = \boldsymbol{\Sigma}_Y - \boldsymbol{\Sigma}_Y \mathbf{C}^{\top} \mathbf{B}^{-1} \mathbf{C} \boldsymbol{\Sigma}_Y, \quad \text{with} \quad \mathbf{B} = \mathbf{C} \boldsymbol{\Sigma}_Y \mathbf{C}^{\top}. \tag{14}$$

Correspondingly, the theoretical RMSE for the disaggregated series is:

$$RMSE(\mathbf{Y}|\mathbf{Z}, \mathbf{Y}_a) = \sqrt{\frac{1}{nT} \text{ trace} \left[Var(\mathbf{Y}|\mathbf{Z}, \mathbf{Y}_a) \right]}. \qquad (15)$$

For the aggregated series $\mathbf{Y}_a = \mathbf{CY}$, the variance structure becomes:

$$Var(\mathbf{Y}_a) = \mathbf{B} = \mathbf{C}\mathbf{\Sigma}_Y\mathbf{C}^\top, \quad Var(\mathbf{Y}_a|\mathbf{Z}) = \mathbf{C}Var(\mathbf{Y}|\mathbf{Z})\mathbf{C}^\top. \qquad (16)$$

Thus, the explained proportion at the aggregate level is:

$$R^2(\mathbf{Y}_a|\mathbf{Z}) = 1 - \frac{\text{trace}\left[Var(\mathbf{Y}_a|\mathbf{Z}) \right]}{\text{trace}\left[Var(\mathbf{Y}_a) \right]}, \qquad (17)$$

and the corresponding RMSE is:

$$RMSE(\mathbf{Y}_a|\mathbf{Z}) = \sqrt{\frac{1}{T} \text{ trace}\left[Var(\mathbf{Y}_a|\mathbf{Z}) \right]}. \qquad (18)$$

The innovation variance σ^2 defined in (4) acts as a scalar for the uncertainty in both the disaggregated and aggregated series, proportionally increasing RMSE as it grows, while the R^2 remains stable if the relative signal-to-noise ratio is unchanged. The autoregressive parameter ϕ_1 affects the serial correlation: higher values of ϕ_1 amplify the temporal dependence within each region, enlarging off-diagonal covariances in (12). If the covariates do not capture this temporal persistence, the conditional variance (14) stays high, resulting in larger RMSE and lower R^2.

The spatial dependence parameter ρ controls the propagation of both signal and noise through the spatial structure defined by $\mathbf{W}$. A higher ρ increases the spatial feedback captured by $\mathbf{A}^{-1}$, which can enhance the portion of variance explained if the covariates align with the spatial pattern. However, this same feedback effect spreads noise spatially, potentially raising RMSE if the covariates are weak.

Aggregation generally averages out idiosyncratic variation across subregions. When positive spatial correlation is present ($\rho > 0$), the aggregate series may exhibit a larger share of its variance explained by covariates that primarily capture the common trend, leading to $R^2(\mathbf{Y}_a)$ exceeding $R^2(\mathbf{Y})$ if local heterogeneity remains unmodelled.

These implications can be qualitatively summarized as shown in Table 1.

Table 1: Expected patterns for R^2 and RMSE under different covariate strength and spatial dependence scenarios.

Z Strength	Spatial	$R^2(\mathbf{Y})$	$R^2(\mathbf{Y}_a)$	RMSE($\mathbf{Y}$)	RMSE($\mathbf{Y}_a$)
Strong	Low ($\rho \approx 0$)	High	High	Low	Low
Strong	High ($|\rho| \to 1$)	High	High	Low	Low
Weak	Low ($\rho \approx 0$)	Low	Low	High	Moderate
Weak	High ($|\rho| \to 1$)	Low	Moderate to High	High	Moderate

In summary, the combined interpretation of (13)–(17) and (15)–(18) clarifies how the covariate signal, temporal dependence, and spatial feedback interact to determine both the proportion of explained variation and the level of residual uncertainty, at both disaggregated and aggregated levels.

4. Simulation

To empirically validate the quality of the model we perform a simulation with synthetic data in R.

4.1. Synthetic data generation

We generate synthetic data with parameters $n, t, \rho, \phi, \theta, \beta_0, \beta_1, \sigma$ in the following way:

To start, we fix the number of municipalities $n = k^2$. We position spatially the municipalities in a grid as presented in Figure 1. We construct the $n \times n$ adjacency matrix $\mathbf{W}$ as $w_{ij} = 1$ when municipalities i, and j are adjacent, and 0 otherwise.

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

Figure 1: Example of spatial positioning of synthetic data with $4^2 = 16$ municipalities. Municipality 6 is adjacent to $1, 2, 3, 5, 7, 9, 10, 11$.

We continue with the $nt \times 2$ matrix $\mathbf{Z}$, where first column entries are one and each entry z_{i2} is the realization of a random variable $X \sim \mathcal{U}[0, 1]$.

To compute the $nt \times 1$ matrix $\mathbf{U}$ as the t values of n distinct $AR(1)$ models with autoregressive parameter ϕ, and innovations $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$ for all values of i. The generation of $AR(1)$ data is performed with the `arima.sim` instruction, part of the `stats` package, version 4.4.3.

We obtain the synthetic data $\mathbf{Y}$ as the result of

$$\mathbf{Y} = \mathbf{A}^{-1}\mathbf{Z}\beta + \mathbf{A}^{-1}\mathbf{U}$$

where $\beta = [\beta_0, \beta_1]^T$ and

$$\mathbf{A}^{-1} = \mathbb{I}_t \otimes (\mathbb{I}_n - \rho\mathbf{W})^{-1}.$$

The usage of $\mathbf{Z}, \mathbf{Y}, \mathbf{W}, \mathbf{U}$ corresponds with the notation of Equation 3 (?). To fasten the computation the package `furrr` was used to parallelize the process in a local Linux Server.

4.2. Simulation parameters and metrics

For all combinations of values:

- $n \in \{9, 16, 25, 36, 49, 64\}$

- $t \in \{12, 24, 36, \dots, 144\}$

- $\rho \in \{0, \pm 0.25, \pm 0.5, \pm 0.75, \}$

- $\phi \in \{0, \pm 0.25, \pm 0.5, \pm 0.75, \}$

- $\beta_0 = 1$

- $\beta_1 = \{0, 0.5, 1, 5, 10, 50, 100\}$

- $\sigma = \{0.1, \sqrt{0.1}, 1\}$

we have generated synthetic data and executed the proposed model to

see the quality of our results. We have considered the following metrics

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \sum_{t=1}^{T} (Y_{it} - \hat{Y}_{it})^2}{\sum_{i=1}^{n} \sum_{t=1}^{T} (Y_{it} - \bar{Y})^2}$$

$$MAPE = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \left| \frac{Y_{it} - \hat{Y}_{it}}{Y_{it}} \right| \times 100$$

$$RRMSE = \frac{RMSE}{\bar{Y}} = \frac{1}{\bar{Y}} \sqrt{\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} (Y_{it} - \hat{Y}_{it})^2}$$

$$RMSE = \sqrt{\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} (Y_{it} - \hat{Y}_{it})^2}$$

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{\sum_{t=1}^{T} (Y_{it} - \hat{Y}_{it})}{\sum_{t=1}^{T} Y_{it}} \right) \tag{19}$$

Due to the high number of parameter combinations ($6 \cdot 12 \cdot 7 \cdot 7 \cdot 7 \cdot 3 = 74088$) we grouped the results in 3 categories depending on the value of $\frac{\beta_1}{\sigma^2}$ as by the signal-to-noise ratio. Lower values indicate weaker covariate effects relative to the residual variance. The specific choice of values was made to balance the number of results per category (see Table 2).

$$\text{ratio} = \frac{\beta}{\sigma^2} \quad \text{with} \quad \begin{cases} \text{Low} & \text{if ratio} < 5 \\ \text{Medium} & \text{if } 5 \leq \text{ratio} < 50 \\ \text{High} & \text{if } 50 \leq \text{ratio} \leq 500 \\ \text{Very High} & \text{if ratio} > 500 \end{cases}$$

Table 2: Number of results per category.

Category	Count
Low	21168
Medium	17640
High	21168
Very High	14112

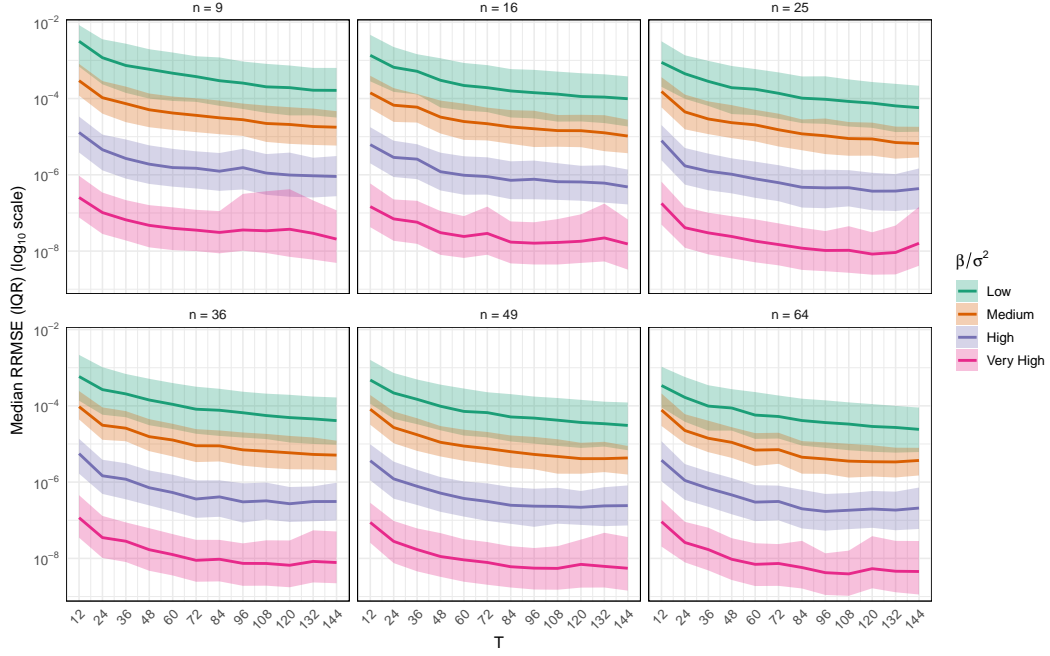Figure 2 shows that the behavior of the mediated RMSE and the quartiles

Figure 2: Median relative root mean squared error (RRMSE) with interquartile ranges (shaded ribbons) plotted on a log scale, across different time lengths T and sample sizes n. Colors represent different signal-to-noise ratio levels measured by β/σ^2.

is similar regardless of the size of the spatial grid according to the quotient between $\frac{\beta}{\sigma^2}$. It is clear that the higher this value, the better the reconstruction of the time series. It is also shown that as T grows, there is a slight decrease in the RMSE of the disaggregation, although it is a smooth logarithmic trend.

Figure 2 shows that the behavior of the median RMSE and the quartiles is similar regardless of the spatial grid size based on the ratio between $\frac{\beta}{\sigma^2}$. It is placed on a logarithmic scale to avoid very small numbers. It is evident that the higher this value, the better the reconstruction of the time series. It also shows that, as T increases, a slight decrease in the RMSE of the disaggregation is observed, although it is a smooth logarithmic trend.

The small values show that the disaggregation yields very low metric results. Table 4.2 shows the summary metrics for each of the ratio levels between $\frac{\beta}{\sigma^2}$. Since the RMSE is a non-relative measure, it increases as the ratio increases. This is explained because as the ratio increases, the variable Y_{it} in the simulation takes on larger values. Therefore, by relativizing this

12

Table 3: Average performance metrics by covariate strength level and number of spatial units (n).

β/σ^2	n	RMSE	log $RRMSE$	MAPE	R^2	χ^2
Low	9	0.856	-8.030	5.124	-0.036	1.515
	16	0.834	-8.591	3.377	-0.002	1.851
	25	0.845	-9.011	4.590	-0.031	1.560
	36	0.843	-9.384	6.488	-0.047	1.553
	49	0.843	-9.693	4.384	-0.048	1.539
	64	0.859	-9.941	6.599	-0.110	1.629
Medium	9	0.716	-10.162	1.143	0.704	0.371
	16	0.754	-10.626	0.894	0.658	0.446
	25	0.786	-11.018	0.897	0.612	0.415
	36	0.819	-11.344	0.950	0.573	0.429
	49	0.826	-11.638	1.076	0.567	0.428
	64	0.895	-11.833	1.382	0.450	0.474
High	9	1.177	-13.214	0.285	0.956	0.320
	16	1.080	-13.739	0.230	0.953	0.163
	25	1.118	-14.096	0.230	0.950	0.256
	36	1.077	-14.420	0.360	0.942	0.142
	49	1.172	-14.704	0.191	0.938	0.129
	64	1.502	-14.864	0.207	0.906	0.277
Very High	9	1.513	-16.493	0.120	0.969	0.627
	16	1.365	-17.143	0.069	0.961	0.311
	25	1.013	-17.590	0.076	0.968	0.323
	36	0.671	-17.925	0.085	0.983	0.092
	49	0.866	-18.197	0.139	0.974	0.100
	64	1.495	-18.310	0.071	0.951	0.344

RMSE, the RRMSE and the log $RRMSE$ are shown to illustrate the difference between the models. This is much smaller as the ratio increases, as seen in Figure 2. The MAPE is also shown, where the reduction is even more significant, with values less than 1% for the High and Very High values. The R^2 is also shown, which obtains very good values for High and Very High. It presents medium-high R^2 values when the ratio is Medium, and negative values close to zero when the ratio is Low. This is explained because, when the variables $\boldsymbol{Z}$ are not informative in the proposed model, the disaggregation is similar to performing a disaggregation with the mean, that is, a proxy for naivety.

The value of χ^2 proposed in the equation (19) is also presented, which weights the goodness of fit for each spatial unit. This is of particular interest because it allows us to assess whether the spatial units are well disaggregated. It is shown that the model disaggregates all variables very well for each spatial unit, even with large n sizes.
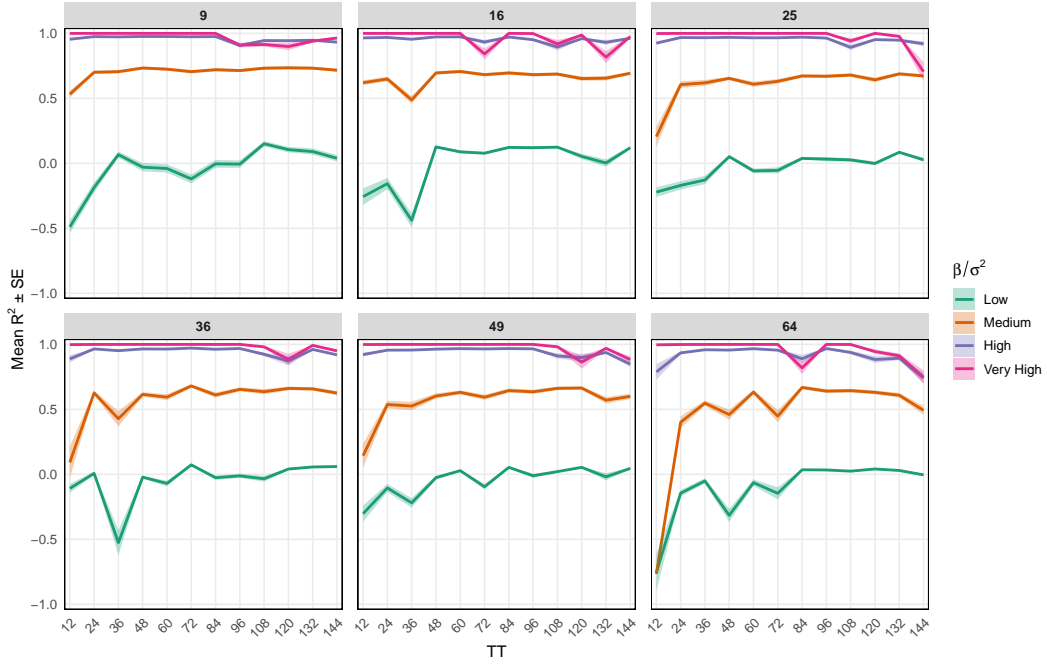

Figure 3: Mean observed R^2 with standard error bands across different time lengths (T), grouped by the number of spatial units (n) shown in facets. Colours represent levels of the ratio β/σ^2.

What was stated above regarding R^2 is reaffirmed in Figure 4. It shows that the behavior of this metric is very similar regardless of the number of spatial units. Furthermore, only in the case of a low coefficient does the model yield mediocre results, because the information provided by the covariates is too limited to reconstruct the series.

It should be noted that all the metrics shown above were calculated using the average of all ρ values. This demonstrates the consistency of the method regardless of the value of the spatial correlation. However, we will now discuss the behavior of the RRMSE with respect to the spatial component of the model.
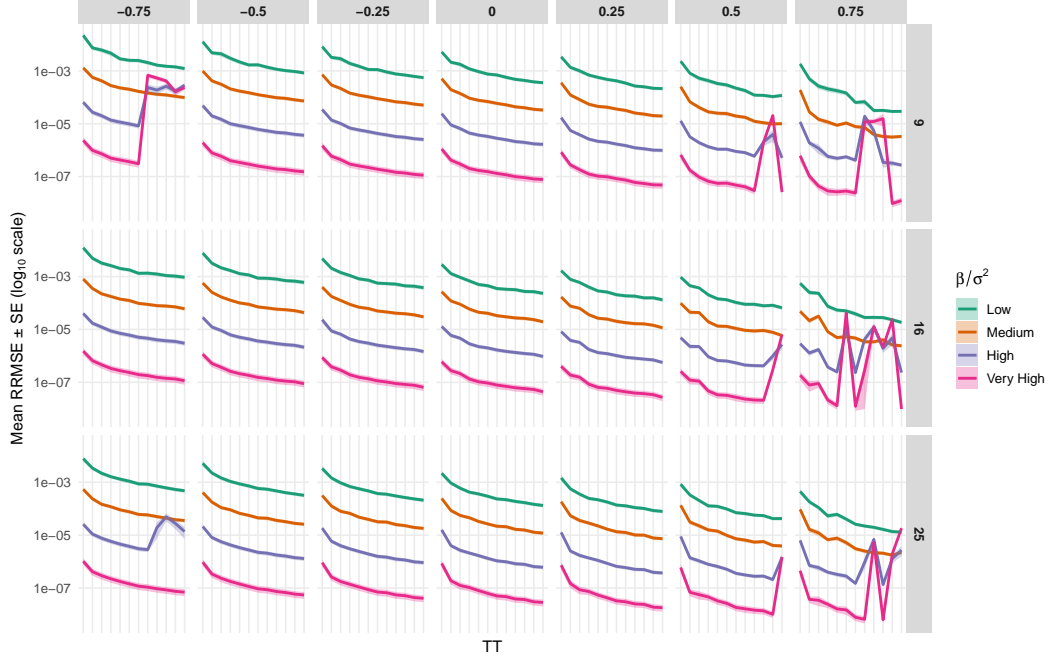
Figure 4: Mean relative root mean squared error (RRMSE) with standard error bands on a log scale across different time lengths (T), grouped by the number of samples ($n = \{9, 16, 25\}$) and spatial correlation (ρ) shown in facets. Colors represent levels of the ratio β/σ^2.

Figure 4 shows the RRMSE values on a logarithmic scale. They are subdivided by the number of spatial units. Here it is shown for 9, 16, and 25. It is observed that the behavior according to the quotient β/σ^2 is similar to what we had already observed in the previous figures; that is, the higher the quotient, the better the breakdown of the time series. But this behavior only occurs for ρ values of -0.5, 0.25, 0, 0.25. When ρ is -0.75, very anomalous behavior occurs when n is 9. And when ρ is 0.5, anomalous behavior occurs with very large T values. As for the value of $\rho = 0.75$, when the T is greater than 36, the model is completely damaged at high ratios.

In Figure 5 it is shown analogously to Figure 4 for 36, 49 and 64. It is observed that the behavior is good for all values of rho except when this value is 0.75 and T is greater than 48. With these last two graphs it is observed that the spatial part alters the model when it is very strong, because the matrix $\mathbf{A}$ is close to being singular, and therefore, presents estimability
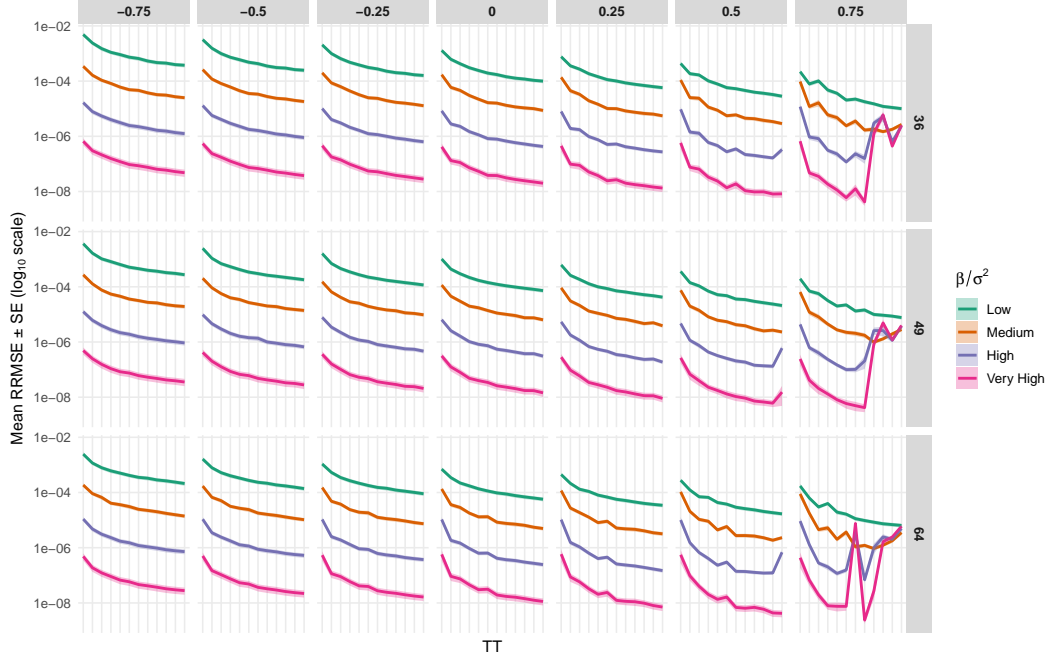
Figure 5: Mean relative root mean squared error (RRMSE) with standard error bands on a log scale across different time lengths (T), grouped by the number of samples ($n = \{36, 49, 64\}$) and spatial correlation (ρ) shown in facets. Colors represent levels of the ratio β/σ^2.

problems. Furthermore, in a practical sense, if the variables $\boldsymbol{Z}$ do not capture the spatiality, they will make the model unstable. Therefore, this model should be applied with caution if the estimator of ρ is close to 1 and the sample size in the time series is medium or large. In all other scenarios, the disaggregation is as good as the covariates that allow the disaggregation of the series of interest.

5. Application- GDP in Spain

This document describes the auxiliary variables employed to disaggregate the Gross Domestic Product (GDP) in Spain at the level of Autonomous Communities (NUTS-2) from 2002 to 2023. These covariates are selected for their economic relevance, availability at the regional level, and statistical relationship with the GDP components. All data sources are official and publicly available.

16

- Industrial Production Index (IPI): The IPI is a monthly short-term indicator that measures the evolution of industrial activity in Spain, excluding construction. It covers sections B (Mining and quarrying), C (Manufacturing), D (Electricity and gas supply), and division 36 of section E (Water supply; sewerage, waste management and remediation activities) of the CNAE-2009 classification. The IPI reflects changes in the physical volume of production, eliminating the effects of prices. Over 11500 industrial establishments are surveyed monthly to construct this index.[1]

- Population: The annual population register provides the number of residents in each Autonomous Community. Population size affects the scale of production and consumption, thus impacting regional GDP.[2][*ATN: Las fuentes mejor como citas?*]

- Industrial Energy Consumption: This survey provides data on electricity, natural gas, fuel oil, renewable sources, petroleum products, and other fuels used by industrial establishments. It covers companies with at least 20 employees in sections B and C of CNAE-2009. Energy consumption is a proxy for industrial activity and productive intensity. From 2022, the survey has annual periodicity.[3]

- Business Demography: The Statistics on Mercantile Companies provide monthly data on the number of companies created, dissolved, and capital modifications. From this, we extracted the number of new companies, the amount of capital subscribed, and the capital disbursed. Business dynamism is closely related to regional investment and economic activity.[4]

---

[1]Source: Instituto Nacional de Estadística (INE). https://www.ine.es/dyngs/INEbase/operacion.htm?c=Estadistica_C&cid=1254736145519&idp=1254735576715

[2]Source: INE Annual Municipal Register. https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177095&menu=ultiDatos&idp=1254735572981

[3]Source: INE Energy Consumption Survey. https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736146240&menu=ultiDatos&idp=1254735576715

[4]Source: INE Statistics on Mercantile Companies. https://ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177026&idp=1254735576550

- Household Food Expenditure: This dataset provides annual series of food consumption in monetary terms by households, disaggregated by Autonomous Community. It captures regional differences in household demand and consumption patterns, which are relevant to the GDP via the expenditure approach.[5]

- Labour Force and Activity Rates: The Labour Force Survey collects quarterly information on the working-age population, including the number of economically active individuals by sex, and gender-specific activity rates. These indicators reflect the availability and participation of human capital in the regional economy.[6]

These variables serve as proxies of sectoral, demographic, and labour market conditions across regions and over time, enabling a robust estimation of the regional GDP trajectories within the disaggregation framework.

5.1. Spatial Weight Matrices

Instead of relying on conventional definitions of spatial proximity—such as contiguity or fixed-distance thresholds—we developed a custom spatial weight matrix designed to reflect spatial relationships based on the socio-economic and demographic structure of each region. To construct this matrix, we employed Gower distance, which is particularly well-suited for spatial analysis involving variables of mixed types. Unlike Euclidean distance, Gower distance accommodates heterogeneous data structures, making it especially appropriate for comparing regions with complex socio-economic profiles. For example, Comber et al. (2011) applied Gower distance to analyze spatial disparities in access to health services by incorporating multiple socio-economic dimensions, while Dai and Wang (2011) used it to assess inequalities in food accessibility based on a combination of demographic and geographic indicators. Originally introduced by Gower (1971), this distance metric offers a

---

[5]Source: Ministry of Agriculture, Fisheries and Food (MAPA). `https://www.mapa.gob.es/es/alimentacion/temas/consumo-tendencias/panel-de-consumo-alimentario/series-anuales`

[6]Source: INE Labour Force Survey (EPA). `https://www.ine.es/dyngs/INEbase/operacion.htm?c=Estadistica_C&cid=1254736176918&menu=ultiDatos&idp=1254735976595`

flexible and interpretable framework for capturing spatial similarity in contexts where conventional proximity measures may be inadequate.

To compute the Gower distance between each pair of autonomous communities, we used a set of eleven indicators corresponding to the year 2021: population density, birth rate per 1,000 inhabitants, mortality rate per 1,000 inhabitants, life expectancy, risk of poverty, Human Development Index (HDI), Gini index, total public debt (in millions of euros), public deficit (in millions of euros), unemployment rate according to the Labour Force Survey (LFS), and the Consumer Price Index (CPI). This multidimensional set captures structural differences across regions and supports a more substantive definition of spatial similarity. The distance matrix was calculated using the `daisy()` function from the `cluster` package in R Maechler et al. (2024), which offers a robust implementation for mixed data types. All variables were assigned equal weights, ensuring that each contributed equally to the overall dissimilarity measure. In Figure 6, we present the resulting Gower distance matrix between the autonomous communities using a heatmap, where each cell represents the pairwise dissimilarity between two regions. Darker blue tones indicate higher similarity (lower distance), while lighter red tones denote greater dissimilarity (higher distance). A hierarchical clustering dendrogram accompanies the heatmap, grouping regions with similar socio-economic and demographic profiles. This visualization facilitates the identification of clusters of autonomous communities that share comparable structural characteristics.

On the rightmost side of the heatmap, Castilla-La Mancha, Islas Baleares, La Rioja, Aragón, and Cantabria appear closely clustered and share predominantly blue-toned cells, indicating a high degree of similarity in their regional profiles. These regions display comparable values in population density, birth and mortality rates, life expectancy, and unemployment. Similarly, Navarra, the Basque Country, Asturias, Galicia, and Castilla y León are grouped together in the dendrogram, suggesting structural proximity with respect to low poverty risk, high HDI scores, and similar CPI levels.

In contrast, Ceuta and Melilla consistently exhibit reddish shades in their pairwise comparisons with all other autonomous communities, highlighting their distinctiveness. These differences are likely due to their high population

---

[6]Note: Autonomous community names are presented in Spanish as per their official designation.
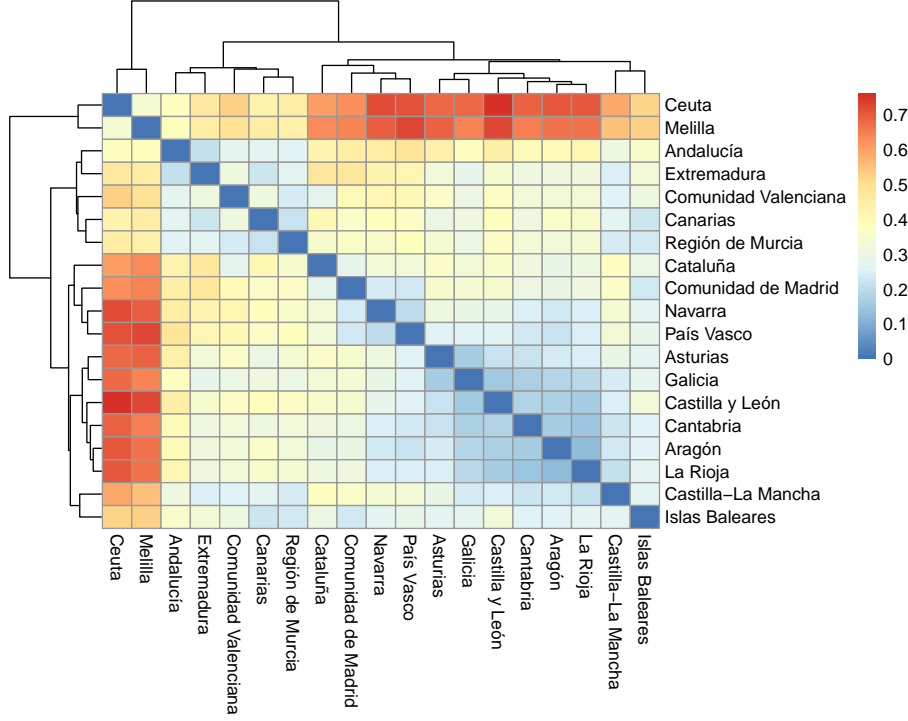
Figure 6: Heatmap of Gower distances between Spanish autonomous communities. Each cell represents the dissimilarity between a pair of regions based on eleven socio-economic and demographic variables for the year 2021. Darker blue shades indicate higher similarity (lower distance), while lighter red shades reflect greater dissimilarity (higher distance). The dendrogram shows the hierarchical clustering of regions with similar profiles.

densities, elevated birth rates, and markedly different labor market indicators, such as unemployment. Andalusia also shows substantial distances from many other regions, suggesting a differentiated socio-economic profile characterized by higher unemployment, lower HDI, and above-average poverty risk. These patterns underscore the value of incorporating a wide range of indicators when defining spatial similarity, as doing so reveals not only geographic groupings but also deeper structural disparities across regions.

5.2. Selection of Covariates

Although some energy-related variables exhibited missing values in certain years and regions, these gaps were relatively limited in both frequency and magnitude. To preserve the internal structure and temporal consistency

of each series, imputation was conducted independently for each region and variable using linear interpolation over time. This method, while simple, is widely used when the primary concern is maintaining continuity without introducing artificial trends or external information (Little and Rubin, 2002; Junger and De Leon, 2015). More sophisticated imputation strategies could potentially exploit cross-variable or cross-region correlations, but these were deliberately avoided to prevent leakage of information that could bias the temporal disaggregation model.

Ceuta and Melilla were excluded from the analysis due to their unique administrative and economic profiles, which differ substantially from those of the 17 main Autonomous Communities and may lead to unstable or unrepresentative estimates.

Because many covariates exhibited strong multicollinearity—especially due to the use of multiple correlated economic indicators and energy series—a Principal Component Analysis (PCA) was applied to the standardized covariates. PCA is a classical technique for dimension reduction that transforms a correlated set of variables into a smaller number of uncorrelated components that retain most of the original variance. This transformation improves numerical stability and interpretability in regression-type models, especially when the number of predictors is large relative to the sample size or when predictors are highly collinear (Jolliffe, 2011). In the context of temporal disaggregation, the use of PCA has been shown to improve estimation by regularizing the predictor space and avoiding overfitting (Durbin, 1960; Proietti, 2006).

Instead of selecting covariates based on out-of-sample predictive error—which is not directly available in this context—we relied on structural diagnostics of the fitted models. In particular, we focused on the estimated innovation variance ($\hat{\sigma}^2$) and the autocorrelation parameter ($\hat{\rho}$) from the temporal model. Simulation studies performed in parallel to this application indicated that high values of $\hat{\rho}$ (close to 1) in conjunction with large innovation variances were associated with unstable or biased disaggregation, especially when the series included regions with sparse or weak signal. Consequently, among the different combinations of covariates explored, we prioritized those yielding lower values of $\hat{\sigma}^2$ and moderate values of $\hat{\rho}$, avoiding those approaching the unit root boundary.

This selection criterion ensures that the model retains a reasonable degree of temporal smoothness without being overly driven by persistence, while also mitigating the propagation of uncertainty during the disaggregation process.

The final model included a reduced set of principal components derived from the full pool of covariates, ensuring interpretability and stability without relying on subjective or ad hoc selection mechanisms.

5.3. Results of disaggregation

From the full set of available covariates, a subset of six variables was selected for inclusion in the final disaggregation model based on their interpretability and contribution to the explained variance. The selected variables were: Household Food Expenditure, Total Energy Consumption, Number of Active Men, Electricity Consumption, Coal and Derivatives Consumption, and Gas Consumption. This selection captures key aspects of household demand, labor force participation, and energy use, which are all closely related to regional GDP dynamics.

Table 4: Principal Component Analysis results: standard deviations, explained variance, and loadings

Variable / Statistic	PC1	PC2	PC3	PC4	PC5	PC6
Standard Deviation	2.196	0.842	0.563	0.320	0.196	0.103
Explained Variance (%)	80.4	11.7	5.7	1.7	0.6	0.2
Consumo Euros	-0.409	0.468	-0.147	0.369	-0.672	-0.065
Total Energy Consumption	-0.436	-0.319	-0.087	0.160	0.071	0.819
Number of Active Males	-0.373	0.651	-0.185	-0.231	0.591	0.030
Electricity Consumption	-0.426	-0.344	0.041	0.562	0.368	-0.498
Coal and Derivatives Consumption	-0.397	0.004	0.849	-0.326	-0.115	-0.046
Gas Consumption	-0.406	-0.370	-0.463	-0.602	-0.215	-0.273

The Principal Component Analysis (PCA) results (see Table 4 reveal that the first three components explain approximately 97.8% of the total variance in the selected covariates, with PC1 accounting for the majority (80.4%), PC2 contributing 11.7%, and PC3 adding 5.7%. The remaining components collectively explain less than 2.5% of the variance and were therefore excluded from the analysis.

PC1 appears to capture the overall economic activity and energy consumption patterns, as evidenced by strong negative loadings for total energy consumption, electricity consumption, and gas consumption, as well as the number of active males. This component likely reflects general production and consumption trends closely linked to GDP.

PC2 shows notable positive loadings on the number of active males and electricity consumption, and a negative loading on total energy consumption. This suggests it may represent structural changes in labor market participation and energy usage efficiency that can affect regional GDP dynamics differently than overall consumption.

PC3 is characterized by a strong positive loading on coal and derivatives consumption and a moderate negative loading on gas consumption. This component could be related to the energy mix and its transition effects, potentially capturing regional heterogeneities in industrial activity and energy sources relevant to GDP fluctuations.

Including the first three principal components balances the need to capture the dominant variation in the data while avoiding overfitting noise. Despite the smaller explained variance of PC3 compared to PC1 and PC2, it contains meaningful information about energy source composition and labor structure that is economically relevant for modeling regional GDP. The last three components, with negligible explained variance (less than 2.5% combined), likely represent noise or idiosyncratic fluctuations and were excluded to maintain model parsimony and interpretability.

Thus, the selection of these three components ensures that the disaggregation model incorporates the main systematic factors affecting GDP across regions and time while filtering out less informative variation, which could otherwise deteriorate model stability and predictive performance.

The disaggregation methodology was then implemented using the first three principal components as predictors. The estimates of the model parameters are shown in Table 5.

Table 5: Estimated parameters from the disaggregation model

Parameter	Estimate
ρ	-0.099
PC1 coefficient	3.6125
PC2 coefficient	3.9910
PC3 coefficient	0.8781
σ^2	0.9155
ϕ_1	0.9983

To assess model performance, we compared the observed and disaggregated GDP series using two metrics: Mean Absolute Percentage Error (MAPE)

and Root Mean Squared Error (RMSE) relative to the mean GDP. These results are reported in Table 6. In addition to the Gower-based dissimilarity matrix, we also considered a conventional spatial contiguity matrix based on shared borders between regions. However, this alternative yielded consistently worse results in terms of both error metrics, reinforcing the advantage of the Gower approach for capturing interregional similarities in this application.

Table 6: Disaggregation error metrics with and without anchoring

W	Scenario	MAPE	Relative RMSE
Spatial neighboord	Without anchoring	0.8426	0.6513
Spatial neighboord	With anchoring in 2002	0.2301	0.2930
Gower	Without anchoring	0.1548	0.1824
Gower	With anchoring in 2002	0.1068	0.1459

In the second scenario, anchoring was performed using the 17 observed GDP values for the year 2002. This information was used to improve identification and evaluate the effects of anchoring on the disaggregation estimates. The results show a clear improvement in both error metrics, reflecting the benefit of incorporating partial observed data at specific time points.

The results of the disaggregation procedure applied to Spanish regional GDP reveal several important insights. First, Table 6 clearly shows that incorporating anchoring—by including the observed GDP values for the year 2002—significantly improves the accuracy of the estimates. Both MAPE and RRMSE are considerably reduced, demonstrating the stabilizing effect of anchoring in the reconstruction of regional series.

Figures 7 and 8 illustrate the evolution of the disaggregated GDP for the 17 autonomous communities, grouped by their average GDP level. It is apparent that regions with larger economies, such as *Madrid*, *Cataluña*, and *Andalucía*, exhibit smoother and more stable trajectories. In contrast, regions with lower GDP, such as *La Rioja*, *Cantabria*, and *Extremadura*, show greater variability and occasional deviations, even when anchoring is used. This pattern is further supported by Table 7, where we observe that the regions with higher average GDP tend to have lower MAPE and RRMSE, and higher R^2 values, indicating better predictive performance.

A particularly noteworthy finding is that the disaggregation approach based on only three principal components extracted from auxiliary variables
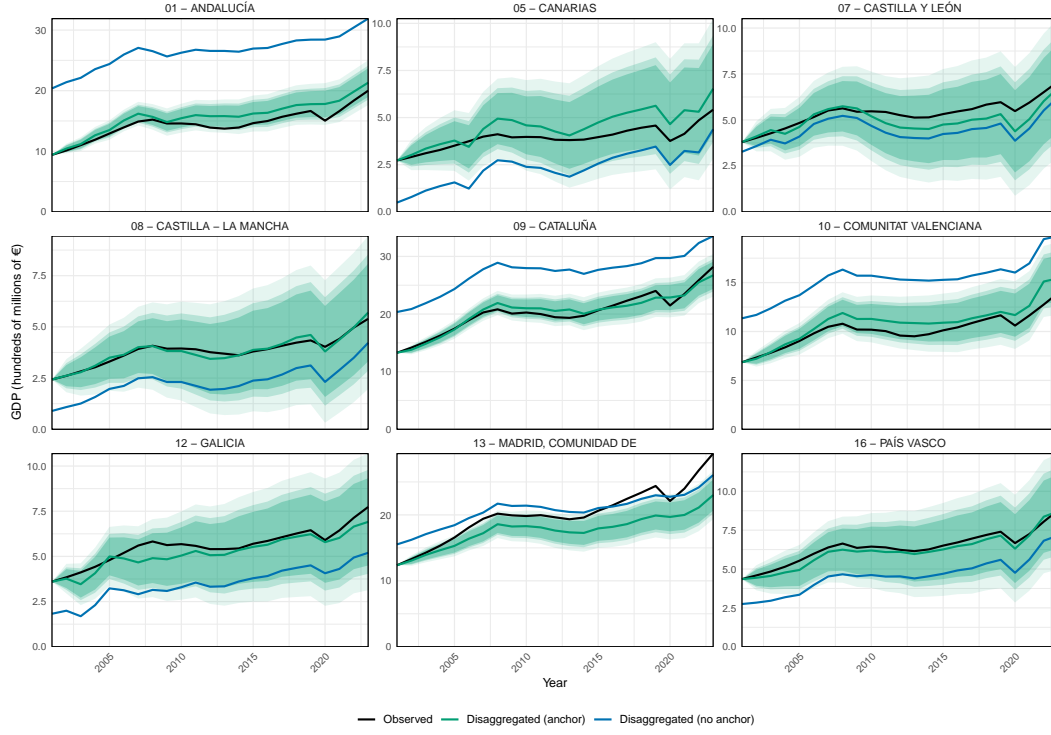
Figure 7: Evolution of GDP (in hundreds of millions of euros) over the years 2002–2023 for the 9 Spanish autonomous communities with the highest average GDP, estimated using the proposed disaggregation methodology.

was sufficient to explain the majority of the temporal and spatial variability in the regional GDP series. This dimensionality reduction preserved the key economic signals while discarding noise, making the model both efficient and interpretable. The effectiveness of this low-dimensional representation, combined with the spatial structure and the anchoring mechanism, is a central contribution of this work.

Overall, the proposed methodology offers a robust and flexible framework for temporal disaggregation in settings where only partial information is available. It succeeds in capturing both long-term economic trends and regional differences, while leveraging auxiliary information and limited direct observations to improve accuracy and interoperability.
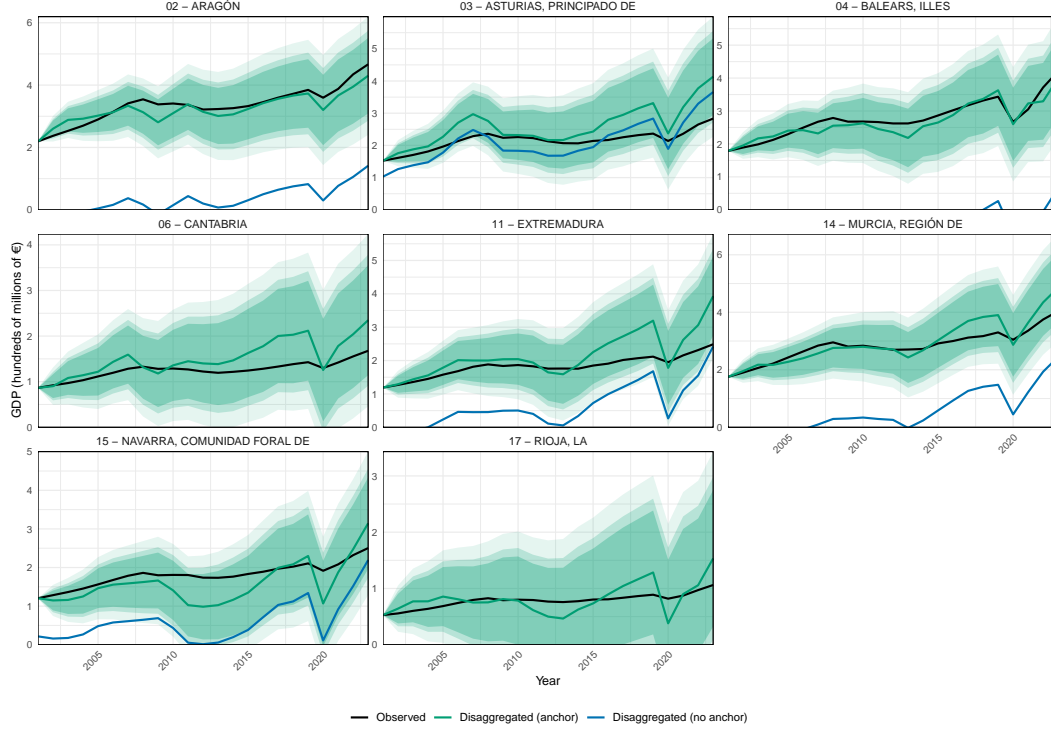
Figure 8: Evolution of GDP (in hundreds of millions of euros) over the years 2002–2023 for the 8 Spanish autonomous communities with the lowest average GDP, estimated using the proposed disaggregation methodology.

Limitations

Despite the promising results, the proposed disaggregation framework has some limitations. First, while anchoring improves estimation accuracy, its performance is highly dependent on the quality and representativeness of the observed reference year(s). If the anchored year corresponds to an atypical economic condition (e.g., recession or shock), this may introduce bias in the recovered trajectories.

Second, the current methodology assumes that the spatial weight matrix W is fixed and known. Although using geographic contiguity or Gower distances provides reasonable structure, more sophisticated approaches could involve estimating or dynamically adjusting W based on economic similarity or temporal changes, which was not explored in this work.

Furthermore, although the principal components capture most of the vari-

CCAA	MAPE (%)	RRMSE (%)	R^2 (%)	Average GDP
09	3.26	4.09	94.34	202.48
13	11.03	14.55	96.81	200.97
01	7.90	9.36	95.59	144.68
10	7.54	9.59	95.61	101.33
16	4.15	4.52	97.68	63.85
12	6.65	8.37	90.09	55.69
07	7.99	9.84	73.86	53.40
05	15.02	18.04	89.58	39.27
08	3.06	4.24	95.17	38.18
02	5.87	7.51	85.92	33.52
14	7.89	12.14	89.15	28.28
04	6.08	7.43	89.09	27.60
03	18.66	26.40	84.22	21.57
11	18.06	28.50	81.85	18.22
15	17.63	22.99	61.91	18.04
06	20.52	28.45	77.24	12.52
17	20.23	27.06	40.35	7.82

Table 7: Accuracy metrics of the disaggregated GDP estimates (with anchor) for each Spanish autonomous community, identified by CCAA code. MAPE = Mean Absolute Percentage Error; RRMSE = Relative Root Mean Squared Error; R^2 = Coefficient of Determination. GDP is expressed in millions of euros.

ation in auxiliary covariates, there is an implicit assumption that the relation between these covariates and the GDP is linear and stable over time. This may not hold in periods of structural economic change. Despite this limitation Figure 7 shows how the model correctly identifies the trends caused by the 2008 financial crisis and the 2020 COVID pandemic. Lastly, while the method offers strong empirical performance, further research could explore formal model selection strategies for determining the optimal number of components or penalization strength.

6. Conclusions

Beyond its empirical efficacy, the proposed disaggregation methodology is based on a rigorous theoretical framework. The estimation procedure is based on a linear model structure with a SAR model and an AR model in

the error. This hybrid formulation allows for a fundamental balance between flexibility and identifiability.

Importantly, under standard regularity conditions and with an increasing time dimension T, the estimators of the model parameters are consistent and asymptotically normal, even in the absence of Gaussianity in the disturbances. This robustness to distributional assumptions constitutes a key theoretical advantage, especially relevant in economic applications where outliers or biased innovations are common. The anchoring mechanism, modelled as a constrained estimation problem, guarantees identifiability and reduces variance while preserving consistency. Together, these features make the method not only practically useful but also statistically reliable, combining interpretability, computational feasibility, and theoretical guarantees.

To empirically validate our proposal, a practical evaluation of the model has been performed. With the the usage of synthetic data, results have shown how a controlled level of spatial correlation (ρ) permits efficient series data disaggregation.

Furthermore, the problem of disaggregating the Spanish GDP by Autonomous Communities has been considered to further evaluate the validity our model with real data. In this scenario the usage of PCA and anchoring show how the usage of our model in combination with common techniques in econometrics permit a good disaggregation.

As future work, we are interested in how to strengthen the model for values of ρ closer to 1 and to further refine the capacity of the proposal under dynamic scenarios, that is: where the weight matrix W is not fixed and when the relationship between covariates and the estimated value fluctuates over time.

Software implementation

Supplementary files

1. Supplementary file 1: R files with the three simulations of this paper,
2. Supplementary file 2: R files with application in GDP,

References

J. P. LeSage, R. K. Pace, Introduction to Spatial Econometrics, CRC press, 2015.

N. Cressie, C. K. Wikle, Statistics for Spatio-Temporal Data, John Wiley & Sons, 2011.

A. E. Gelfand, P. J. Diggle, M. Fuentes, P. Guttorp, Handbook of environmental and ecological statistics, CRC Press, 2019.

N.-J. Hsu, H.-C. Huang, R. S. Tsay, Matrix autoregressive spatio-temporal models, Journal of Computational and Graphical Statistics 30 (2021) 1143–1155.

G. C. Chow, A.-l. Lin, Best linear unbiased distribution and extrapolation of economic time series by related series, Review of Economics and Statistics (1971) 372–375.

Eurostat, Manual on Regional Accounts Methods, Luxembourg, 2013. URL: https://ec.europa.eu/eurostat/documents/3859598/5937641/KS-GQ-13-001-EN.PDF, corrected edition 2019. Eurostat Manuals and Guidelines, KS-GQ-13-001-EN.

A. Cuevas, E. M. Quilis, A. Espasa, Quarterly regional gdp flash estimates by means of benchmarking and chain linking, Journal of Official Statistics 31 (2015) 627–647. doi:10.1515/JOS-2015-0038.

B. Han, B. Howe, Sarn: Structurally-aware recurrent network for spatio-temporal disaggregation, in: Proceedings of the 32nd ACM International Conference on Advances in Geographic Information Systems, 2024, pp. 338–349.

E. M. Quilis, Temporal disaggregation of economic time series: The view from the trenches, Statistica Neerlandica 72 (2018) 447–470.

L. Bisio, Reconstructing a short-term indicator by state-space models: An application to estimate hours worked by quarterly national accounts, Journal of Official Statistics 40 (2024) 530–553.

R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2024. URL: http://www.R-project.org/.

R. H. Byrd, P. Lu, J. Nocedal, C. Zhu, A limited memory algorithm for bound constrained optimization, SIAM Journal on scientific computing 16 (1995) 1190–1208.

A. C. Harvey, Forecasting, structural time series models and the kalman filter (1990).

J. Durbin, S. J. Koopman, Time series analysis by state space methods, Oxford university press, 2012.

A. Comber, C. Brunsdon, M. Charlton, A spatial analysis of variations in health access: linking geography, socio-economic status and access to health services, International Journal of Health Geographics 10 (2011) 1–11. doi:`10.1186/1476-072X-10-44`.

D. Dai, F. Wang, Geographic disparities in accessibility to food stores in southwest mississippi, Environment and Planning B: Planning and Design 38 (2011) 659–677. doi:`10.1068/b36088`.

J. C. Gower, A general coefficient of similarity and some of its properties, Biometrics 27 (1971) 857–871. doi:`10.2307/2528823`.

M. Maechler, P. Rousseeuw, A. Struyf, M. Hubert, K. Hornik, cluster: Cluster Analysis Basics and Extensions, 2024. URL: `https://CRAN.R-project.org/package=cluster`, r package version 2.1.8.

R. J. A. Little, D. B. Rubin, Statistical Analysis with Missing Data, 2nd ed., Wiley, 2002.

W. Junger, A. P. De Leon, Imputation of missing data in time series for air pollutants, Atmospheric Environment 102 (2015) 96–104.

I. Jolliffe, Principal component analysis, in: International encyclopedia of statistical science, Springer, Berlin, 2011, pp. 1094–1096.

J. Durbin, Estimation of parameters in time-series regression models, Journal of the Royal Statistical Society. Series B 22 (1960) 139–153.

T. Proietti, Temporal disaggregation by state space methods: Dynamic regression and cointegration, Statistical Methods & Applications 15 (2006) 59–87.

L.-F. Lee, Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models, Econometrica 72 (2004) 1899–1925.

Appendix A. Proof of Theorem 1

Proof. First,

$$Y_{it} = \rho \sum_{j=1}^{n} w_{ij} Y_{jt} + \mathbf{z}_{it}^{\top}\boldsymbol{\beta} + u_{it}$$

$$\sum_{i=1}^{n} Y_{it} = \rho \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} Y_{jt} + \sum_{i=1}^{n}\mathbf{z}_{it}^{\top}\boldsymbol{\beta} + \sum_{i=1}^{n} u_{it}$$

$$Y_t = \rho \sum_{j=1}^{n} Y_{jt} \sum_{i=1}^{n} w_{ij} + \mathbf{z}_t^{\top}\boldsymbol{\beta} + u_t$$

If Assumption 2 are not satisfied by the matrix $\mathbf{W}$, then, $\sum_{i=1}^{n} w_{ij} = 1$ for $i = 1, \ldots, n$, and then,

$$Y_t = \rho \sum_{j=1}^{n} Y_{jt} + \mathbf{z}_t^{\top}\boldsymbol{\beta} + u_t \tag{A.1}$$

$$Y_t = \rho Y_t + \mathbf{z}_t^{\top}\boldsymbol{\beta} + u_t \tag{A.2}$$

$$(1-\rho)Y_t = +\mathbf{z}_t^{\top}\boldsymbol{\beta} + u_t \tag{A.3}$$

$$\tag{A.4}$$

hence, the model $Y_t = \mathbf{z}_t^{\top}\frac{1}{1-\rho}\boldsymbol{\beta} + \tilde{u}_t$ always satisfies the Equation (2) for all $\rho \in (-1, 1)$, and the model will be non identifiable. Let $\ln L(\boldsymbol{\theta})$ defined as:

$$\ln L(\boldsymbol{\theta}) = -\frac{1}{2}\log |\boldsymbol{\Sigma}_Y| - \frac{1}{2}(\mathbf{Y}_a - \boldsymbol{\mu})^{\top}\boldsymbol{\Sigma}_Y^{-1}(\mathbf{Y}_a - \boldsymbol{\mu})$$

where

$$\boldsymbol{\mu} = \mathbf{CA}^{-1}\mathbf{Z}\boldsymbol{\beta}$$

$$\boldsymbol{\Sigma}_Y = \mathbf{CA}^{-1}(\boldsymbol{\Sigma}_U \otimes \mathbb{I}_n)(\mathbf{A}^{-1})^{\top}\mathbf{C}^{\top}$$

Hence, the gradient of $\ln L(\boldsymbol{\theta})$ is:

$$\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} = \left(\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\beta}}\right)^{\top} \boldsymbol{\Sigma}_Y^{-1} \mathbf{e} = (\mathbf{Z}^{\top}(\mathbf{A}^{-1})^{\top}\mathbf{C}^{\top})\boldsymbol{\Sigma}_Y^{-1}\mathbf{e} \tag{A.5}$$

$$\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \sigma^2} = \frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \sigma^2}\right) - \frac{1}{2}\mathbf{e}^{\top}\frac{\partial \boldsymbol{\Sigma}_Y^{-1}}{\partial \sigma^2}\mathbf{e}$$

$$= \frac{T}{2\sigma^2} - \frac{1}{2\sigma^2}\mathbf{e}^{\top}\boldsymbol{\Sigma}_Y^{-1}\mathbf{e} \tag{A.6}$$

$$\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \phi_1} = \frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \phi_1}\right) - \frac{1}{2}\mathbf{e}^{\top}\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \phi_1}\boldsymbol{\Sigma}_Y^{-1}\mathbf{e} \tag{A.7}$$

$$\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \phi_1} = \mathbf{C}\mathbf{A}^{-1}\left(\frac{\partial \boldsymbol{\Sigma}_U}{\partial \phi_1}\right)(\mathbf{A}^{-1})^{\top}\mathbf{C}^{\top}$$

$$\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \rho} = \frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \rho}\right) - \left(\frac{\partial \boldsymbol{\mu}}{\partial \rho}\right)^{\top}\boldsymbol{\Sigma}_Y^{-1}\mathbf{e} + \frac{1}{2}\mathbf{e}^{\top}\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \rho}\boldsymbol{\Sigma}_Y^{-1}\mathbf{e} \tag{A.8}$$

where:

$$\mathbf{e} = \mathbf{Y}_a - \boldsymbol{\mu} = \mathbf{C}\mathbf{A}^{-1}\mathbf{U} = \mathbf{C}\mathbf{A}^{-1}(\mathbf{R} \otimes \mathbb{I}_n)\boldsymbol{\epsilon} = \mathbf{M}\boldsymbol{\epsilon}$$

$$\boldsymbol{\epsilon} = (\epsilon_{11}, \epsilon_{12}, \cdots, \epsilon_{1T}, \epsilon_{21}, \cdots, \epsilon_{2T}, \cdots, \epsilon_{nT})^{\top}$$

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \phi_1 & 1 & 0 & \cdots & 0 \\ \phi_1^2 & \phi_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_1^{T-1} & \phi_1^{T-2} & \phi_1^{T-3} & \cdots & 1 \end{pmatrix}$$

Now, under Assumption 1, let $\mathbf{P}$ and $\mathbf{L}$ matrices of size $T \times T$, and $\mathbf{N}$ a matrix of size $1 \times T$, γ_1 the skewness coefficient and γ_2 the kurtosis coefficient, it is

clear that:

$$\mathbb{E}(\mathbf{e}) = \mathbf{0} \tag{A.9}$$

$$\mathbb{E}(\mathbf{e}\mathbf{e}^\top) = \mathbb{E}(\mathbf{M}\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top\mathbf{M}^\top) = \sigma^2\mathbf{M}\mathbf{M}^\top = \mathbf{C}\boldsymbol{\Sigma}_Y\mathbf{C}^\top \tag{A.10}$$

$$\mathbb{E}(\mathbf{e}^\top\mathbf{e}) = \mathrm{tr}(\mathbb{E}(\mathbf{e}\mathbf{e}^\top)) = \sigma^2\mathrm{tr}(\mathbf{M}\mathbf{M}^\top) = \sigma^2\sum_{t=1}^{T}(\mathbf{M}\mathbf{M}^\top)_{tt} \tag{A.11}$$

$$\mathbb{E}(\mathbf{N}\mathbf{e}\mathbf{e}^\top\mathbf{P}\mathbf{e}) = \sum_{i=1}^{n}\sum_{t=1}^{T}(\mathbf{N}\mathbf{M})_{it}(\mathbf{M}^\top\mathbf{P}\mathbf{M})_{it,it}\mathbb{E}(\epsilon_{it}^3)$$

$$= \sigma^3\gamma_1\sum_{i=1}^{n}\sum_{t=1}^{T}(\mathbf{N}\mathbf{M})_{it}(\mathbf{M}^\top\mathbf{P}\mathbf{M})_{it,it} \tag{A.12}$$

$$\mathbb{E}(\mathbf{e}^\top\mathbf{L}\mathbf{e}\mathbf{e}^\top\mathbf{P}\mathbf{e}) = \sum_{i=1}^{n}\sum_{t=1}^{T}(\mathbf{M}^\top\mathbf{L}\mathbf{M})_{it,it}(\mathbf{M}^\top\mathbf{P}\mathbf{M})_{it,it}\mathbb{E}(\epsilon_{it}^4)+$$

$$\sum_{i\neq j}^{n}\sum_{t\neq k}^{T}(\mathbf{M}^\top\mathbf{L}\mathbf{M})_{jk,it}(\mathbf{M}^\top\mathbf{P}\mathbf{M})_{it,jk}\mathbb{E}(\epsilon_{it}^2)\mathbb{E}(\epsilon_{jk}^2)$$

$$= (3+\gamma_2)\sigma^4\sum_{i=1}^{n}\sum_{t=1}^{T}(\mathbf{M}^\top\mathbf{L}\mathbf{M})_{it,it}(\mathbf{M}^\top\mathbf{P}\mathbf{M})_{it,it}+$$

$$\sigma^4\sum_{i\neq j}^{n}\sum_{t\neq k}^{T}(\mathbf{M}^\top\mathbf{L}\mathbf{M})_{jk,it}(\mathbf{M}^\top\mathbf{P}\mathbf{M})_{it,jk} \tag{A.13}$$

Hence, using Equations (A.5)–(A.8) and (A.9)–(A.13) it is obtained that:

$$\mathbf{E}\left(\frac{\partial\ln L(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}}\frac{\partial\ln L(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}^\top}\right) = (\mathbf{Z}^\top(\mathbf{A}^{-1})^\top\mathbf{C}^\top)\boldsymbol{\Sigma}_Y^{-1}\mathbf{C}\mathbf{A}^{-1}\mathbf{Z}$$

$$= -\mathbf{E}\left(\frac{\partial^2\ln L(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^\top}\right) \tag{A.14}$$

$$\mathbf{E}\left(\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}}\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \sigma^2}\right) = -\frac{1}{2}\gamma_1\sigma^3\sum_{i=1}^{n}\sum_{t=1}^{T}((\mathbf{Z}^\top(\mathbf{A}^{-1})^\top\mathbf{C}^\top)\boldsymbol{\Sigma}_\mathbf{Y}^{-1}\mathbf{M})_{it}\times$$

$$(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it} = \mathcal{K}_2 \tag{A.15}$$

$$\mathbf{E}\left(\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}\partial \sigma^2}\right) = \mathbf{E}\left((\mathbf{Z}^\top(\mathbf{A}^{-1})^\top\mathbf{C}^\top)\boldsymbol{\Sigma}_Y^{-1}\mathbf{e}\right) = \mathbf{0}$$

$$\mathbf{E}\left(\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}}\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \phi_1}\right) = -\frac{1}{2}\gamma_1\sigma^3\sum_{i=1}^{n}\sum_{t=1}^{T}((\mathbf{Z}^\top(\mathbf{A}^{-1})^\top\mathbf{C}^\top)\boldsymbol{\Sigma}_\mathbf{Y}^{-1}\mathbf{M})_{it}\times$$

$$(\mathbf{M}^\top\boldsymbol{\Sigma}_\mathbf{Y}^{-1}\frac{\partial \boldsymbol{\Sigma}_\mathbf{Y}}{\partial \phi_\mathbf{1}}\boldsymbol{\Sigma}_\mathbf{Y}^{-1}\mathbf{M})_{it,it}\mathcal{K}_3 \tag{A.16}$$

$$\mathbf{E}\left(\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}\partial \phi_1}\right) = \mathbf{E}\left((\mathbf{Z}^\top(\mathbf{A}^{-1})^\top\mathbf{C}^\top)\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \phi_1}\boldsymbol{\Sigma}_Y^{-1}\mathbf{e}\right) = \mathbf{0}$$

$$\mathbf{E}\left(\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}}\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \rho}\right) = \mathbf{Z}^\top(\mathbf{A}^{-1})^\top\mathbf{C}^\top)\boldsymbol{\Sigma}_Y^{-1}\left(\frac{\partial \boldsymbol{\mu}}{\partial \rho}\right) +$$

$$-\frac{1}{2}\gamma_1\sigma^3\sum_{i=1}^{n}\sum_{t=1}^{T}((\mathbf{Z}^\top(\mathbf{A}^{-1})^\top\mathbf{C}^\top)\boldsymbol{\Sigma}_\mathbf{Y}^{-1}\mathbf{M})_{it}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \phi_1}\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}$$

$$= \mathbf{E}\left(\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}\partial \rho}\right) - \frac{1}{2}\gamma_1\sigma^3\sum_{i=1}^{n}\sum_{t=1}^{T}((\mathbf{Z}^\top(\mathbf{A}^{-1})^\top\mathbf{C}^\top)\boldsymbol{\Sigma}_\mathbf{Y}^{-1}\mathbf{M})_{it}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \phi_1}\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}$$

$$\mathbf{E}\left(\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}\partial \rho}\right) - \mathcal{K}_4 \tag{A.17}$$

$$\mathbf{E}\left(\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \sigma^2}\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \sigma^2}\right) = \frac{1}{4}(3+\gamma_2)\sum_{i=1}^{n}\sum_{t=1}^{T}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}^2+$$

$$\frac{1}{4}\sum_{i\neq j}^{n}\sum_{t\neq k}^{T}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{jk,it}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,jk}$$

$$= \mathbf{E}\left(\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \sigma^4}\right) + \frac{\gamma_2}{4}\sum_{i=1}^{n}\sum_{t=1}^{T}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}^2$$

$$= \mathbf{E}\left(\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \sigma^4}\right) + \mathcal{K}_5 \tag{A.18}$$

$$\mathbf{E}\left(\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \sigma^2}\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \phi_1}\right) = -\frac{T}{4\sigma^2}\mathrm{tr}\left(\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \phi_1}\right)+$$

$$-\frac{1}{4\sigma^2}(3+\gamma_2)\sigma^4\sum_{i=1}^{n}\sum_{t=1}^{T}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \phi_1}\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}+$$

$$\frac{\sigma^4}{4\sigma^2}\sum_{i\neq j}^{n}\sum_{t\neq k}^{T}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{jk,it}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \phi_1}\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,jk}$$

$$= \mathbf{E}\left(\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \sigma^2\partial \phi_1}\right) + \frac{\gamma_2\sigma^2}{4}\sum_{i=1}^{n}\sum_{t=1}^{T}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \phi_1}\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}$$

$$= \mathbf{E}\left(\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \sigma^2\partial \phi_1}\right) + \mathcal{K}_6 \tag{A.19}$$

$$\mathbf{E}\left(\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \sigma^2}\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \rho}\right) = -\frac{T}{4\sigma^2}\operatorname{tr}\left(\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \rho}\right) +$$

$$-\frac{1}{2\sigma^2}\gamma_1\sigma^3\sum_{i=1}^{n}\sum_{t=1}^{T}(\frac{\partial \boldsymbol{\mu}}{\partial \rho}\boldsymbol{\Sigma}_Y^{-1}M)_{it}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}$$

$$-\frac{1}{4\sigma^2}(3+\gamma_2)\sigma^4\sum_{i=1}^{n}\sum_{t=1}^{T}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \rho}\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}+$$

$$\frac{\sigma^4}{4\sigma^2}\sum_{i\neq j}^{n}\sum_{t\neq k}^{T}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{jk,it}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \rho}\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,jk}$$

$$= \mathbf{E}\left(\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \sigma^2\partial \rho}\right) - \frac{1}{2\sigma^2}\gamma_1\sigma^3\sum_{i=1}^{n}\sum_{t=1}^{T}(\frac{\partial \boldsymbol{\mu}}{\partial \rho}\boldsymbol{\Sigma}_Y^{-1}M)_{it}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}+$$

$$\frac{\gamma_2\sigma^2}{4}\sum_{i=1}^{n}\sum_{t=1}^{T}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \rho}\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}$$

$$= \mathbf{E}\left(\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \sigma^2\partial \rho}\right) - \mathcal{K}_7 \tag{A.20}$$

Analogously to Equation (A.18) it is obtained that:

$$\mathbf{E}\left(\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \phi_1}\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \phi_1}\right) = \mathbf{E}\left(\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \sigma^4}\right) + \frac{\gamma_2}{4}\sum_{i=1}^{n}\sum_{t=1}^{T}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \rho}\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}^2$$

$$= \mathbf{E}\left(\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \sigma^4}\right) + \mathcal{K}_8 \tag{A.21}$$

and analogously to Equation (A.20), it is obtained that:

$$\mathbf{E}\left(\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \phi}\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \rho}\right) = \mathbf{E}\left(\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \phi \partial \rho}\right) -$$

$$\frac{1}{2\sigma^2}\gamma_1\sigma^3 \sum_{i=1}^{n}\sum_{t=1}^{T}(\frac{\partial \boldsymbol{\mu}}{\partial \rho}\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \phi_1}\boldsymbol{\Sigma}_Y^{-1}M)_{it}(\mathbf{M}^\top \boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}+$$

$$\frac{\gamma_2\sigma^2}{4}\sum_{i=1}^{n}\sum_{t=1}^{T}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \phi_1}\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \rho}\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}$$

$$= \mathbf{E}\left(\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \phi \partial \rho}\right) - \mathcal{K}_9 \tag{A.22}$$

and, finally,

$$\mathbf{E}\left(\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \rho}\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \rho}\right) = \mathbf{E}\left(\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \rho^2}\right) -$$

$$\frac{1}{2\sigma^2}\gamma_1\sigma^3 \sum_{i=1}^{n}\sum_{t=1}^{T}(\frac{\partial \boldsymbol{\mu}}{\partial \rho}\boldsymbol{\Sigma}_Y^{-1}M)_{it}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \rho}\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}+$$

$$\frac{\gamma_2\sigma^2}{4}\sum_{i=1}^{n}\sum_{t=1}^{T}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \rho}\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}(\mathbf{M}^\top\boldsymbol{\Sigma}_Y^{-1}\frac{\partial \boldsymbol{\Sigma}_Y}{\partial \rho}\boldsymbol{\Sigma}_Y^{-1}\mathbf{M})_{it,it}$$

$$= \mathbf{E}\left(\frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \rho^2}\right) - \mathcal{K}_{10} \tag{A.23}$$

Then using Equations (A.14) to (A.23), and Under assumptions 1, 2, 3, 4 and applied the Theorem 3.2 of Lee (2004) it is obtained that:

$$\boldsymbol{\Sigma}_\theta = -\lim_{T\to\infty}\mathbb{E}\left(\frac{1}{T}\frac{\partial^2 \ln L(\theta)}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}^\top}\right) \tag{A.24}$$

$$\boldsymbol{\Delta}_\theta = \lim_{T\to\infty}\boldsymbol{\Delta}_{\theta,T} \tag{A.25}$$

$$\boldsymbol{\Delta}_{\theta,T} = \begin{pmatrix} 0 & \mathcal{K}_2 & \mathcal{K}_3 & \mathcal{K}_4 \\ \mathcal{K}_2 & \mathcal{K}_5 & \mathcal{K}_6 & \mathcal{K}_7 \\ \mathcal{K}_3 & \mathcal{K}_6 & \mathcal{K}_8 & \mathcal{K}_9 \\ \mathcal{K}_4 & \mathcal{K}_7 & \mathcal{K}_9 & \mathcal{K}_{10} \end{pmatrix}$$

$\square$

37