Note to the Reader: *This document is a technical excerpt from my Master's Thesis "Modelling price interdependencies between the European carbon and electricity markets: a Global Vector Autoregressive approach".*

- *References to external chapters or sections (e.g., "Appendix", (some) "Equations") refer to the full original manuscript and should be disregarded in this standalone document.*
- *The bibliography provided at the end has been filtered to include only the references cited within this technical appendix.*

1.1. Econometric modelling

This section presents the technical underpinnings and theoretical foundations of the applied model. While certain generalisations are introduced for clarity, the exposition remains as faithful as possible to the specific implementation used in this study, highlighting its essential features and maintaining a close alignment between the theoretical framework and its practical realisation.

1.1.1. Country-specific models

The GVAR model is built upon a collection of country-specific VARX models. These are then subsequently aggregated to obtain the final global VAR. In this study, however, each of the 23 country-specific models (as this is the number of entities included within this research) takes the simpler form of an ARX. The reason behind this choice stems from the nature of the variables employed.

For each Member State, the focal variable is represented by the specific wholesale electricity price return. Country-specific control variables—which are renewable generation indices for wind and solar as well as climatic measures for temperature and precipitation—are treated as exogenous. In addition, though, a set of global factors—returns on EUAs, natural gas, coal, Brent oil, and the STOXX 600 index—applies uniformly across all markets. Within a standard GVAR, any series designated as endogenous in a country-specific model remains endogenous in the final global system too. If one were to treat these global factors as endogenous in every country's VARX, the global aggregation would imply multiple, inconsistent equations for the same variable: a situation which is, of course, not reasonable nor feasible. One might instead include endogenously these common factors in just one country's model, since the final aggregation would result in only one equation to describe the dynamic process for each of these five variables. This procedure, however, would introduce artificial heterogeneity among the selected countries since each individual model would be fundamentally quite different from the one where these global series are endogenously included.

Given that cross-country comparability is particularly relevant for this thesis's purposes, all five global factors are specified as exogenous in each nation's model,

alongside the respective country-specific exogenous controls. Consequently, the endogenous variable of interest is just one for each of the selected entities, which is why the starting point of the development of the whole model considers 23 ARX specifications instead of the same number of VARX.

An argument could be made on the fact that this specification overlooks the possibility of reverse causality: for example, that fluctuations in wholesale electricity prices could influence EUA (or other input) prices other than the converse. As previously mentioned within the review of the pertinent research, empirical studies generally support the original direction of effect as the predominant relationship in historical data. Nevertheless, much of this literature predates recent market developments, and it is conceivable that the directional nexus has evolved. For the reasons outlined above, I leave this avenue of investigation for further research.

To make the econometric exposition clearer, I begin by classifying the available data into four groups of variables:

- domestic: these serve as the dependent series in each of the 23 ARX models—namely, the wholesale electricity price returns. They enter endogenously into the global VAR. I will indicate with y_{it} the value of the domestic variable for country i at time t;
- foreign: country-specific variables, which for each ARX model are defined as a linear combination of other countries' domestic variables. An explanation of how these are computed and what they represent is provided later on. Similarly to domestic variables, I will indicate foreign ones with y_{it}^{*};
- country-specific exogenous: all the exogenous factors that vary across entities—namely, wind and solar generation as well as temperature and precipitation indices. I will call them $wind_{it}, solar_{it}, temp_{it}$ and $prec_{it}$ respectively, and to streamline the exposition, they will be included in the vector $\boldsymbol{x}_{it}^{\boldsymbol{c}}$:

$$\boldsymbol{x}_{it}^{\boldsymbol{c}} = \begin{bmatrix} wind_{it} \\ solar_{it} \\ temp_{it} \\ prec_{it} \end{bmatrix};$$

- global exogenous: those that are common to all the 23 countries and that are treated as exogenous—namely, EUA, natural gas, coal, Brent oil and STOXX 600 returns. These will be indicated with $EUA_t, gas_t, coal_t, brent_t$ and $stoxx_t$ respectively, and summarised in the vector $\boldsymbol{x}_t^{\boldsymbol{g}}$:

$$\boldsymbol{x}_t^{\boldsymbol{g}} = \begin{bmatrix} EUA_t \\ gas_t \\ coal_t \\ brent_t \\ stoxx_t \end{bmatrix}.$$

Given these modelling choices and the emphasis of this thesis on explanation rather than prediction, since the EUA price is specified as a global exogenous variable in both the country-specific ARX models and ultimately in the global VAR, no structural identification scheme is required. The final model remains in its reduced form because treating this variable as exogenous inherently precludes reverse causality—an issue that would arise if it were included as endogenous instead. The interpretation of results, therefore, proceeds via response functions generated by shocks to the exogenous EUA series. Lastly, since an ARX is formally equivalent to a univariate VARX, its use at the country level and subsequent aggregation into a global VAR framework remains entirely valid.

For each country i, the analysis commences with the specification of a basic ARX(p_i, k_i) model:

$$y_{it} = \mu_{i0} + \sum_{j=1}^{6} \delta_{ij} d_t^j + \sum_{j=1}^{p_i} \rho_{ij} y_{i,t-j} + \sum_{j=0}^{k_i} \left(\lambda_{ij} y_{i,t-j}^* + \boldsymbol{\beta}_{ij}^{\boldsymbol{c}\prime} \boldsymbol{x}_{i,t-j}^{\boldsymbol{c}} + \boldsymbol{\beta}_{ij}^{\boldsymbol{g}\prime} \boldsymbol{x}_{t-j}^{\boldsymbol{g}} \right) + \varepsilon_{it}. \quad (0.1)$$

The number of estimated models is denoted by N (throughout this thesis, N will be equal to 23), therefore $i = 1, 2, \dots, N$.

In Equation ((0.1) the first two terms represent the deterministic components of the individual models, which for all countries are given by the constant term and six seasonal dummy variables (d_t^j with $j = 1, 2, \dots, 6$). The third term represents the autoregressive component of the ARX, where p_i denotes the maximum order of lags for country i. Similarly, k_i represents the maximum order of lags of the exogenous part. This includes both the two groups of "purely" exogenous variables (i.e., country-specific and global) plus the foreign variable. While the latter is treated as exogenous in this first step of individual models' estimation, when computing the global model, it will enter as endogenous within the whole system of equations. To streamline the estimation of the optimal lag lengths (which will be detailed later in this subsection) and to maintain clarity throughout the development of this model, I considered a common maximum lag order k_i on all exogenous components—whether foreign, country-specific, or global. $\boldsymbol{\beta}_{ij}^{\boldsymbol{c}}$ and $\boldsymbol{\beta}_{ij}^{\boldsymbol{g}}$ respectively include the coefficients associated with the two sets of country-specific and global exogenous variables, and are such that:

$$\boldsymbol{\beta}_{ij}^{\boldsymbol{c}} = \begin{bmatrix} \beta_{ij}^{wind} \\ \beta_{ij}^{solar} \\ \beta_{ij}^{temp} \\ \beta_{ij}^{prec} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\beta}_{ij}^{\boldsymbol{g}} = \begin{bmatrix} \beta_{ij}^{EUA} \\ \beta_{ij}^{gas} \\ \beta_{ij}^{coal} \\ \beta_{ij}^{brent} \\ \beta_{ij}^{stoxx} \end{bmatrix}.$$

At last, considerations about the distribution of ε_{it} are provided at a later stage as they are associated with subsequent steps of the modelling process too.

I now turn to a discussion of the key features of this initial estimation stage, both to complete the exposition and to motivate the choices made in fitting the N individual ARX models.

Deterministic components and seasonality

I remind the reader that y_{it} represents the log-returns of wholesale electricity prices for country i. Accordingly, each of the 23 models includes an intercept—allowing the processes to have a nonzero mean—but omits any trend, since the visual inspection of the return series confirms the absence of a systematic trend.

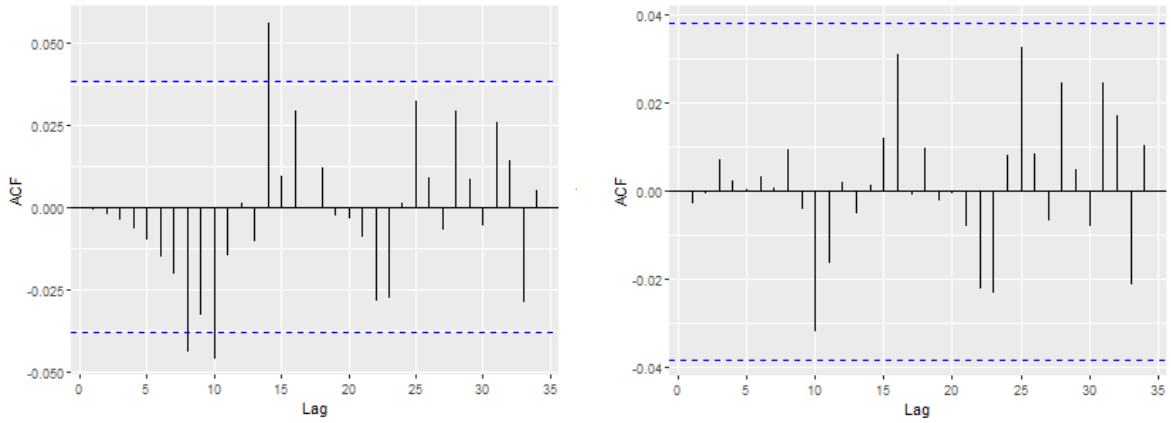
The seasonal component, however, deserves particular attention. Although the descriptive plots of the electricity price series did not display evident seasonality patterns, the significant level of noise in the daily return and the extensive length of the time series might obscure this feature. In fact, the dynamics that characterise electricity markets imply the presence of a seasonal component affecting electricity prices—and, consequently, their returns. For instance, weekday industrial and commercial activity drives higher consumption levels, whereas weekends face reduced demand, which, is reflected in the day-ahead price-formation mechanism. To account for these effects, each country's ARX model incorporates six (to avoid perfect collinearity with the constant term) dummy variables. In every case, these seasonal regressors showed statistical significance, confirming their importance in all the individual models.

To be more precise, even if seasonality were detected in only one or a few countries, the corresponding weekly dummies would still have to be included in every equation. While this would not have been strictly necessary if one confined themselves to the estimation of individual models, it becomes relevant when all the ARX specifications are aggregated into a global VAR. This is because in a standard VAR, the deterministic components specified for any jointly endogenous variable apply uniformly across all equations—regardless of whether some series exhibit the assumed pattern—so it is possible to assess how deviations of one variable from, for instance, its seasonal cycle affect deviations of another around its mean, even if the latter lacks seasonality. The same logic governs, of course, other deterministic terms. Since the 23 country-level ARX will be aggregated into a single GVAR, these conventions must be respected.

Upon examining the empirical autocorrelation function (ACF) of the residuals from the country-level ARX models, I observed that, in some cases, notwithstanding the inclusion of weekly dummies, seasonal effects were still not fully captured: persistent

and statistically significant autocorrelation spikes appeared at 7-day or 14-day lags.[1] This pattern implies that an innovation on day t is correlated with innovations on days $t-7$ or $t-14$, clearly indicating that the weekly seasonal cycle has not been fully accounted for; Figure 0.1a provides a practical example encountered when dealing with the ARX of Denmark, but a similar phenomenon happened for Czechia, France, Greece, Italy, Lithuania, the Netherlands, Norway and Portugal too.

The persistence of such a significant autocorrelation despite the inclusion of seasonal dummies indicates that the weekly cycle of electricity-price returns may, in some cases, evolve rather than remain constant. Given that my sample spans 2635 daily observations—thus covering over seven years—it is plausible that shifts in weekly demand or supply dynamics have occurred. Fixed deterministic dummies effectively capture stable and regular patterns, but they cannot accommodate temporal changes in seasonality, which therefore may manifest in the residuals. To address this issue, I implemented within the affected country models a first-order seasonal autoregressive polynomial in L^7. This strategy successfully removed the significant autocorrelations in the innovations, yielding cleaner residuals' diagnostics (see Figure 0.1b for a comparison of the Danish case).



(a) Seasonal dummies only.

(b) Seasonal dummies and polynomial.

Figure 0.1 Empirical ACF of Denmark's ARX residuals.

Contrary to what one might conclude, introducing a seasonal AR(1) polynomial at lag 7 does not unduly complicate the subsequent GVAR computation. In each ARX model where it is applied, the maximum autoregressive lag simply increases from p_i

[1] When performing this type of diagnostics, it is reasonable to expect a few values of the empirical residual autocorrelation function to be outside the confidence bands, since multiple testing across many lags are performed and there is always the possibility of a type I error. However, the recurrence of significant spikes precisely at multiples of seven days is unlikely to be attributable to this aspect.

to $p_i + 7$ with all intermediate domestic-lag coefficients constrained to zero. Moreover, the coefficients from lags 7 to $p_i + 7$ derive directly from the seasonal-AR parameter multiplied by the relevant lower-order domestic coefficients, so that only one additional parameter—the seasonal-AR coefficient—requires estimation. As a result, this approach (which is not explicitly shown in Equation (0.1) to avoid clutter) adds minimal complexity while effectively eliminating the residual autocorrelation. Seasonal dummies, however, remain in every ARX specification—both because they proved significant in all cases and because, as said, deterministic components must be handled consistently in light of the subsequent development of the global VAR. I will denote by s_i the order of the seasonal polynomial (recalling that this refers to a polynomial in L^7) for model i. Therefore, in my case, s_i equals one for all countries where I applied this further seasonal adjustment, and zero otherwise.

Constructing foreign variables

The foreign variables serve to capture potential spillover effects between each country and the rest of the environment. As such, they represent the level of openness of the market under analysis to all other markets, effectively measuring its exposure to the remaining $N - 1$ economies. Although treated as exogenous in each country-level model, these foreign aggregates are the linchpin that allows to transform a collection of isolated ARX specifications into a cohesive GVAR that explicitly accounts for cross-country interdependencies.

Each foreign variable is constructed as a weighted sum of the domestic variables of all other economies. The question then becomes: which weights should be applied? I derived a country-pair weights matrix from bilateral metrics that reflect the strength of each interconnection. Given the frequent macroeconomic applications of GVAR models, this matrix is typically built upon trade flows: each w_{ij} is defined as the share of country j in the total trade level (accounting for both imports and exports) of country i on the premise that in these contexts, spillovers operate through trade linkages. For day-ahead electricity markets, however, trade volumes are not the most suitable conduit. I therefore constructed the weights starting from cross-border net transfer capacities of electricity.

In the obtained weights matrix $\boldsymbol{W}$—where w_{ij} is represented by the value in the cell corresponding to row i and column j—all the elements on the diagonal are considered to be zero. Foreign variables are hence defined as:

$$y_{it}^{*} = \sum_{j=1}^{N} w_{ij} y_{jt} \qquad \text{such that:} \quad \sum_{j=1}^{N} w_{ij} = 1 \quad \text{and} \quad w_{ii} = 0 \quad \forall i. \tag{0.2}$$

Estimation-related details

I estimated each country-level ARX model using the Arima() function from Hyndman et al.'s "forecast" package in R [5]. Parameter estimation proceeded via conditional maximum likelihood—equivalent to ordinary least squares (OLS). I could have used exact maximum likelihood as well, but this choice stems from an evaluation of the benefits and disadvantages of both alternatives, contextualised within the scope of the investigation.

Exact maximum likelihood is, in principle, more efficient because it exploits all the available observations in the sample, considering just for the first $p_i + 7s_i$ observations the unconditional distribution of the process. Conditional maximum likelihood instead conditions on those first initial values without imposing any distributional assumption upon them, thus making use of fewer observations for estimation purposes. Using the unconditional distribution for the first data points, however, implies the assumption that those initial observations were drawn from a process that has persisted since the infinite past with the same mechanism that characterises all subsequent values of the time series, an assumption that OLS does not hinge upon. This idea is particularly relevant if the analysis is developed over a period of time, which is judged as interesting because changing dynamics might have occurred, and shifts in the interplay between the observed variables may have happened. In fact, one of the reasons for selecting the temporal coverage studied throughout this thesis is precisely that, beginning in 2018, the EU ETS underwent substantial structural shifts. Hence, I chose conditional estimation, despite relinquishing some information from the initial observations.

Regarding the error terms, the Arima() routine enforces the assumption of $\varepsilon_{it} \overset{i.i.d}{\sim} \mathcal{N}(0; \sigma_i^2)$ as it does not allow alternative distributional forms, nor specifications for the innovations. However, the residuals of the individual estimated models often exhibit heavier tails than the Gaussian for several countries. Nonetheless, I maintained the normal-innovation assumption in this first estimation stage because, to build the global VAR, what is of interest are the coefficient estimates for the ARX regressors (that will feed into the GVAR calculation process). And even if the distribution of the errors deviates from normality, the consistency of OLS estimates still holds as long as the regressors are not perfectly multicollinear—which is more of a fundamental requirement to be able to compute the OLS estimator at all—and are exogenous, i.e.,

uncorrelated with the error term.[2] The variables incorporated into the model were also selected to support the plausibility of this assumption.

That said, when deriving the final results, it is essential to go beyond mere point estimates of the coefficients, out of which response functions can be obtained. These alone offer little insight if not coupled with a proper measure of their uncertainty. And it is here that departures from normality in the residuals become problematic. To overcome this challenge, I implemented a bootstrap procedure that preserves the empirical distribution and covariance structure of the innovations, thereby yielding reliable uncertainty estimates for all inferred quantities. Subsection 1.2.4 provides a detailed account of this methodology.

Optimal lag order selection

In Equation (0.1) it can be noticed how each country's ARX model requires choosing the two lag orders p_i and k_i for the autoregressive and exogenous components. Although p_i and k_i may differ across countries, this poses no obstacle for the global aggregation process: the final GVAR has a maximum autoregressive lag of $p = \max\left(\max_i(p_i + 7s_i), \max_i k_i\right)$ and a maximum lag for its exogenous component of $k = \max_i k_i$, with any "missing" coefficients automatically set to zero for countries with shorter lag lengths (Subsection 1.1.2 further elaborates on this point).

To select the proper p_i and k_i for each of the N individual models, I first specified a sufficiently large p_{max} and k_{max}, then proceeded to estimate every ARX(p_i, k_i) which is nested in the ARX(p_{max}, k_{max}). For each Member State, I then chose the model whose Bayesian Information Criterion (BIC) was minimised.[3] In this stage, though, I made sure that all estimations conditioned on the exact same initial observations—namely, the first $p_{max} + 7s_i$ for each evaluation of model i. This ensures comparability of the information criterion results, since, being based on the likelihood, it means that it is a function of the data used to estimate the parameters. And if the underlying estimation

---

[2] In autoregressive settings however, the exogeneity condition $E[\boldsymbol{\varepsilon} \mid \boldsymbol{X}] = \mathbf{0}$ does not strictly hold, since $\boldsymbol{X}$ (the matrix of regressors) includes lagged values of the focal variable $\boldsymbol{y}$, implying that each component ε_t of the vector $\boldsymbol{\varepsilon}$ is correlated with future observations of the regressors. Nonetheless, under stationarity—a property I have verified for all series—the bias that consequently characterizes the coefficient estimates vanishes asymptotically. Given that the sample size is not small, I therefore proceeded with the conditional maximum likelihood estimation as described.

[3] The BIC has been favoured over the Akaike information criterion (AIC) because, although the AIC is still designed to account for parsimony of the evaluated model, it may not sufficiently penalize model complexity [6]. In contrast, since the BIC applies a penalty term that increases with the sample size (while in the case of AIC, this term remains fixed), it reduces the probability of selecting an overparametrized model when a correctly specified, more parsimonious alternative is available.

dataset changes, then the BIC values cannot be properly compared. Once selected, each model was nonetheless re-estimated over its full usable sample.

Although the choices of p_i and k_i suggested by this approach generally yielded acceptable residual diagnostics, a few country models still exhibited some autocorrelation spikes in their empirical residuals ACFs. For those cases, I incrementally raised p_i or k_i by one or two lags until the residuals displayed the desired white-noise properties. After obtaining the N models, a number of considerations were drawn regarding the resulting specifications.

Even though it is not a formal requirement, it is sound practice to compare the magnitude of each country's own autoregressive coefficients (so for its lagged domestic variable): ρ_{ij}, with that for the lagged foreign aggregate: λ_{ij}. Intuitively, one expects a country's own past dynamics to exert a stronger influence than the historical behaviour of its interconnected partners. That said, because the contemporaneous foreign variable y_{it}^{*} is also part of the regressors of each model, and since it is constructed as a weighted average of current domestic values across interlinked markets, its impact (due to its temporal coincidence with the values of the focal variable) is clearly significant and may well exceed that of a past domestic observations. This is especially true in light of the considerations about the European electricity market architecture and integration dynamics. Nonetheless, the relevance of the comparison still holds if one considers lagged values of the domestic and foreign parameters: in most cases, the coefficients of $y_{i,t-j}$ should dominate those of $y_{i,t-j}^{*}$.

For many countries, the optimal exogenous lag order k_i was found to be zero, meaning that only the contemporaneous foreign values entered their ARX models—an outcome that aligns with the expectation of a stronger inertia in domestic lagged (compared to foreign lagged) returns. In the remaining countries, the maximum optimal exogenous lag was found to be 1. In all these instances, the coefficient of the one-day lagged foreign aggregate was uniformly smaller in magnitude than that of the own lagged return. By means of an example, for Italy, the estimate for the AR coefficient $\rho_{13,1}$ (13 represents Italy in the countries' numbering $1, 2, \ldots, N$) is -0.419 while the corresponding foreign parameter $\lambda_{13,1}$ estimate is -0.005. Similarly, Hungary shows the two estimates for $\rho_{12,1}$ and $\lambda_{12,1}$ to be -0.372 and -0.014 respectively. And both series represent log-returns of electricity prices (the foreign aggregate being a weighted average of such returns), so these differences cannot be attributed to scale effects. Although I conducted a rough comparison of statistical significance—using the standard errors reported by the Arima() output to compute p-values—it should be noted that formal inference on these coefficients' variance is not a reliable reference due to the non-Gaussian nature of the innovations (as discussed above). Nonetheless, in every model the own lag coefficients displayed higher "significance levels" than their foreign lag counterparts, further reinforcing the predominance of domestic dynamics. These ad hoc significance checks, however, were performed for indicative purposes only and given their limited accuracy, they are not reported.

One other prudent practice is to inspect the obtained optimal lag orders p_i. Whereas financial equity markets typically exhibit only very short-memory dynamics—often limited to one or two lags—commodity markets such as electricity may require slightly higher autoregressive orders, although still rarely very large. The same principle applies to the exogenous lag order k_i.

As just discussed, no country model saw the selection of a k_i greater than one. There were only seven markets (Austria, Greece, Hungary, Italy, Lithuania, Latvia, and Poland) that retained a single-day exogenous lag ($k_i = 1$), while all others included only contemporaneous exogenous terms.[4] The BIC-guided procedure for choosing p_i yielded autoregressive orders of five or fewer in the majority of country models, with many settling between two and four lags. A handful of exceptions arose, however: some countries showed optimal lag levels of six, only two (the Netherlands and Denmark) went up to seven, and Latvia required even eight lags. Aside from this, even when I occasionally increased p_i to improve the residual autocorrelation structure, the highest lag length across all models remained relatively moderate, peaking at eight.

One reason why the observed lag orders only partly match the expectations can be the fact that the underlying process of the modelled series may be better described by the inclusion of a moving average (MA) component in addition to autoregression. In other words, the inclusion of past shocks in the equations may provide a better representation of some of the dynamic series. Indeed, an invertible MA polynomial—i.e., the polynomial in L multiplying the error term ε_t—can be shown to be equivalent to an infinite-order AR process with exponentially decaying coefficients, and hence well approximated by a finite AR(p) provided that p is sufficiently large. In principle, one could therefore estimate ARMAX models to capture this behaviour more parsimoniously, since the MA structure would require fewer total parameters. However, the AR approximation delivers a highly similar in-sample fit, the only drawback being the greater number of coefficient estimates required. The estimated models have consequently been retained as ARX specifications across all countries. This, on the one hand avoids introducing structural heterogeneity that would largely complicate the global VAR assembly, given that the introduction of a MA component would be suitable only in a few of the N cases. On the other hand, it better aligns with the prevailing practice in the GVAR literature that works with a foundation of AR processes, while it neglects MA components.

---

[4] Recall that, apart from the foreign aggregates—treated identically to the day-ahead electricity price returns—each country's exogenous series (both global and country-specific) are built so that their value at t, from the perspective of the models, matches the previous calendar day (i.e., what market participants see at $t-1$ when bidding for electricity prices that will be delivered at day t).

Accounting for potential cointegration

One of the primary reasons for working with log-returns rather than the original level series is the advantage of operating on stationary processes. Alongside unit-root tests on the log-returns, I also carried out unit-root tests on the logged variables (i.e., without differencing). While the global series do not provide enough evidence to reject the null hypothesis of a unit-root process, many of the logged electricity price series do so, providing evidence in favour of their stationarity. Nonetheless, I preserved log-returns as the main analytical framework. However, the conclusion about the stationarity properties of most logged electricity prices (coupled with the non-stationarity of the corresponding global series) excluded the possibility of modelling cointegration relationships between electricity price logarithms and other potential logged inputs.

The developed framework, however, does not preclude the possibility of modelling cointegration—i.e., the incorporation of error-correction mechanisms into each model accounting for potential long-run equilibrium adjustments between electricity prices and other non-differenced series. Once the error-correction term (ECT) is introduced into the dynamic equation for the variable in the differences, it can be shown that an ARX(p_i, k_i) model as specified in Equation (0.1)—but augmented with an ECT—is equivalent to an ARX$(p_i + 1, k_i + 1)$ in the non-differenced variables. The coefficients of the latter representation can be directly derived from those of the original ARX(p_i, k_i) formulation. This equivalence holds for cointegration relationships between the domestic variable and any exogenous series—whether country-specific, global, or even foreign. Therefore, by starting from the N individual error-correction models, one can reconstruct the corresponding ARX representations in the non-differenced data and use these as inputs for the global VAR construction.[5]

For reasons of brevity—and because it falls outside the requirements of this specific analysis—a detailed technical exposition is omitted. Nevertheless, this paragraph was deemed necessary to highlight the fact that, if needed, cointegration can be seamlessly incorporated into the current econometric framework.

---

[5] Notice that, the foreign variables y_{it}^* can be either obtained as described in Equation (0.2), therefore working directly with returns; or, alternatively, they can be calculated as $\Delta ln^*(P_{it})$ where $ln^*(P_{it}) = \sum_{j=1}^{N} w_{ij} \ln(P_{jt})$. In other words, applying the linear combination to obtain the foreign variables before or after having taken the differences of logarithms does not change the final result. This preserves the underlying framework for constructing a GVAR using non-differenced variables, since the computation of the foreign aggregates remains conceptually unchanged.

Additional remarks

A couple of supplementary aspects have been addressed during the country-specific modelling phase; one of them relates to the treatment of outlying observations. Extreme wholesale electricity prices (and consequently returns) can arise during supply disruptions or episodic demand spikes such as during severe cold snaps or heatwaves, especially if coupled with low renewable generation, or conversely, during periods of exceptionally low demand accompanied by an unusually high renewable output. The results revealed substantial departures from the Gaussian assumption—primarily driven by heavy tails—highlighting that, while extreme observations remain rare, they occur far more frequently than what a normal distribution would predict. Incorporating the resulting outliers into a regression analysis can substantially distort parameter estimates, potentially leading to erroneous conclusions about the true relationship between input costs and wholesale prices.

This issue has been extensively examined in the relevant literature around similar topics. For instance, B. Guo and G. Castagneto Gissey (2021) [2] impose an upper and lower bound on their electricity price series, replacing any observations that fall outside these thresholds with the nearest limit. They report that several parameter estimates differ significantly when outliers are winsorized, ultimately basing most of their conclusions on a model which approximately trims 5% of the total observations (from both tails of the distribution). A very similar procedure was carried on by one of the same authors [1], substituting implausible values with more representative ones to then proceed with the estimation. P.-A. Jouvet and B. Solier (2013) [4] point out that the pronounced volatility of spot electricity prices can heavily distort estimates of carbon cost pass-through; consequently, they drop the outer 5% of observations (from both tails) before re-estimating their model on the remaining 90%, which materially improved in fit. Likewise, A. Bublitz et al. (2017) [3] and S. Tehrani et al. [7] find that excluding outliers yields a superior model specification, emphasising on their treatment as a recommended best practice before estimation.

On the same logic as B. Guo and G. Castagneto Gissey (2021) [2] and P.-A. Jouvet and B. Solier (2013) [4], after conducting multiple estimation experiments, I established the final model specification by setting country-specific upper and lower thresholds so as to leave 5% of the observations from both the left and right tails of the distributions. Any daily return falling outside these bounds was winsorized to the nearest threshold value. Crucially, no observations were removed: the length of the time series remains identical to the original data, only with adjusted extreme values. This strategy enabled first to obtain more parsimonious ARX specifications as the required lag orders were sometimes reduced compared to the original starting point. Second, it enhanced the stability of the GVAR, a critical prerequisite to ensure that the model accurately reflects real-world dynamics; this latter element is discussed in Subsection 1.2.1.

One other aspect concerns peculiar characteristics of a few of the included countries. Some Member States exhibited virtually zero renewable output for specific technologies. In particular, solar generation in Norway and wind generation in Slovenia and Slovakia are negligible due to their minimal installed capacity, making any meaningful impact on wholesale electricity prices implausible. Moreover, including these series in the country-level ARX models leads to estimation problems as for instance, Norway's solar output is disclosed by the data provider as a constant sequence of zeros. Consequently, for these three cases, I excluded the specific renewable generation variable from the ARX estimation. However, as the GVAR construction requires a structurally consistent variable set for all individual models, the coefficients for these omitted regressors were simply imputed as zero in the output of their ARX specification.

Once all the N individual models were estimated, I extracted their coefficient estimates and employed these to construct the global VAR. The details of this procedure are provided in the next subsection.

1.1.2. Obtaining the GVAR model

After estimating each of the N country-specific ARX models, the GVAR is constructed treating all domestic variables as endogenous. For this purpose, it is useful to define the $N \times 1$ "global" vector containing every country's domestic variable:

$$\boldsymbol{y}_t = \begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{Nt} \end{bmatrix}.$$

Starting from the country-specific ARX models[6]

$$\begin{aligned} y_{it} = \mu_{i0} + \sum_{j=1}^{6} \delta_{ij} d_t^j + \rho_{i1} y_{i,t-1} + \cdots + \rho_{ip_i} y_{i,t-p_i} + \lambda_{i0} y_{it}^* + \cdots + \lambda_{ik_i} y_{i,t-k_i}^* \\ + \boldsymbol{\beta}_{i0}^{\boldsymbol{c}\prime} \boldsymbol{x}_{it}^{\boldsymbol{c}} + \cdots + \boldsymbol{\beta}_{ik_i}^{\boldsymbol{c}\prime} \boldsymbol{x}_{i,t-k_i}^{\boldsymbol{c}} + \boldsymbol{\beta}_{i0}^{\boldsymbol{g}\prime} \boldsymbol{x}_t^{\boldsymbol{g}} + \cdots + \boldsymbol{\beta}_{ik_i}^{\boldsymbol{g}\prime} \boldsymbol{x}_{t-k_i}^{\boldsymbol{g}} + \varepsilon_{it}, \end{aligned} \tag{0.3}$$

define the 2×1 vector $\boldsymbol{z}_{it}$ as:

[6] Or, as previously discussed, from the corresponding ARX representations recovered from each individual country-level error correction models. In addition, Equation ((0.3) omits seasonal autoregressive terms from the exposition (thus assuming $s_i = 0$), to make things clearer. Although some models include them, their incorporation does not alter the overall framework, as they merely add lagged variables subject to the same coefficient restrictions described earlier.

$$\boldsymbol{z}_{it} = \begin{bmatrix} y_{it} \\ y_{it}^{*} \end{bmatrix}.$$

Equation ((0.3) can therefore be written as:

$$\begin{aligned} \boldsymbol{A}_{i0}\boldsymbol{z}_{it} = \mu_{i0} + \sum_{j=1}^{6} \delta_{ij} d_t^j + \boldsymbol{A}_{i1}\boldsymbol{z}_{i,t-1} + \cdots + \boldsymbol{A}_{iq_i}\boldsymbol{z}_{i,t-q_i} \\ + \boldsymbol{\beta}_{i0}^{\boldsymbol{c}\,\prime}\boldsymbol{x}_{it}^{\boldsymbol{c}} + \cdots + \boldsymbol{\beta}_{ik_i}^{\boldsymbol{c}\,\prime}\boldsymbol{x}_{i,t-k_i}^{\boldsymbol{c}} + \boldsymbol{\beta}_{i0}^{\boldsymbol{g}\,\prime}\boldsymbol{x}_{t}^{\boldsymbol{g}} + \cdots + \boldsymbol{\beta}_{ik_i}^{\boldsymbol{g}\,\prime}\boldsymbol{x}_{t-k_i}^{\boldsymbol{g}} + \varepsilon_{it}, \end{aligned} \tag{0.4}$$

where

$$q_i = \max(p_i, k_i),$$

and

$$\boldsymbol{A}_{i0} = \begin{bmatrix} 1 & -\lambda_{i0} \end{bmatrix}, \qquad \boldsymbol{A}_{ij} = \begin{bmatrix} \rho_{ij} & \lambda_{ij} \end{bmatrix} \text{ for } j = 1, \dots, q_i.$$

Within Equation ((0.4), whenever the lag orders p_i and k_i differ, resulting in a value for q_i which is greater than one of the two original orders, any "missing" coefficients for the $\boldsymbol{A}_{ij}$ matrices are set to zero. For instance, since each model in this analysis has at most $k_i = 1$, consider the hypothetical case of $p_i = 2$ and $k_i = 1$. In this scenario, the $\boldsymbol{A}_{i0}$ and $\boldsymbol{A}_{ij}$ matrices take the form:

$$\begin{aligned} \boldsymbol{A}_{i0} &= \begin{bmatrix} 1 & -\lambda_{i0} \end{bmatrix}, \\ \boldsymbol{A}_{i1} &= \begin{bmatrix} \rho_{i1} & \lambda_{i1} \end{bmatrix}, \\ \boldsymbol{A}_{i2} &= \begin{bmatrix} \rho_{i2} & 0 \end{bmatrix}. \end{aligned}$$

By extension, the same principle applies when a seasonal polynomial is included ($s_i = 1$) bearing in mind that ρ_{ij} represents the entire coefficient multiplying the term $y_{i,t-j}$ and that in this case $q_i = \max(p_i + 7s_i, k_i)$ while the maximum autoregressive lag part of Equation (0.3) would be $\rho_{i,(p_i+7s_i)}y_{i,t-(p_i+7s_i)}$. Therefore, if a seasonal polynomial is present, ρ_{ij} would be given (for the last p_i lags) by the product of the single seasonal AR parameter ρ_i^s and the shorter-lag AR coefficients with a negative sign in front. In other words, with a seasonal polynomial defined as $(1 - \rho_i^s L^7)$, only the seasonal coefficient ρ_i^s is directly estimated; all higher-order seasonal lags are obtained by multiplying ρ_i^s by the corresponding non-seasonal $\rho_{i1}, \dots, \rho_{ip_i}$. If in the previous example, where $p_i = 2$ and $k_i = 1$, a seasonal polynomial is also present, the $\boldsymbol{A}_{i0}$ and $\boldsymbol{A}_{ij}$ matrices would be:

$$\begin{aligned} \boldsymbol{A}_{i0} &= \begin{bmatrix} 1 & -\lambda_{i0} \end{bmatrix}, & \boldsymbol{A}_{i3} &= \begin{bmatrix} 0 & 0 \end{bmatrix}, & \boldsymbol{A}_{i7} &= \begin{bmatrix} \rho_i^s & 0 \end{bmatrix}, \\ \boldsymbol{A}_{i1} &= \begin{bmatrix} \rho_{i1} & \lambda_{i1} \end{bmatrix}, & &\vdots & \boldsymbol{A}_{i8} &= \begin{bmatrix} -\rho_i^s\rho_{i1} & 0 \end{bmatrix}, \\ \boldsymbol{A}_{i2} &= \begin{bmatrix} \rho_{i2} & 0 \end{bmatrix}, & \boldsymbol{A}_{i6} &= \begin{bmatrix} 0 & 0 \end{bmatrix}, & \boldsymbol{A}_{i9} &= \begin{bmatrix} -\rho_i^s\rho_{i2} & 0 \end{bmatrix}. \end{aligned}$$

The vector $\boldsymbol{z}_{it}$ can, however, be obtained directly from the global vector $\boldsymbol{y}_t$ in the following way:

$$\boldsymbol{z}_{it} = \boldsymbol{W}_i \boldsymbol{y}_t, \tag{0.5}$$

where $\boldsymbol{W}_i$ is a $2 \times N$ link matrix of model i defined from the entries w_{ij} of the global weights matrix $\boldsymbol{W}$. In particular, $\boldsymbol{W}$ can be seen as:

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{\omega}_1' \\ \boldsymbol{\omega}_2' \\ \vdots \\ \boldsymbol{\omega}_N' \end{bmatrix},$$

where $\boldsymbol{\omega}_i$ is the $N \times 1$ vector containing all the weights referred to country i, so that $\boldsymbol{\omega}_i'$ actually represents the i-th row of the matrix $\boldsymbol{W}$; recall that, as specified in Equation ((0.2) all the elements within $\boldsymbol{\omega}_i$ sum to 1 and $w_{ii} = 0$ for all i.

$$\boldsymbol{\omega}_i = \begin{bmatrix} w_{i1} \\ w_{i2} \\ \vdots \\ w_{iN} \end{bmatrix}.$$

It follows that each foreign variable y_{it}^* can be written as:

$$y_{it}^* = \boldsymbol{\omega}_i' \boldsymbol{y}_t, \tag{0.6}$$

a more compact form of the one reported in Equation ((0.2). $\boldsymbol{W}_i$ is therefore constructed with the following specification:

$$\boldsymbol{W}_i = \begin{bmatrix} \boldsymbol{e}_i' \\ \boldsymbol{\omega}_i' \end{bmatrix}, \tag{0.7}$$

where $\boldsymbol{e}_i$ here denotes the $N \times 1$ vector with a 1 in the i-th coordinate and zeros elsewhere. For all i, each $\boldsymbol{W}_i$ is a matrix of known constants. From this definition of $\boldsymbol{W}_i$, it immediately follows that the identity ((0.5) is verified for each i. As a consequence, Equation ((0.4) can be expressed through the equivalence ((0.5) in this manner:

$$\begin{aligned} \boldsymbol{A}_{i0} \boldsymbol{W}_i \boldsymbol{y}_t = \mu_{i0} + \sum_{j=1}^{6} \delta_{ij} d_t^j + \boldsymbol{A}_{i1} \boldsymbol{W}_i \boldsymbol{y}_{t-1} + \cdots + \boldsymbol{A}_{iq_i} \boldsymbol{W}_i \boldsymbol{y}_{t-q_i} \\ + \boldsymbol{\beta}_{i0}^{c\,\prime} \boldsymbol{x}_{it}^c + \cdots + \boldsymbol{\beta}_{ik_i}^{c\,\prime} \boldsymbol{x}_{i,t-k_i}^c + \boldsymbol{\beta}_{i0}^{g\,\prime} \boldsymbol{x}_t^g + \cdots + \boldsymbol{\beta}_{ik_i}^{g\,\prime} \boldsymbol{x}_{t-k_i}^g + \varepsilon_{it}, \end{aligned} \tag{0.8}$$

yielding to an equivalent equation for each entity, which is, however, expressed as a function of the global vector containing all N domestic variables rather than solely the specific domestic and foreign variables of country i. At this point, all the individual specifications are stacked on top of each other to obtain the following model:

$$\boldsymbol{G}_0\boldsymbol{y}_t = \boldsymbol{\mu}_0 + \boldsymbol{\delta}\boldsymbol{d}_t + \boldsymbol{G}_1\boldsymbol{y}_{t-1} + \cdots + \boldsymbol{G}_p\boldsymbol{y}_{t-p} + \boldsymbol{B}_0^{\boldsymbol{c}}\boldsymbol{\gamma}_t^{\boldsymbol{c}} + \cdots + \boldsymbol{B}_k^{\boldsymbol{c}}\boldsymbol{\gamma}_{t-k}^{\boldsymbol{c}} + \boldsymbol{B}_0^{\boldsymbol{g}}\boldsymbol{x}_t^{\boldsymbol{g}} + \cdots + \boldsymbol{B}_k^{\boldsymbol{g}}\boldsymbol{x}_{t-k}^{\boldsymbol{g}} + \boldsymbol{\varepsilon}_t, \tag{0.9}$$

where

$$\boldsymbol{G}_0 = \begin{bmatrix} \boldsymbol{A}_{10}\boldsymbol{W}_1 \\ \boldsymbol{A}_{20}\boldsymbol{W}_2 \\ \vdots \\ \boldsymbol{A}_{N0}\boldsymbol{W}_N \end{bmatrix}, \qquad \boldsymbol{G}_j = \begin{bmatrix} \boldsymbol{A}_{1j}\boldsymbol{W}_1 \\ \boldsymbol{A}_{2j}\boldsymbol{W}_2 \\ \vdots \\ \boldsymbol{A}_{Nj}\boldsymbol{W}_N \end{bmatrix} \text{ for } j = 1, \dots, p,$$

are $N \times N$ matrices,

$$\boldsymbol{B}_j^{\boldsymbol{c}} = \begin{bmatrix} \boldsymbol{\beta}_{1j}^{\boldsymbol{c}\,\prime} & \boldsymbol{0}' & \boldsymbol{0}' & \cdots & \boldsymbol{0}' \\ \boldsymbol{0}' & \boldsymbol{\beta}_{2j}^{\boldsymbol{c}\,\prime} & \boldsymbol{0}' & \cdots & \boldsymbol{0}' \\ \boldsymbol{0}' & \boldsymbol{0}' & \boldsymbol{\beta}_{3j}^{\boldsymbol{c}\,\prime} & \cdots & \boldsymbol{0}' \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0}' & \boldsymbol{0}' & \boldsymbol{0}' & \cdots & \boldsymbol{\beta}_{Nj}^{\boldsymbol{c}\,\prime} \end{bmatrix} \quad \text{and} \quad \boldsymbol{B}_j^{\boldsymbol{g}} = \begin{bmatrix} \boldsymbol{\beta}_{1j}^{\boldsymbol{g}\,\prime} \\ \boldsymbol{\beta}_{2j}^{\boldsymbol{g}\,\prime} \\ \boldsymbol{\beta}_{3j}^{\boldsymbol{g}\,\prime} \\ \vdots \\ \boldsymbol{\beta}_{Nj}^{\boldsymbol{g}\,\prime} \end{bmatrix} \text{ for } j = 0, \dots, k,$$

are $N \times (N \times 4)$ and $N \times 5$ matrices respectively; 4 being the number of country-specific exogenous variables and 5 being the global exogenous variables in each ARX specification. Each $\boldsymbol{0}$ in the matrix $\boldsymbol{B}_j^{\boldsymbol{c}}$ is a 4×1 vector of zeros. For illustrative purposes, the matrix $\boldsymbol{B}_j^{\boldsymbol{c}}$ can be seen as:

$$\boldsymbol{B}_j^{\boldsymbol{c}} = \begin{bmatrix} \beta_{1j}^{wind} & \beta_{1j}^{solar} & \beta_{1j}^{temp} & \beta_{1j}^{prec} & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \beta_{2j}^{wind} & \beta_{2j}^{solar} & \beta_{2j}^{temp} & \beta_{2j}^{prec} & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & \beta_{Nj}^{wind} & \beta_{Nj}^{solar} & \beta_{Nj}^{temp} & \beta_{Nj}^{prec} \end{bmatrix},$$

while $\boldsymbol{B}_j^{\boldsymbol{g}}$ is:

$$\boldsymbol{B}_j^{\boldsymbol{g}} = \begin{bmatrix} \beta_{1j}^{EUA} & \beta_{1j}^{gas} & \beta_{1j}^{coal} & \beta_{1j}^{brent} & \beta_{1j}^{stoxx} \\ \beta_{2j}^{EUA} & \beta_{2j}^{gas} & \beta_{2j}^{coal} & \beta_{2j}^{brent} & \beta_{2j}^{stoxx} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \beta_{Nj}^{EUA} & \beta_{Nj}^{gas} & \beta_{Nj}^{coal} & \beta_{Nj}^{brent} & \beta_{Nj}^{stoxx} \end{bmatrix}.$$

$\boldsymbol{\gamma}_t^{\boldsymbol{c}}$ and $\boldsymbol{x}_t^{\boldsymbol{g}}$ are the vectors containing the values for the country-specific exogenous (for all N countries) and global exogenous variables, respectively, at time t. In my case $\boldsymbol{\gamma}_t^{\boldsymbol{c}}$ is an $(N \times 4) \times 1$ vector while $\boldsymbol{x}_t^{\boldsymbol{g}}$ is a 5×1 vector already defined in Subsection 1.1.1:

$$\boldsymbol{\gamma}_t^{\boldsymbol{c}} = \begin{bmatrix} \boldsymbol{x}_{1t}^{\boldsymbol{c}} \\ \boldsymbol{x}_{2t}^{\boldsymbol{c}} \\ \vdots \\ \boldsymbol{x}_{Nt}^{\boldsymbol{c}} \end{bmatrix} = \begin{bmatrix} wind_{1t} \\ solar_{1t} \\ temp_{1t} \\ prec_{1t} \\ wind_{2t} \\ solar_{2t} \\ temp_{2t} \\ prec_{2t} \\ \vdots \\ wind_{Nt} \\ solar_{Nt} \\ temp_{Nt} \\ prec_{Nt} \end{bmatrix}.$$

$\boldsymbol{\mu}_0$ and $\boldsymbol{\delta d}_t$ constitute the deterministic components of the "intermediate" model described by Equation (0.9), and are constructed such as:

$$\boldsymbol{\mu}_0 = \begin{bmatrix} \mu_{10} \\ \mu_{20} \\ \vdots \\ \mu_{N0} \end{bmatrix}, \qquad \boldsymbol{\delta} = \begin{bmatrix} \delta_{11} & \delta_{12} & \cdots & \delta_{16} \\ \delta_{21} & \delta_{22} & \cdots & \delta_{26} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{N1} & \delta_{N2} & \cdots & \delta_{N6} \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{d}_t = \begin{bmatrix} d_t^1 \\ d_t^2 \\ \vdots \\ d_t^6 \end{bmatrix},$$

$\boldsymbol{\mu}_0$ being an $N \times 1$ vector of constants, $\boldsymbol{\delta}$ being an $N \times 6$ matrix collecting the coefficient estimates associated with the seasonal dummies and $\boldsymbol{d}_t$ being a 6×1 vector of seasonal dummy variables at time t.

The error component $\boldsymbol{\varepsilon}_t$ is an $N \times 1$ vector containing each individual model's error:

$$\boldsymbol{\varepsilon}_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{Nt} \end{bmatrix}.$$

In Equation (0.9), p and k denote, respectively, the maximum autoregressive order and the maximum lag order for the exogenous regressors in the GVARX, where "exogenous" now refers only to each country's own specific variables and the global variables common to all entities (the foreign variables having been "endogenised"). It is worthwhile to highlight that each country-level equation includes—alongside the global ones—only its own country-specific regressors. p and k are obtained as:

$$p = \max_i q_i \tag{0.10}$$

where, as touched upon earlier, accounting for the most general case in which a seasonal AR polynomial is present,

$$q_i = \max\left(p_i + 7s_i, k_i\right); \tag{0.11}$$

and

$$k = \max_i k_i. \tag{0.12}$$

On the same line of reasoning developed when discussing Equation (0.4), the final model must employ a single set of lag orders. To preserve all the information coming from each of the N country models, these lag orders are chosen as the maxima across every individual lag (both for the autoregressive and exogenous parts of the models). Therefore, any country whose original lag orders fall below the global maxima (p, k) simply has its corresponding higher-order coefficients set to zero. This ensures that no country-specific information is lost and that the GVARX construction remains well posed.

From Equation (0.9), because $\boldsymbol{G}_0$ is a non-singular and known matrix (its elements being functions of the weights matrix $\boldsymbol{W}$ and the estimated ARX coefficients), one can premultiply both sides of the equation by its inverse $\boldsymbol{G}_0^{-1}$. This yields to a GVARX(p, k) which can be expressed as:

$$\begin{aligned}
&\boldsymbol{y}_t = \boldsymbol{a}_0 + \boldsymbol{D}\boldsymbol{d}_t + \boldsymbol{F}_1\boldsymbol{y}_{t-1} + \cdots + \boldsymbol{F}_p\boldsymbol{y}_{t-p} \\
&\qquad\qquad + \boldsymbol{H}_0^{\boldsymbol{c}}\boldsymbol{\gamma}_t^{\boldsymbol{c}} + \cdots + \boldsymbol{H}_k^{\boldsymbol{c}}\boldsymbol{\gamma}_{t-k}^{\boldsymbol{c}} + \boldsymbol{H}_0^{\boldsymbol{g}}\boldsymbol{x}_t^{\boldsymbol{g}} + \cdots + \boldsymbol{H}_k^{\boldsymbol{g}}\boldsymbol{x}_{t-k}^{\boldsymbol{g}} + \boldsymbol{u}_t, \\
\text{where} \quad &\boldsymbol{a}_0 = \boldsymbol{G}_0^{-1}\boldsymbol{\mu}_0, \\
&\boldsymbol{D} = \boldsymbol{G}_0^{-1}\boldsymbol{\delta}, \\
&\boldsymbol{F}_j = \boldsymbol{G}_0^{-1}\boldsymbol{G}_j, \quad \text{for } j = 1, \dots, p, \\
&\boldsymbol{H}_j^{\boldsymbol{c}} = \boldsymbol{G}_0^{-1}\boldsymbol{B}_j^{\boldsymbol{c}} \text{ and } \boldsymbol{H}_j^{\boldsymbol{g}} = \boldsymbol{G}_0^{-1}\boldsymbol{B}_j^{\boldsymbol{g}}, \quad \text{for } j = 0, \dots, k, \\
&\boldsymbol{u}_t = \boldsymbol{G}_0^{-1}\boldsymbol{\varepsilon}_t.
\end{aligned} \tag{0.13}$$

The resulting $N \times 1$ vector equation expresses each element of $\boldsymbol{y}_t$ also as a function of other countries' domestic variables since $\boldsymbol{y}_t$ collects the electricity prices of all N regions.

By exploiting the first-stage ARX estimates and the country-specific link matrices $\boldsymbol{W}_i$, it is possible to effectively sidestep the dimensionality challenge that would arise from direct estimation. If one were to estimate a VAR (or VARX) with 23 endogenous series, each autoregressive lag would introduce 23^2 additional parameters to be estimated; every extra lag would significantly reduce the degrees of freedom of the specified model, eventually leading to an overfitting scenario. The fact that the GVARX is not estimated but calculated starting from a set of more parsimonious, individually estimated ARX models effectively obviates this challenge. This methodological advantage is the key feature of the GVAR framework, and it is the main rationale for adopting and developing it throughout this thesis. Equation (0.13) lends itself to a wide range of applications whose potential is discussed in Subsection 1.2.2.

Companion form of the GVARX model

The model in Equation (0.13) can be equivalently expressed in a more compact form. Starting from the obtained GVARX(p, k), define:

$$\boldsymbol{\Psi}_t = \begin{bmatrix} \boldsymbol{y}_t \\ \boldsymbol{y}_{t-1} \\ \vdots \\ \boldsymbol{y}_{t-p+1} \end{bmatrix}, \qquad \boldsymbol{\Gamma}_t = \begin{bmatrix} \boldsymbol{\gamma}_t^{\boldsymbol{c}} \\ \boldsymbol{\gamma}_{t-1}^{\boldsymbol{c}} \\ \vdots \\ \boldsymbol{\gamma}_{t-k}^{\boldsymbol{c}} \end{bmatrix}, \qquad \boldsymbol{X}_t = \begin{bmatrix} \boldsymbol{x}_t^{\boldsymbol{g}} \\ \boldsymbol{x}_{t-1}^{\boldsymbol{g}} \\ \vdots \\ \boldsymbol{x}_{t-k}^{\boldsymbol{g}} \end{bmatrix} \quad \text{and} \quad \boldsymbol{U}_t = \begin{bmatrix} \boldsymbol{u}_t \\ \boldsymbol{0} \\ \vdots \\ \boldsymbol{0} \end{bmatrix},$$

where $\boldsymbol{\Psi}_t$ represents the $(p \times N) \times 1$ vector containing all the endogenous variables considered in the GVARX(p, k) model, encompassing each endogenous vector both contemporaneously and at lags $1, 2, \dots, p-1$. $\boldsymbol{\Gamma}_t$ is the $((k+1) \times (N \times 4)) \times 1$ vector which stacks all the contemporaneous and lagged country-specific exogenous up to k. $\boldsymbol{X}_t$ is the $((k+1) \times 5) \times 1$ vector that does the same as $\boldsymbol{\Gamma}_t$ but for the global exogenous variables. At last, $\boldsymbol{U}_t$ is the $(p \times N) \times 1$ vector containing as the first N elements the error terms within $\boldsymbol{u}_t$ defined in Equation (0.13).

Then one can express (0.13) as:

$$\boldsymbol{\Psi}_t = \boldsymbol{C} + \boldsymbol{S}\boldsymbol{d}_t + \boldsymbol{F}\boldsymbol{\Psi}_{t-1} + \boldsymbol{\Phi}^{\boldsymbol{c}}\boldsymbol{\Gamma}_t + \boldsymbol{\Phi}^{\boldsymbol{g}}\boldsymbol{X}_t + \boldsymbol{U}_t \tag{0.14}$$

corresponding to a GVARX$(1, 0)$ model, but in which the focal-variable vector is represented by $\boldsymbol{\Psi}_t$ and the exogenous variable vectors are $\boldsymbol{\Gamma}_t$ and $\boldsymbol{X}_t$. In this specification:

$$\boldsymbol{F} = \begin{bmatrix} \boldsymbol{F}_1 & \boldsymbol{F}_2 & \cdots & \boldsymbol{F}_{p-1} & \boldsymbol{F}_p \\ \boldsymbol{I}_N & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_N & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{I}_N & \boldsymbol{0} \end{bmatrix},$$

is the $(p \times N) \times (p \times N)$ companion coefficient matrix,

$$\boldsymbol{\Phi}^{\boldsymbol{c}} = \begin{bmatrix} \boldsymbol{H}_0^{\boldsymbol{c}} & \boldsymbol{H}_1^{\boldsymbol{c}} & \cdots & \boldsymbol{H}_k^{\boldsymbol{c}} \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Phi}^{\boldsymbol{g}} = \begin{bmatrix} \boldsymbol{H}_0^{\boldsymbol{g}} & \boldsymbol{H}_1^{\boldsymbol{g}} & \cdots & \boldsymbol{H}_k^{\boldsymbol{g}} \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} \end{bmatrix},$$

are the $(p \times N) \times ((k+1) \times (N \times 4))$ and $(p \times N) \times ((k+1) \times 5)$ matrices containing the coefficients associated with the country-specific and global exogenous estimates, respectively. Lastly,

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{a}_0 \\ \boldsymbol{0} \\ \vdots \\ \boldsymbol{0} \end{bmatrix} \quad \text{and} \quad \boldsymbol{S} = \begin{bmatrix} \boldsymbol{D} \\ \boldsymbol{0} \\ \vdots \\ \boldsymbol{0} \end{bmatrix}$$

represent the $(p \times N) \times 1$ vector including the constant terms of the model, and the $(p \times N) \times 6$ matrix accounting for the seasonal coefficients.

1.2. Model validation and application

1.2.1. Diagnostic assessment

Stability of the model

The companion form of the obtained GVARX model in (0.14) is advantageous for several reasons. For instance, by reducing the original GVARX(p, k) to a GVARX$(1, 0)$ representation, the stability (i.e., stationarity) assessment of the system is facilitated.

Under standard VAR theory, the stationarity condition requires that all roots of the autoregressive polynomial lie outside the unit circle; that is, they all have to be larger than one in modulus. In the companion-matrix formulation, the autoregressive polynomial is obtained from the matrix polynomial $\boldsymbol{I}_{pN} - \boldsymbol{F}L$. Solving the determinantal equation

$$\det\left(\boldsymbol{I}_{pN} - \boldsymbol{F}L\right) = 0, \tag{0.15}$$

one can find the roots of the AR polynomial associated with the estimated GVARX. Equivalently, condition ((0.15) can be written as:

$$\det\left(\boldsymbol{I}_{pN}\lambda_{eig} - \boldsymbol{F}\right) = 0, \tag{0.16}$$

where λ_{eig} denote the eigenvalues of $\boldsymbol{F}$ (and hence of the GVARX). Solving Equation ((0.16) yields to the complex solutions $\lambda_{eig} = a \pm bi$ whose corresponding moduli are calculated as $\left|\lambda_{eig}\right| = \sqrt{a^2 + b^2}$. If, to assess the system's stability, the eigenvalues are computed, the stationarity requirement is that all their moduli should be less than one, as this ensures that all the solutions of Equation ((0.15) are outside the unit circle.[7]

When some of the eigenvalue moduli approach unity, the GVARX—while strictly speaking still stable—operates close to the threshold of instability, yielding potentially uninformative results. This was how I was able to conclude that excluding outliers

[7] To see this, consider that the determinant in Equation ((0.15) can be written as $\det\left(L\left(\boldsymbol{I}_{pN}\frac{1}{L}\right) - \boldsymbol{F}L\right) = \det\left(L(\boldsymbol{I}_{pN}\frac{1}{L} - \boldsymbol{F})\right) = L^{pN}\det\left(\boldsymbol{I}_{pN}\frac{1}{L} - \boldsymbol{F}\right)$. And since $L = 0$ cannot be a solution of Equation ((0.15) as $\det(\boldsymbol{I}_{pN}) = 1$, by defining $\lambda_{eig} = \frac{1}{L}$ one can easily verify that having all the solutions of $\det\left(\boldsymbol{I}_{pN} - \boldsymbol{F}L\right) = 0$ greater than one in modulus is equivalent to having all the solutions of $\det\left(\boldsymbol{I}_{pN}\lambda_{eig} - \boldsymbol{F}\right) = 0$ lower than one in modulus.

from the electricity price return series yielded a "more stable" GVARX specification, as outlined in Subsection 1.1.1: the system's eigenvalues fell deeper inside the unit circle. An unstable GVARX is unsuitable for empirical analysis because it fails to reflect real-world dynamics. For instance, response functions derived from this system may diverge over the forecast horizon instead of converging to a steady state, a behaviour which is usually inconsistent with the real observed phenomena. A GVARX specification that properly reflects reality must therefore exhibit stability, ensuring that its dynamic behaviour remains bounded and aligned with expectations.

Instability can nonetheless arise even if each ARX model is stable (i.e., all their AR polynomial roots lie outside the unit circle). This happens because when constructing the GVARX, additional cross-country linkages via the $\boldsymbol{W}_i$ link matrices are introduced. When these interdependencies are incorporated into the companion matrix $\boldsymbol{F}$, they can alter its spectral properties and potentially push some eigenvalues outside the unit circle. In other words, although each ARX may be well-behaved on its own, aggregating them into a global system can amplify shocks at each step rather than dampen them, leading to overall instability.

In my specific case, the eigenvalue decomposition of the companion coefficient matrix yields for the largest eigenvalue a modulus of 0.882, confirming the overall system's stability. Naturally, each of the constituent ARX models is also stable.

The companion matrix representation, however, proves valuable for more than just conducting stability diagnostics. A second advantage lies, for instance, in its facilitation of response function computation: further details on this aspect are provided in Subsection 1.2.3.

Residual autocorrelations

Beyond assessing stability, a key diagnostic is to examine the residuals of the obtained model—an analysis foreshadowed in Subsection 1.1.1 during the discussion of seasonal components and appropriate lag selection in the ARX models. Since these diagnostic plots are rather unwieldy at this stage, they are presented in the Appendix.

Section X displays the empirical autocorrelation functions of the GVARX residual series $\boldsymbol{u}_t$ for each country, with the corresponding 95% confidence bands for the null hypothesis of zero autocorrelation. Although the GVARX is constructed from a set of individual ARX models whose residuals might exhibit, to a lesser extent, the presence of autocorrelation, the aggregation step can slightly "transform" these residuals, causing some ACF plots to reveal more pronounced autocorrelation levels. Conversely, in certain instances, the autocorrelations may actually diminish as well. I used a maximum lag of 30—chosen as a multiple of seven, plus a couple of additional lags—to ensure the detection of any remaining seasonal patterns that may not have been fully captured.

Some degree of autocorrelation seems to be present; not for all countries, but in a few of them—most notably for Czechia, France, Norway, Poland, Romania (at a surprisingly short lag), and Sweden. For the remaining countries, although some empirical autocorrelations fall outside the 95% confidence bands at certain lags, these deviations are minor. As noted in Footnote 1, given the multiple hypothesis tests conducted across numerous lags, I expect that a number of autocorrelation coefficients will fall outside the 95% confidence bounds due to type I errors. With 30 lags tested for each country, observing some values outside the confidence bands is anticipated. What I seek to avoid is instead a systematic autocorrelation pattern.

Among the countries whose ACF bands exceed more evidently the confidence limits, these exceedances occur in isolation rather than clustering into a recognisable structure beyond the bounds. Moreover, across the full set of countries, there is no systematic behaviour that would, for instance, indicate misspecified seasonality. No common residual autocorrelation pattern emerges across nations either, suggesting that any individual lag outside the confidence bands does not reflect a shared misspecification for the European landscape.

The isolated autocorrelation spike at lag 1 for Romania is somewhat unexpected. While this deviation is singular—since the rest of the ACF plot exhibits no notable irregularities—it is surprising given that the individual ARX specification already includes autoregressive dynamics of order greater than one. There also remain entities (e.g. Austria, Hungary, the Netherlands, and Slovakia) that exhibit no significant autocorrelation up to lag 30.

Although all these residual patterns seem neither to be systematic nor clustered in a manner that would imply clear misspecification, a clearer diagnostic assessment could be found after applying some refinements discussed earlier, which—for reasons highlighted in that same chapter—have not been included within the present thesis.

Residual distributions

As noted in Subsection 1.1.1, the residuals from the individual country models often exhibit heavier tails compared to a Gaussian distribution and this feature is likewise reflected in the distributions of the GVARX residual vector $\boldsymbol{u}_t$. A formal normality test here, is not performed as the departure from normality is already quite evident from a qualitative assessment of the plots.

For a preliminary discussion on how I handled this characteristic, I refer to Subsection 1.1.1, while for a more specific description of the methodology employed to assess the uncertainty around the model's estimates, see Subsection 1.2.4.

Residual scatterplots

In addition to the residual diagnostic measures already discussed, I also examined the scatterplot of GVARX residuals against their fitted values. Within this assessment, a

random dispersion of points around zero—with no discernible pattern—serves as a robust "sanity check" that the GVARX structure is functioning correctly. Conversely, the emergence of patterns indicates that some component has been omitted from the model. This assessment is particularly useful for a multitude of controls. For instance, it allows to evaluate the suitability of a linear specification (which, I remind, is applied to log-returns, since most variables are modelled as log differences[8]). If the true relationships are approximately linear, residuals should cluster around the horizontal zero line without exhibiting curvature or waves. A systematic profile—such as a U-shaped plot—would suggest the need to include nonlinear terms or interactions.

Furthermore, this plot facilitates the detection of potential heteroscedasticity. In an ideal scenario, the "cloud" of residuals maintains a roughly constant vertical dispersion across all fitted-value levels. However, if the points fan out or contract as the fitted values increase—forming a funnel-like shape—this indicates that the residual variance depends on the predicted level (i.e., there would be evidence for a heteroscedastic pattern). Notice that, by expressing the target variable as log differences (effectively percentage changes in electricity prices with respect to the underlying price level), I ensure that the errors are likewise measured in relative, rather than absolute, terms. Without the logarithmic transformation, the model would operate on level differences, and the residuals would reflect absolute deviations (observed minus predicted differences). In many cases, absolute errors grow with the level of the focal variable; by modelling percentage changes instead, I normalise the residuals to the underlying price level, which often helps in mitigating heteroscedasticity. This consideration is one more reason in favour of taking logarithms before differencing: in essence, when price fluctuations are larger at higher price levels, log-transforming the series can greatly stabilise the variance of the residuals other than the variance of the series itself.

The scatterplots reveal residual clouds that are well-dispersed around zero for all countries, with no systematic patterns—a positive indication of the GVARX's validity. Moreover, the vertical ranges of these clouds differ across countries, mirroring the relative volatility levels already observed in the summary statistics of the target variable return series. For example, Finland, Sweden, and Denmark—each exhibiting higher empirical return volatility—also display wider (yet uniform across the fitted values) residual dispersions, while Italy and Greece, which evidenced lower return volatility, show consistently tighter residual clouds.

[8] Only the climatic control variables enter as raw differences.

1.2.2. Potential applications

Although the econometric framework developed here is suited—and specifically tailored—to the thesis's research question, its applicability extends beyond this immediate scope.

One natural application of the developed specification is, of course, interpretative: one can examine how shocks to exogenous variables (both country-specific and global) propagate through the system and affect each country's electricity price, all while accounting for the interconnected and interdependent nature of the markets. Building on this, the model enables computation of impulse-response functions for domestic shocks as well, quantifying the temporal profile of how a disturbance in one country's electricity price impacts every other country in the GVARX. This approach allows for an assessment of cross-country spillovers, measuring the effect of an economy's shock on its neighbours while preserving the geographical interconnection dynamics.

Importantly, these analyses are inherently dynamic rather than static. In fact, in contexts such as electricity markets, adjustments unfold over time: an initial shock may produce an immediate effect, but further variations often materialise as the system responds and reverts toward equilibrium. The GVARX framework captures this evolving behaviour, making it a tool for understanding both the contemporaneous and dynamic responses of electricity prices to diverse shocks.

In addition, although forecasting lies beyond this thesis's focus, the developed approach is readily adaptable to prediction exercises, with certain considerations. Because the model also includes contemporaneous exogenous variables, forecasting, for instance, one step ahead (i.e., at $T+1$) requires the exogenous values at time $T+1$ which may not be available in the information set up to time T: $\mathcal{I}_T$. Therefore, one has to identify a suitable value for the contemporaneous exogenous regressors in order to produce forecasts. In some cases, these exogenous inputs may be policy levers or controllable interventions, allowing to specify their future paths directly. More commonly, however, also considering the "typical" context of a GVARX application, they lie outside direct control, necessitating auxiliary forecasting models for $E[\boldsymbol{X}_{T+1} \mid \mathcal{I}_T]$, where $\boldsymbol{X}_t$ here denotes the set of the exogenous variables at time t. Such auxiliary forecasts might come from bespoke time-series models, institutional forecasts, or scenario analyses (e.g. optimistic, neutral, and pessimistic paths for key exogenous drivers). One must then feed these predicted values into the GVARX. A caveat of this two-step approach is that it complicates the propagation of uncertainty and thus the construction of confidence intervals.

Within this thesis's specific application, however, all exogenous variables are time-aligned to reflect the information set available at the day-ahead price formation stage. Therefore, when forecasting the price for delivery on day $T+1$, the exogenous inputs for $T+1$ are populated with the data observed up to calendar day T and as such can be considered available for one-step ahead forecasts. I did this under the assumption

that no new data arriving at calendar time $T + 1$ can influence the already announced price coming from the previous day's auction.

The GVARX framework is well-suited to a variety of decision-making and strategic applications which are of interest for various stakeholders: energy regulators and public authorities at both national and European levels, TSOs, infrastructure investors, energy traders or more generally market participants. Some examples of potential uses include policy design and impact assessment as the model enables to simulate cross-border policy measures—such as higher carbon taxes or revised gas-market regulations—and quantifies their effects on electricity prices throughout the interconnected system. This could, for instance, be a way to evaluate the effectiveness of the emission trading scheme in shaping the outcomes of the power markets.

Infrastructure planning and investment appraisal represent one other area of interest. The developed specification allows for assessing how changes in transmission capacity or new renewable generation projects influence price dynamics across neighbouring countries, thereby informing optimal siting and development of cross-border interconnectors or national generation assets.

The opportunity of quantifying a country's exposure to both domestic and global shocks—whether meteorological events or fuel-supply disruptions—and tracing how these disturbances propagate through the network of linked markets helps in identifying critical dependencies and potential contagion pathways. This directly translates into a resilience and shock propagation analysis.

At last, it can be seen as a strategic hedging and trading tool as well. For market participants, the model provides insights into temporal and spatial shock transmission, allowing traders to gauge inter-market dependencies and design hedging strategies that reflect the dynamic spillovers of price shocks across regions.

1.2.3. Specific empirical application

To address the research question of this thesis, attention has been paid to interpreting the impact of the global EUA price on wholesale electricity prices across all 23 countries included in the GVARX. Leveraging the model's structure, this translated into the construction and analysis of impulse response functions for the exogenous EUA return series (EUA_t) on each country's electricity-price return y_{it}.

When measuring the impact of an EU-allowance price shock on electricity prices, the fact that the model operates on log-returns must be considered. In particular, it is both relevant and methodologically appropriate to assess the dynamic impact of a sustained level-increase in EUA prices on the electricity level-price across countries.

To this end, two elements stand out for their importance: first, an impulse in the first differences of a variable (i.e., a one-time shock of magnitude η at $t = 0$, preceded and followed by values which correspond to a constant equilibrium level at 0 for all $t \neq 0$)

is mathematically equivalent to a step function of size η in the non-differenced series. In other words, applying a discrete jump of η to the non-differenced variable at $t = 0$ (after a constant pre-shock baseline for $t < 0$) and maintaining that elevated level for all $t > 0$, produces the same dynamic profile as an impulse of magnitude η in the differences. Second, because my model is specified in log-returns, each impulse response coefficient at horizon h can be interpreted as an elasticity. Specifically, the response at h measures the percentage change in the electricity price P_h relative to P_{h-1} induced by a $100\eta\%$ step-shock in the EUA level-price at $t = 0$, with respect to the previous EUA level-price at $t = -1$.

Accordingly, I based the analysis on impulse-response functions computed in returns, thus measuring the effect on electricity-price returns of an impulse to the EUA-return series. Equivalently, one can see this procedure as assessing the target's response to a step change in the logarithms of the original EUA prices. However, because the target is specified in log-returns, interpreting raw impulse response values can be cumbersome. To provide more intuitive results, I instead computed cumulative impulse response functions. This transforms the output from the period-to-period percentage change in electricity prices:

$$\ln(P_h) - \ln(P_{h-1})$$

to the total percentage change from the pre-shock equilibrium level:

$$\ln(P_h) - \ln(P_{-1})$$

with an impulse that, as said, is assumed to manifest itself at $t = 0$ and for $h = 0, 1, 2, \dots, H$ where H represents the extent of the horizon on which the response analysis is carried on. In practical terms, the cumulative response measures the overall impact of a permanent, percentage-level increase in the EUA price on electricity prices—relative to the equilibrium prevailing immediately before the shock—over the chosen forecast horizon.

This approach allows me to interpret the results as follows: "given a $100\eta\%$ sustained increase in the EUA price, what is the resulting percentage increase in electricity prices across the various countries relative to their pre-shock levels?" And: "how does that percentage response evolve over time?"

Technical derivation of the cumulative impulse response function

The following discussion presents a technical exposition of the computations used to derive the impulse response results described above. As noted in Subsection 1.2.1, expressing the GVARX in its companion form (see Equation ((0.14)) also simplifies the calculation of response functions.

Specifically, abstracting from deterministic components and the stochastic part of the model $\boldsymbol{U}_t$ (which will still be considered in the results uncertainty assessment

described in Subsection 1.2.4), consider the system's initial state as being at an equilibrium level with all returns set to 0 for all $t < 0$. At $t = 0$, I impose a one-off shock of magnitude η on the EUA log-returns ($EUA_0 = \eta$)[9]—variable included as the first element of the vector $\boldsymbol{x}_t^{\boldsymbol{g}}$. For all subsequent periods $t = 1, 2, ...$, I assume no further shocks and that the EUA return series remains at zero, maintaining the new equilibrium level. To trace the dynamic adjustment of the system in response to this shock, from (0.14) one can verify how

$$\begin{aligned}
&\boldsymbol{\Psi}_0 = \boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_1,\\
&\boldsymbol{\Psi}_1 = \boldsymbol{F}\boldsymbol{\Psi}_0 + \boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_{1+5} = \boldsymbol{F}\boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_1 + \boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_{1+5},\\
&\boldsymbol{\Psi}_2 = \boldsymbol{F}\boldsymbol{\Psi}_1 + \boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_{1+(5\cdot 2)} = \boldsymbol{F}^2\boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_1 + \boldsymbol{F}\boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_{1+5} + \boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_{1+(5\cdot 2)},\\
&\quad\vdots\\
&\boldsymbol{\Psi}_k = \boldsymbol{F}\boldsymbol{\Psi}_{k-1} + \boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_{1+(5k)} = \boldsymbol{F}^k\boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_1 + \boldsymbol{F}^{k-1}\boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_{1+5} + \cdots +\\
&\qquad\qquad + \boldsymbol{F}\boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_{1+(5\cdot(k-1))} + \boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_{1+(5k)},
\end{aligned} \tag{0.17}$$

where $\boldsymbol{e}_i$ represents the $((k+1)\times 5)\times 1$ (same dimensions of $\boldsymbol{X}_t$) vector containing all zeros except a 1 as the i-th entry. I set this according to the structure of the vector $\boldsymbol{X}_t$ which I recall being constructed as

$$\boldsymbol{X}_t = \begin{bmatrix} \boldsymbol{x}_t^{\boldsymbol{g}} \\ \boldsymbol{x}_{t-1}^{\boldsymbol{g}} \\ \vdots \\ \boldsymbol{x}_{t-k}^{\boldsymbol{g}} \end{bmatrix}, \qquad \text{with} \qquad \boldsymbol{x}_t^{\boldsymbol{g}} = \begin{bmatrix} EUA_t \\ gas_t \\ coal_t \\ brent_t \\ stoxx_t \end{bmatrix}.$$

Continuing the procedure ((0.17) for subsequent steps after the maximum exogenous lag k of the GVARX(p, k), one obtains

$$\begin{aligned}
&\boldsymbol{\Psi}_{k+1} = \boldsymbol{F}\boldsymbol{\Psi}_k = \boldsymbol{F}^{k+1}\boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_1 + \boldsymbol{F}^k\boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_{1+5} + \cdots +\\
&\qquad + \boldsymbol{F}^2\boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_{1+(5\cdot(k-1))} + \boldsymbol{F}\boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_{1+(5k)},\\
&\quad\vdots
\end{aligned} \tag{0.18}$$

[9] η might be for example 0.1, representing a 10% increase in the EUAs price. Notice, however, that if this shock is set as being too big, the percentage change approximation breaks down. This case, however, is dismissed as not interesting since such extreme shocks in the timespan of a single day are unlikely to occur.

$$\boldsymbol{\Psi}_{k+j} = \boldsymbol{F}\boldsymbol{\Psi}_{k+j-1} = \boldsymbol{F}^{k+j}\boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_1 + \boldsymbol{F}^{k+j-1}\boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_{1+5} + \cdots +$$
$$+ \boldsymbol{F}^{j+1}\boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_{1+(5\cdot(k-1))} + \boldsymbol{F}^{j}\boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_{1+(5k)},$$

which can be summarised as

$$\boldsymbol{\Psi}_{k+j} = \sum_{i=0}^{k} \boldsymbol{F}^{k+j-i}\boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_{1+5i}, \tag{0.19}$$

or more simply, for a given $h = 0, 1, \ldots, H$ as

$$\boldsymbol{\Psi}_{h} = \sum_{i=0}^{\min(k,h)} \boldsymbol{F}^{h-i}\boldsymbol{\Phi}^{\boldsymbol{g}}\eta\boldsymbol{e}_{1+5i.} \tag{0.20}$$

Since equation ((0.20) sums to $\min(k, h)$, it holds for any $h \geq 0$, and can thus be used to compute impulse responses at each step h up to the predefined horizon H. Moreover, by substituting the unit vector $\boldsymbol{e}_{1+5i}$ with $\boldsymbol{e}_{2+5i}, \boldsymbol{e}_{3+5i}$, and so on, up until $\boldsymbol{e}_{5+5i}$, one can derive impulse response functions for the other global exogenous variables (following the order of the vector $\boldsymbol{x}_t^{\boldsymbol{g}} = [EUA_t, gas_t, coal_t, brent_t, stoxx_t]'$. The same approach naturally extends to shocks in country-specific exogenous variables.

Recall that $\boldsymbol{\Psi}_t$ is defined as

$$\boldsymbol{\Psi}_t = \begin{bmatrix} \boldsymbol{y}_t \\ \boldsymbol{y}_{t-1} \\ \vdots \\ \boldsymbol{y}_{t-p+1} \end{bmatrix};$$

therefore, by extracting its first N entries, the impulse response values for all N target variables at each step h can be obtained (represented by the vector $\boldsymbol{y}_h$).

Once the values of y_{ih} have been calculated for a given country i using Equation ((0.20), it should be noted that each y_{ih} corresponds to $\ln(P_{ih}) - \ln(P_{i,h-1})$ since y_{ih} has been defined according to Equation X. By calculating the cumulated impulse response function at step h, given by

$$\mathcal{CIRF}_i(h) = \sum_{j=0}^{h} y_{ij} = \ln(P_{i0}) - \ln\left(P_{i,-1}\right) + \ln(P_{i1}) - \ln(P_{i0}) + \cdots +$$
$$+ \ln(P_{ih}) - \ln\left(P_{i,h-1}\right) = \ln(P_{ih}) - \ln\left(P_{i,-1}\right), \tag{0.21}$$

it is noticeable how each value of $\mathcal{CIRF}_i(h)$ for country i represents the percentage difference of the wholesale day-ahead price at h with respect to the pre-shock price level $P_{i,-1}$.

In principle, the cumulative impulse-response function $\mathcal{CIRF}_i(h)$ continues to evolve indefinitely. However, when the analysis uses a sufficiently large horizon H (which usually does not need to be a very high number; often $H = 20$ is enough if the GVARX is not close to instability), one typically observes an initial transient period after the shock, followed by convergence to a steady-state level—that is, a new equilibrium. This convergence occurs provided that the GVARX is stable (i.e., all eigenvalues of the companion matrix $\boldsymbol{F}$ lie inside the unit circle). Consequently, the "long-term" effect of the shock is given by the post-transient equilibrium level of $\mathcal{CIRF}_i(h)$.

1.2.4. Extracting the results uncertainty

Point estimates of the cumulative impulse response functions (CIRFs) alone, while may still be somewhat informative, are neither sufficient nor reliable for practical decision-making if not accompanied by proper measures of uncertainty. Therefore, in addition to computing each $\mathcal{CIRF}_i(h)$ as in Equation (0.21) for $h = 0, 1, \ldots, H$, and for all $i = 1, \ldots, N$, confidence intervals must be constructed around these responses.

As noted in Subsections 1.1.1 and 1.2.1, the lack of a credible distributional assumption for the GVARX residuals $\boldsymbol{u}_t$ undermines the validity of standard-error estimates from OLS—making analytically derived confidence intervals potentially misleading. Under Gaussian innovations—since OLS estimates are expressed as a linear combination of the focal variable, which in turn can be seen as a linear transformation of the errors—the obtained coefficients' estimators would themselves be normally distributed. But this assumption is violated. Although asymptotic theory guarantees the joint normality of VAR coefficient estimates in large samples (under mild assumptions), significant departures from normality in the innovations can materially distort the finite-sample distribution of the estimators. This makes uncertainty estimates' generation problematic: had the Gaussian assumption held, even a less time-consuming and computationally intensive Monte Carlo approach—drawing parameter vectors from the assumed sampling distribution without re-estimating each ARX in every iteration—could have been sufficient to generate uncertainty measures for the GVARX and thereby the CIRFs. However, since this is not the case, I adopted an alternative methodology to empirically approximate the sampling distribution of the CIRFs.

To quantify uncertainty, I implemented a bootstrap-based procedure: this nonparametric method generates confidence intervals directly from the data without relying on distributional assumptions. I bootstrapped the GVARX to obtain the empirical distribution of the CIRFs and, as a consequence, the percentile-based confidence intervals around them. In detail, I adopted the following procedure.

For a given day t, denote the estimated GVARX residuals by $\widehat{\boldsymbol{u}_t}$. These are recentred to obtain $\widetilde{\boldsymbol{u}_t}$ where

$$\widetilde{\boldsymbol{u}_t} = \begin{bmatrix} \widetilde{u_{1t}} \\ \widetilde{u_{2t}} \\ \vdots \\ \widetilde{u_{Nt}} \end{bmatrix} \quad \text{with} \quad \widetilde{u_{it}} = \widehat{u_{it}} - \overline{u_i} \quad \text{and} \quad \overline{u_i} = \frac{1}{T}\sum_{t=1}^{T} \widehat{u_{it}} \quad \forall i = 1, 2, \dots, N. \quad (0.22)$$

T represents the length of the residual time series $\widehat{\boldsymbol{u}_t}$ and this procedure is carried out for all $t = 1, \dots, T$ so that one obtains the series $\{\widetilde{\boldsymbol{u}_1}, \widetilde{\boldsymbol{u}_2}, \dots, \widetilde{\boldsymbol{u}_T}\}$. Recentring the residuals before each bootstrap draw is essential to ensure that the resampled residuals population have zero mean. For each country i, the estimated residuals series $\widehat{u_{it}}$ may exhibit a slight, nonzero average; therefore, sampling them without recentring would propagate this bias into every iteration. By subtracting the sample mean of $\widehat{u_{it}}$ before resampling, I prevent the introduction of any systematic drift in the simulated series.

The series $\{\widetilde{\boldsymbol{u}_1}, \widetilde{\boldsymbol{u}_2}, \dots, \widetilde{\boldsymbol{u}_T}\}$ is then resampled through random draws with replacement to obtain $\{\widetilde{\boldsymbol{u}_1}^b, \widetilde{\boldsymbol{u}_2}^b, \dots, \widetilde{\boldsymbol{u}_T}^b\}$: a series (of the same length as the original GVARX residuals) of bootstrap errors. The intuition behind sampling with replacement is that the observed sample serves as the best available proxy for the true population. By drawing observations with replacement from this sample, I effectively mimic the act of repeatedly sampling from the original population itself. Conversely, if one were to implement random draws without replacement, later samplings would depend on the results of the initial ones. The former procedure ensures that the bootstrap residuals retain the same empirical distribution observable in the original sample.

A bootstrap time series $\boldsymbol{y}_t^b$ is then built from the bootstrap residuals as follows:

$$\begin{aligned} \boldsymbol{y}_t^b = \widehat{\boldsymbol{a}_0} + \widehat{\boldsymbol{D}}\boldsymbol{d}_t + \widehat{\boldsymbol{F}_1}\boldsymbol{y}_{t-1} + \cdots + \widehat{\boldsymbol{F}_p}\boldsymbol{y}_{t-p} \\ + \widehat{\boldsymbol{H}_0^c}\boldsymbol{\gamma}_t^c + \cdots + \widehat{\boldsymbol{H}_k^c}\boldsymbol{\gamma}_{t-k}^c + \widehat{\boldsymbol{H}_0^g}\boldsymbol{x}_t^g + \cdots + \widehat{\boldsymbol{H}_k^g}\boldsymbol{x}_{t-k}^g + \widetilde{\boldsymbol{u}_t}^b, \end{aligned} \quad (0.23)$$

for all $t = 1, \dots, T$. Notice that because lagged regressors are unavailable for the first $\max(p, k)$ observations, the series $\boldsymbol{y}_t^b$ is initialised with the original $\boldsymbol{y}_t$ over those periods. For $t > \max(p, k)$, each value of $\boldsymbol{y}_t^b$ is constructed using the corresponding centred bootstrap residuals.

For each iteration $b = 1, \dots, B$, all the N country-specific ARX models are re-estimated on the bootstrap series $\boldsymbol{y}_t^b$. The corresponding GVARX is calculated based on the procedure highlighted in Subsection 1.1.2 and its $\mathcal{CIRF}_i^b(h)$ is obtained for all $h = 1, \dots, H$ according to Equations ((0.20) and ((0.21). It is important to note that, at each iteration, all foreign variables must be recalculated using the series $\boldsymbol{y}_t^b$ (rather than the original $\boldsymbol{y}_t$) before re-estimating the individual ARX models, since these foreign regressors are functions of the iteratively updated domestic variables.

This method allows me to approximate the distribution of the impulse response estimates; these bootstrap estimates are then ordered from smallest to largest, and the $(1-\alpha)100\%$ confidence interval is constructed by extracting the $\frac{\alpha}{2}$ and $(1-\frac{\alpha}{2})$

quantiles from this distribution. By generating many replications (i.e., provided that B is sufficiently large), I can directly obtain a robust uncertainty measure for each response function without relying on any parametric distributional assumption. A key strength of the bootstrap procedure is its conceptual straightforwardness; this, however, comes at the expense of a computational burden—especially in my case, where each of the b bootstrap iterations requires re-estimating 23 ARX models and rebuilding the associated GVARX. This is why I limited the results of this thesis to a number of iterations $B = 1000$, which should, in any case, be enough to retrieve reliable uncertainty estimates.

Given the potential applications of the obtained GVARX outlined in Subsection 1.2.2, it is clear that, while this bootstrap procedure for uncertainty quantification has been tailored to the specific context of this study, it can be readily extended and remains fully applicable to forecasting cases or any other use case in which the obtained model is suitable for analysis.

Bibliography

[1] G. Castagneto Gissey, "How competitive are EU electricity markets? An assessment of ETS Phase II," *Energy Policy*, vol. 73, pp. 278–297, 2014, doi: https://doi.org/10.1016/j.enpol.2014.06.015.

[2] B. Guo and G. Castagneto Gissey, "Cost pass-through in the British wholesale electricity market," *Energy Economics*, vol. 102, p. 105497, 2021, doi: https://doi.org/10.1016/j.eneco.2021.105497.

[3] A. Bublitz, D. Keles, and W. Fichtner, "An analysis of the decline of electricity spot prices in Europe: Who is to blame?," *Energy Policy*, vol. 107, pp. 323–336, Aug. 2017, doi: https://doi.org/10.1016/j.enpol.2017.04.034.

[4] P.-A. Jouvet and B. Solier, "An overview of CO2 cost pass-through to electricity prices in Europe," *Energy Policy*, vol. 61, pp. 1370–1376, Oct. 2013, doi: https://doi.org/10.1016/j.enpol.2013.05.090.

[5] R. Hyndman *et al.*, "forecast: Forecasting Functions for Time Series and Linear Models," *cran.r-project.org*, 2025. https://cran.r-project.org/web/packages/forecast/index.html

[6] A. Cottrell and R. Lucchetti, "Gretl User's Guide: Gnu Regression, Econometrics and Time-series," 2008. Available: https://gretl.sourceforge.net/

[7] S. Tehrani, J. Juan, and E. Caro, "Electricity Spot Price Modeling and Forecasting in European Markets," *Energies*, vol. 15, no. 16, p. 5980, Aug. 2022, doi: https://doi.org/10.3390/en15165980.