NCCC-134
APPLIED COMMODITY PRICE ANALYSIS, FORECASTING AND MARKET RISK MANAGEMENT

The Effects of Temporal Data Aggregation on Price Transmission Analysis

by

Clemens H. Hoffmann and Stephan von Cramon-Taubadel

Suggested citation format:
Hoffmann, C. H. and S. von Cramon-Taubadel. 2023. "The Effects of Temporal Data Aggregation on Price Transmission Analysis." Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. [http://www.farmdoc.illinois.edu/nccc134].

The effects of temporal data aggregation on price transmission analysis

Clemens Heinrich Hoffmann Stephan von Cramon-Taubadel

Abstract: Agricultural economists often use temporally aggregated data for price transmission analyses due to data availability but the used vector error correction models (VECM) indirectly assume that all price information is included in the analysis. In consequence, internal dependency structures are not captured correctly, and estimated parameter values can be biased. The model with temporally aggregated data requires a moving average dynamic as a remedy because the error terms are autocorrelated. We show how the correct aggregated parameter values can be derived. With Monte Carlo experiments, we compare the true parameter values with estimated ones. For the estimations, we use maximum likelihood method by Johansen and Juselius for VECM and the method by Yap and Reinsel for error correction vector autoregressive moving average (EC-VARMA) models. To demonstrate that temporal aggregation matters in real world data, we perform a price transmission analysis for the French and German hog market. We ourselves aggregate the data and observe a moving average dynamic in the aggregated data sets. Furthermore, we see an improvement in biasness and efficiency for the adjustment parameter with the EC-VARMA models. We conclude that aggregated data need moving average parameters for a correct estimation, but EC-VARMA models are challenging to specify and to estimate, especially in small data sets due to the identification problem.

Keywords: temporal aggregation, VECM, EC-VARMA, autocorrelation, European hog markets

1 Introduction

The key assumption in price transmission analysis is the law of one price, i.e., if a profit can be made by arbitrage there will be market participants who engage in the respective trades. In a perfect world with full information and no time and space dimensions, the price of a homogeneous good is therefore set by the market equilibrium. But in reality, markets are separated by time and space and information is limited for market participants. So, economists relax the assumption, instead of an all-time fulfilled equilibrium a long-run equilibrium is assumed to which prices adjust to.

Such price transmission processes can be modeled in very different ways (Cramon-Taubadel and Goodwin, 2021). At core of many approaches is the vector error correction model (VECM) estimation (Engle and Granger, 1987; Johansen and Juselius, 1990). Price transmission is a time dependent process, a process of information transfer. Market participants act and react based on available price information from their market and all other connected markets. The "natural" frequency of this process is depending on the transaction frequency and the transparency of the markets.

But as the meta-analysis of Mengel and Cramon-Taubadel (2014) indicates, most market integration analyses are performed with monthly data, which is of lower frequency than the underlying markets react to. Therefore, a gap opens between the information that is available for market participants and for us as price transmission analysts. Data frequency matters for price transmission analysis Amikuzuno (2010) finds stronger adjustments to multimarket equilibriums for the Ghanaian tomato market with monthly than with semiweekly price data but he cannot explain this finding.

We want to give some insights on theoretical effects of temporal aggregation and how they affect price transmission analysis, especially why standard VECM estimations might lead to biased results. Hereby, temporal aggregation refers to two aggregation schemes. On the one hand, there is skip sampling, also known as systematically sampling or point in time sampling, and on the other hand, there is average sampling, also known as flat sampling. In multivariate models, there is the additional problem that the price series might be differently aggregated, i.e., they differ in the aggregation scheme or their aggregation periods are not perfectly synchronous. These special multivariate cases we will not cover.

Prices are stock variables, so all findings are applicable for other models with only stock variables. Flow variables, e.g., traded quantities of goods, are differently temporally aggregated. For them all values of one aggregation period are summed up, e.g., the yearly traded quantity is the sum of all monthly traded quantities of the corresponding year. This is similar to average aggregation of stock variables but not completely identical.

Our paper is structured that we give at first an overview on what is known about temporal aggregation from the econometric literature. Than we describe the five steps of a standard procedure to analyze price transmission: testing for stationarity, specification of number of lags, testing for cointegration, estimating the VECM and calculating price transmission measures. From it, we focus on the effects of temporal aggregation on the last two. In the third chapter, we describe a methodic approach on how to derive the aggregated model and the price transmission measures. In the fourth chapter, we give the results of our Monte Carlo experiments, followed by an empirical example with French and German hog prices. Finally, we give our conclusion and some recommendations on temporal aggregation.

1.1 Econometric background

In the econometric literature, the role of data frequency and temporal aggregation in time series models is discussed. Average aggregation causes an autocorrelation pattern for the first differences of a random walk. This effect is induced because the values of the aggregated time series are derived from overlapping intervals of the same series of random errors (Working, 1960), i.e, initial error realizations appear multiple times in consecutive values of the aggregated series. For time series models, this means that a moving average (MA) dynamic is induced. More generally, if the data is generated by a univariate autoregressive (AR) model then the model class will switch to autoregressive moving average (ARMA) models (Amemiya and Wu, 1972). This means data aggregation goes hand in hand with model aggregation and through internal dependencies, the AR polynomial induces a MA dynamic through temporal aggregation. For the different parts of integrated autoregressive moving average (ARIMA) models, the effects of temporal aggregation are outlined by Granger (1988): The AR coefficients loose explanatory power and go to zero, which can be shown through recursive substitution (when a variable depends 0.5 on its predecessor, it also depends 0.25 on its value two periods ago, 0.125 on the one three periods ago, etc.); the MA coefficients stay relevant but their structure can get simplified; and the degree of integration stays the same.

For vector autoregressive moving average (VARMA) models, Lütkepohl (1987) derives a methodology to linear transform and aggregate them. This approach shows that a vector autoregressive (VAR) model becomes a VARMA model when the process is aggregated. Even though, the method includes all kinds of temporal and contemporaneous aggregation schemes, the specification of the aggregated model does not give any insights on cointegration. Marcellino (1999) shows with his way to derive aggregated VARMA processes, that cointegration is stable even though the data is temporally aggregated, i.e. the long run relationship of cointegrated variables does not change. Also, the degree of integration of the variables is stable. We use this approach in our paper.

We find other noticeable work on temporal aggregation (Chambers, 2003; Chambers, 2011; Pons and Sansó, 2005) which is based on a triangular error correction model (Phillips, 1991). Instead of a discrete data, Chambers chooses a continuous approach for the time series. Therefore, he does not rely on the data frequency. He shows that estimations based on flow variables are more efficient than estimation with stock variables or a mixture of both. Of mature concern is the heavy estimation bias when flow or average aggregated data is used, because of its autocorrelation. He applies a modified estimation procedure (Phillips and Hansen, 1990) to cope this cause of bias.

The continuous approach is criticized, especially market reactions tend to be more abrupt, i.e. more discrete (Pons and Sansó, 2005). They also conclude that average aggregation leads to more efficient estimates than skip aggregation. Estimation based on a combination of aggregation schemes showed the worst performance. For them, sample size is a more concerning factor than aggregation schemes.

1.2 A usual price transmission analyses procedure and following research questions for working with temporally aggregated data

Our research focuses on the effects of temporally aggregated data on a usual price transmission analyses procedure. It consists of five steps. At first, the prices are tested whether they are integrated with an Augmented Dickey Fuller (ADF) test. Granger (1988) and Marcellino (1999) point out that the degree of integration does not change with temporal aggregation. Wu et al. (2015) study the effects of skip aggregation on stationarity tests. For several agricultural commodity futures prices, a comparison is drawn between results of their Bayesian Model Averaging unit root test method with ADF tests.

In the second step of the procedure, VAR models with different numbers of lags are estimated. They are compared via model selection criteria. In case of a process with MA dynamic, the true number of lags goes to infinity. Lewis and Reinsel (1985) derived an upper bound for the number of lags, when such a process is approximated with a VAR model. It implies that $p \to \infty$ and $p^3/T \to 0$ when $T \to \infty$. Here, p is the number of lags and T the sample size. Consequently, p follows a function of T with a growth rate smaller than $T^{1/3}$. In case of a sample size of 120 observations, p must be smaller than five, to not overfit the model. This theoretical upper limit is also applicable for VECM models (Saikkonen and Lütkepohl, 1996). Of the three common information criteria, Akaike Information criterion (AIC), Bayesian Information criteria (BIC) and Hannan-Quinn criteria (HQ), AIC usually suggests the highest lag order, which might be higher than the upper bound.

Thirdly, the prices are tested on cointegration. Here, a Johansen test is used that fits to the model specification (i.e. with or without constant or trend in the model). This test is designed for models with finite lag structures. The Johansen test stays asymptotically valid, when the VECM has a MA dynamic (Saikkonen and Lütkepohl, 1996). Yet, the chosen lag order matters for test size and power. Athanasopoulos et al. (2016) developed a non-parametric method to determine the cointegration rank that performes better then the Johansen test for EC-VARMA models.

The fourth step is estimating the VECM with $p-1$ lags. As the usual VECM does not include a MA dynamic, the model is not specified in its actual form if the data is temporally aggregated. The EC-VARMA model has a VECM representation with infinite lags. The effects of the MA dynamic are in this representation attributed to other model parameters. Depending on the structure of the MA matrix polynomial, this leads to stronger and weaker changes of individual AR and adjustment parameter. Marcellino (1999) shows that both temporal aggregation schemes differ in their MA dynamic and their error variance-covariance matrix.

Fifthly and finally, price transmission metrics are calculated to assist interpreting the model. Common ones are the PT measure of Gonzalo and Granger (1995) and the IS measure developed by Hasbrouck (1995). These depend either on parts or the whole VECM model. Both express shares of price transmission attributed to different markets.

From this procedure, we focus on two questions:

1. Does the autocorrelation that is introduced by temporal aggregation bias estimation results of VECMs if MA parameters are omitted?
2. Similarly, are PT-measure and IS measure affected by temporal aggregation?

2 Methods and Data

We want to answer these questions with Monte Carlo experiments. This has the advantage of knowing the non-aggregated data generating process and gives us the possibility to rule out further effects, e.g., of seasonality or heteroscedasticity. We assume that the effects that we find for simple EC-VARMA models, are also apparent in more complex time series models. In general, a EC-VARMA can be expressed by equation 1:

$$\Delta P_t = \alpha\beta^{'} P_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta P_{t-i} + u_t + \sum_{j=1}^{q} M_j u_{t-j} \tag{1}$$

P is the vector of prices and u is the vector of random error terms. The delta indicates that the first difference of a price is used. The alphas are the so-called adjustment parameters, and the betas are the parameters of the long-run equilibriums. The product $\alpha\beta^{'}$ is also referred as Π. This matrix can be split into α and $\beta^{'}$ because of its reduced rank. A constant β_0 can be added to the long-run relationship. Γ_i is to ith matrix of transitory parameters and M_j refers to the jth matrix of MA parameters. For our simulation exercise, we use the estimated model from section 4. It is the following bivariate process:

$$\begin{bmatrix}\Delta P_{1t}\\ \Delta P_{2t}\end{bmatrix} = \begin{bmatrix}-0.11\\ 0.05\end{bmatrix}\begin{bmatrix}P_{1t-1} - 1.01P_{2t-1} - 10.39\end{bmatrix} + \begin{bmatrix}0.44 & -0.13\\ 0.22 & 0.56\end{bmatrix}\begin{bmatrix}\Delta P_{1t-1}\\ \Delta P_{2t-1}\end{bmatrix} + \begin{bmatrix}u_{1t}\\ u_{2t}\end{bmatrix} + \begin{bmatrix}0.26 & 0.15\\ 0.09 & -0.3\end{bmatrix}\begin{bmatrix}u_{1t-1}\\ u_{2t-1}\end{bmatrix} \tag{2}$$

with error means $E\begin{bmatrix}u_{1t}\\ u_{2t}\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}$ and variance-covariance matrix $\Sigma = \begin{bmatrix}5.69 & 0.91\\ 0.91 & 4.85\end{bmatrix}$.

EC-VARMA processes can be approximated with VECMs. Using the same empirical data set, we estimate this VECM:

$$\begin{bmatrix}\Delta P_{1t}\\ \Delta P_{2t}\end{bmatrix} = \begin{bmatrix}-0.08\\ 0.09\end{bmatrix}\begin{bmatrix}P_{1t-1} - 1.03P_{2t-1} - 8.65\end{bmatrix} + \begin{bmatrix}0.72 & 0.06\\ 0.29 & 0.29\end{bmatrix}\begin{bmatrix}\Delta P_{1t-1}\\ \Delta P_{2t-1}\end{bmatrix} + \begin{bmatrix}-0.28 & -0.01\\ 0.00 & 0.16\end{bmatrix}\begin{bmatrix}\Delta P_{1t-2}\\ \Delta P_{2t-2}\end{bmatrix} + \begin{bmatrix}u_{1t}\\ u_{2t}\end{bmatrix} \quad \Sigma = \begin{bmatrix}5.21 & 0.85\\ 0.85 & 4.83\end{bmatrix} \tag{3}$$

For the analysis of temporal aggregation effects, it is better to use the EC-VARMA process, because its temporally aggregated model has less lags. We generate 100 data sets of 48000 “weekly” observations. These are aggregated to a sample size T of 12000 “monthly” data points. For skip aggregation, we use each fourth observation and for average aggregation, we compute the arithmetic means of each consecutive set of four values. Before we use our simulated to investigate our questions, we apply skip and average aggregation theoretically to derive the aggregated models. For all estimations, we use a sample of 12000 values, even though this means we omit the last 36000 values for the non-aggregated samples, to rule out effects from the size of the data set.

2.1 Temporal aggregation of EC-VARMA models

When we want to analyze biasness of coefficient estimates, we need a benchmark. In our case, we need to derive the true coefficients of the aggregated model. Therefore, we utilize the procedure of Marcellino (1999). We present here a summary of this procedure, an application of the steps on an concrete example can be found in the annex. At first, we transfer the EC-VARMA into its level-VAR representation:

$$P_t = \sum_{i=1}^{p} \mathbf{A}_i P_{t-i} + u_t + \sum_{j=1}^{q} M_j u_{t-j} \tag{4}$$

with $\mathbf{A}_1 = \mathbf{I} + \mathbf{\Pi} + \mathbf{\Gamma}_1, \mathbf{A}_i = \mathbf{\Gamma}_i - \ \mathbf{\Gamma}_{i-1},\ \ i = 1,..,p-1$ and $\mathbf{A}_p = -\mathbf{\Gamma}_{p-1}$. Now the model can be expressed with the lag operator L by the following equation:

$$\left(\mathbf{I} - \sum_{i=1}^{p} \mathbf{A}_i L^i\right) P = \mathbf{A}\left(L\right) P = \mathbf{M}\left(L\right) u = \left(\mathbf{I} - \sum_{j=1}^{q} \mathbf{M}_j L^j\right) u \tag{5}$$

Here $\mathbf{A}\left(L\right)$ refers to the AR lag matrix polynomial and $\mathbf{M}\left(L\right)$ to the MA lag matrix polynomial. For models without MA coefficients, $\mathbf{M}\left(L\right)$ is simply an identity matrix. For average aggregation, u is pre-multiplied by a scalar polynomial of degree $l-1$ and coefficients equal to $1/l$. Here, l refers to the length of the aggregation period. Similarly, we can implement a mixture of aggregation schemes. The actual aggregation is then applied by pre-multiplying a specific lag matrix polynomial $\mathbf{B}(L)$:

$$\mathbf{A}^*(L^*)\,P^* = \mathbf{B}(L)\,\mathbf{A}(L)\,P = \mathbf{B}(L)\,\mathbf{M}(L)\,u = \mathbf{M}^*(L^*)\;u^* \tag{6}$$

An asterisk behind a letter means that the part of the model is aggregated. $\mathbf{A}^*(L^*)$ and $\mathbf{A}(L)$ have the same polynomial degree, i.e., the number of AR lags is identical for the aggregated and non-aggregated model. $\mathbf{B}(L)$ just shifts $\mathbf{A}(L)$ so that $\mathbf{A}^*(L^*)$ only depends on commonly known prices, i.e., $\mathbf{A}^*(L)$ has parameters equal to zero with exemptions of the lags for which aggregated data is available. For $l=4$ and $p=2$, $\mathbf{B}(L)$ just shifts $\mathbf{A}(L)$ to the fourth and eighth lag:

$$A^*(L) = B(L)A(L) = I + A_4^* L^4 + A_8^* L^8 \tag{7}$$

Through re-indexing, we move the length of one period from one "week" to one "month":

$$A^*(L^*) = I + A_{1^*}^* L^* + A_{2^*}^* L^{*2} \tag{8}$$

The lag operator is then aggregated by redefinition and omission of all lags with zero coefficient matrices. To derive the coefficients of $\mathbf{B}(L)$, we need two matrices:

$$\underset{k\times k(p+pl)}{A_v = (A_1,\ldots,A_p,0,\ldots,0)} \qquad A_m = \underset{k(p+pl)\times k(pl)}{\begin{pmatrix} -I & A_1 & \cdots & A_p & 0 & \cdots & 0 \\ 0 & -I & A_1 & \cdots & A_p & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -I & A_1 & \cdots & A_p \end{pmatrix}}$$

From these two matrices, we delete each l^{th} column, i.e., if the aggregation period length is two, we will delete the second one, the fourth one, the sixth one and so on. The coefficients matrix B is then determined by:

$$B = (B_1, B_2, \ldots, B_{p*l-p}) = A_{v;-k} A_{m;-k}^{-1}$$

To obtain $\mathbf{M}^*(L^*)\;u^*$, one further step aside the multiplication of $\mathbf{B}(L)$ is necessary. With the variance-covariance matrix and the autocovariance matrices of $\mathbf{B}(L)\,\mathbf{M}(L)\,u$, we set up a non-linear equation system which solutions give $\mathbf{M}^*(L^*)$ and the variance-covariance matrix of u^*. The aggregated level-VARMA representation can now be retransformed into its error correction form. Its parameters can be derived as follows: $\mathbf{\Pi}^* = -(\mathbf{I} - \sum_{\mathbf{i=1}}^{\mathbf{P}} \mathbf{A_i^*})$ and $\mathbf{\Gamma}_i^* = -\sum_{j=i}^{p-1} \mathbf{A_{j+1}^*}$.

2.2 Estimation of models with aggregated data

After deriving the aggregated model, we estimate the VECM with the maximum likelihood method of Johansen and Juselius (1990). We use the number of lags that AIC suggests, taking the upper lag limit for approximation of VARMA processes into account. We compare the means of the estimated parameters with their true values from the benchmark and also give their 5%- and 95%-quantiles and their corresponding standard deviations.

We also give estimates for the correct EC-VARMA models. These model type faces the same identification problem as VARMA models, i.e., not every model specification is estimable. Only with well set parameter restrictions, the estimation algorithm such as the one of Yap and Reinsel (1995) will converge. One simple way is the final moving average (FMA) form used by Kascha and Trenkler (2015). It has the disadvantage that the estimated models are restricted to a one-dimensional MA dynamic. We can bring our predicted models to FMA form but this transformation is not easily reversible so we cannot retrieve the true parameter values.

Alternatives to the FMA form are the reverse echolon form by Poskitt and Lütkepohl (1995) and the scalar component method (SCM) by Athanasopoulos et al. (2016). We use the later because it is more flexible, i.e., the number of AR and MA lags doesn't need to be identical. Therefore the SCM specification often uses less free parameters then reverse echolon form, which increases estimation efficiency. The canonical correlation tests by Tiao and Tsay (1989) allow us to identify scalar components in a data set through testing. Athanasopoulos et al. (2016) show that this testing procedure for stationary VARMA models can be applied to cointegrated data sets without changes. We present the interpretation of the test table in section 4. For our simulated data sets, we can ignore the identification problem due to the huge sample size.

Estimating EC-VARMA is usually done in three steps (Kascha and Trenkler, 2015): At first we estimate a long VAR model with h_T lags. Bartel and Lütkepohl (1998) suggest to choose h_T as the maximum of $log(T)$ and $h(AIC)$, where the later is h_T of the model with the lowest value for the Akaike information criterion.

In second step, we use the the error terms of the long VAR to estimate an EC-VARMA without rank restriction in the Π-matrix. Its estimated parameters and an approximated decomposition of Π into α and β give the start values for step 3.

Thirdly and finally, we use a Newton-algorithm to improve our estimates of our free parameters. Usually convergence is reached after small amount of iterations. The same algorithm is then also used to derive the standard deviations of the short-run parameters. With the estimated parameters, we can then calculate price discovery metrics.

2.3 Price discovery metrics

We look at two price discovery metrics: the PT and the IS measure. Equation 9 gives the definition of the PT measure by Gonzalo and Granger (1995).

$$PT^1 = \frac{\alpha_1}{\alpha_1 - \alpha_2}, \quad PT^2 = 1 - PT^1 = \frac{\alpha_2}{\alpha_2 - \alpha_1} \tag{9}$$

It depends only on the adjustment parameter of the VECM. So it is straight forward to adapt the PT measure to EC-VARMA models.

Another commonly used price discovery metric the IS measure proposed by Hasbrouck (1995). Equation 10 states its definition.

$$IS^1 = \frac{[\Psi C]_1^2}{\Psi \Sigma \Psi'}, \quad IS^2 = \frac{[\Psi C]_2^2}{\Psi \Sigma \Psi'} \tag{10}$$

Here C is the Cholesky factorization of Σ and Ψ the long-run impact matrix of an error term. $[\Psi C]_i$ refers to the element that corresponds with the ith price. ψ itself is computed following equation 11, where $\beta_\perp$ and $\alpha_\perp$ are orthogonal matrices to β and α.

$$\Psi = \beta_\perp \left[\alpha'_\perp \left(I_k - \sum_{i=1}^{p-1} \Gamma_i \right) \beta_\perp \right]^{-1} \alpha'_\perp \tag{11}$$

The IS measure is only defined for error correction models without MA dynamic. Therefore, we adapt its definition for EC-VARMA models. The general VECM equation 12 can be expressed with a lag polynomial $C(L)$ as in equation 13.

$$\Delta P_t = \alpha \beta^{'} P_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta P_{t-i} + u_t \tag{12}$$

$$C(L) y_t = \left((1-L) I_K - \alpha \beta' L - \sum_{i=1}^{p-1} \Gamma_i (1-L) L^i \right) y_t = u_t, \tag{13}$$

Following the Granger Representation Theorem, the inverse of $C(L)$ is

$$C(L)^{-1} = \Psi \sum_{i=1}^{t} L^{(t-i)} + \Psi^* (L) \tag{14}$$

Applying this Inverse to our EC-VARMA model, we derive the vector moving average representation (VMA) of equation 15.

$$y_t = \left(\Psi \sum_{i=1}^{t} L^{(t-i)} + \Psi^* (L) \right) M(L) u_t \tag{15}$$

As in case for the VMA of the VECM without MA dynamic, the VMA of an EC-VARMA can have a constant. For simplicity, we assume that this constant is zero. If the EC-VARMA has two MA lags, its VMA representation is

$$y_t = \left(\Psi \sum_{i=1}^{t} L^{(t-i)} + \Psi^* (L) \right) u_t + \left(\Psi \sum_{i=1}^{t-1} L^{(t-1-i)} + \Psi^* (L) \right) M_1 u_{t-1} + \left(\Psi \sum_{i=1}^{t-2} L^{(t-2-i)} + \Psi^* (L) \right) M_2 u_{t-2} \tag{16}$$

As the IS measure is only interested in the part of the shocks that is permanently incorporated in the prices, we only look at the long-run impact matrix before u_{t-2}:

$$\begin{aligned} &\Psi (I + M_1 + M_2) \sum_{i=1}^{t-2} L^{(t-2-i)} u_{t-2} \\ &\qquad = \left(\beta_\perp \left[\alpha'_\perp \left(I_k - \sum_{i=1}^{p-1} \Gamma_i \right) \beta_\perp \right]^{-1} \alpha'_\perp \right) (I + M_1 + M_2) \sum_{i=1}^{t-2} L^{(t-i)} u_{t-2} \end{aligned} \tag{17}$$

u_t and u_{t-1} will appear again the model, so their long-run impact matrices do not contain their complete permanent effect, which the IS measure demands. In the IS measure calculation, we replace Ψ with $\Psi(I + M_1 + M_2)$.

3 Monte Carlo experiments

Before we give the results for the estimations with aggregated data sets, we recover the initial parameter of the non-aggregated process to have a reference point. In table 1, we give the estimation results for the EC-VARMA model with the true specification and for the approximating VECM. For the 100 data sets, AIC suggests use 58 times four, 32 times five, five times six and another 5 times seven lags for the VECM estimation. We use four lags for all 100 estimations to have comparable results. The true parameter values and the means of the corresponding estimated ones are very close for the EC-VARMA models and still close for the adjustment parameter of the VECMs. The Gammas take most the burden that comes from excluding a MA dynamic. The percentiles and standard deviations (SD) indicate an efficient estimation. The p-values suggest that the results are normal distributed for the short-run parameter.

		VECM					EC-VARMA				
	true value	mean	standard deviation	5%-quantile	95%-quantile	p-value	mean	standard deviation	5%-quantile	95%-quantile	p-value
β_2	1.01	1.01	0.00	1.01	1.01	0.81	1.01	0.00	1.01	1.01	0.80
β_0	10.39	10.36	0.31	9.81	10.80	0.00	10.36	0.32	9.81	10.81	0.00
α_1	-0.11	-0.09	0.00	-0.10	-0.09	0.55	-0.11	0.00	-0.12	-0.11	0.12
α_2	0.05	0.04	0.00	0.04	0.05	0.40	0.05	0.00	0.05	0.05	0.60
$\Gamma_{1,11}$	0.44	0.68	0.01	0.67	0.70	0.72	0.44	0.01	0.42	0.46	0.69
$\Gamma_{1,21}$	0.22	0.31	0.01	0.30	0.33	0.41	0.22	0.01	0.20	0.24	0.77
$\Gamma_{1,12}$	-0.13	0.04	0.01	0.02	0.05	0.61	-0.13	0.01	-0.15	-0.11	0.61
$\Gamma_{1,22}$	0.56	0.85	0.01	0.84	0.87	0.35	0.56	0.01	0.55	0.57	0.78
$\Gamma_{2,11}$	0.00	-0.22	0.01	-0.24	-0.21	0.00	-	-	-	-	-
$\Gamma_{2,21}$	0.00	-0.15	0.01	-0.17	-0.14	0.72	-	-	-	-	-
$\Gamma_{2,12}$	0.00	-0.14	0.01	-0.16	-0.12	0.94	-	-	-	-	-
$\Gamma_{2,22}$	0.00	-0.26	0.01	-0.28	-0.24	0.70	-	-	-	-	-
$\Gamma_{3,11}$	0.00	0.08	0.01	0.06	0.09	0.32	-	-	-	-	-
$\Gamma_{3,21}$	0.00	0.06	0.01	0.05	0.08	0.01	-	-	-	-	-
$\Gamma_{3,12}$	0.00	0.07	0.01	0.05	0.09	0.72	-	-	-	-	-
$\Gamma_{3,22}$	0.00	0.09	0.01	0.07	0.11	0.93	-	-	-	-	-
$\Gamma_{4,11}$	0.00	-0.02	0.01	-0.04	-0.01	0.56	-	-	-	-	-
$\Gamma_{4,21}$	0.00	-0.02	0.01	-0.03	-0.01	0.53	-	-	-	-	-
$\Gamma_{4,12}$	0.00	-0.02	0.01	-0.04	-0.01	0.19	-	-	-	-	-
$\Gamma_{4,22}$	0.00	-0.02	0.01	-0.04	-0.01	0.73	-	-	-	-	-
$M_{1,11}$	0.26	-	-	-	-	-	0.26	0.02	0.24	0.29	0.03
$M_{1,21}$	0.09	-	-	-	-	-	0.09	0.01	0.06	0.11	0.19
$M_{1,12}$	0.15	-	-	-	-	-	0.15	0.01	0.13	0.18	0.93
$M_{1,22}$	0.30	-	-	-	-	-	0.30	0.01	0.28	0.32	0.94

Table 1: Model estimations with non-aggregated data sets, p-values of Jarque-Bera test, source: own simulation and calculation

3.1 Skip aggregation

When we apply Marcellino's framework of temporal aggregation to our error correction process, we derive the following EC-VARMA model (parameter rounded to two decimal places):

$$\begin{bmatrix}\Delta P_{1t}\\ \Delta P_{2t}\end{bmatrix} = \begin{bmatrix}-0.66\\ 0.04\end{bmatrix}\left[P_{1t-1} - 1.01P_{2t-1} - 10.39\right] + \begin{bmatrix}0.03 & -0.06\\ 0.09 & 0.02\end{bmatrix}\begin{bmatrix}\Delta P_{1t-1}\\ \Delta P_{2t-1}\end{bmatrix}$$
$$+ \begin{bmatrix}u_{1t}\\ u_{2t}\end{bmatrix} + \begin{bmatrix}0.28 & -0.09\\ 0.26 & 0.08\end{bmatrix}\begin{bmatrix}u_{1t-1}\\ u_{2t-1}\end{bmatrix} \quad \Sigma = \begin{bmatrix}58.88 & 47.99\\ 47.99 & 108.69\end{bmatrix} \tag{18}$$

Before, we estimate the VECM with the skip aggregated data sets, we use AIC to get a suggestion on the number of lags. It recommends 71 times two, 25 times three and 4 times four lags. Therefore, we choose the specification with two lags for all 100 estimations. Aside the VECM estimations without MA dynamic, we also give the results of the EC-VARMA estimations. Both can be found in table 2. The long-run parameters seem to be unaffected for both models. For the VECM, both adjustment parameters appear to be on average overestimated, especially α_2 when the MA dynamic is not accounted for. Further, the transitory parameters take over effects from the MA parameters when these are omitted.

For the estimation of the EC-VARMA models, we observe the only small deviations from our prediction for the parameters. Even though the large sample size allows us to mostly ignore the identification problem, it still persists and actually specifications with less parameter are indicated by the canonical correlation tests. The percentiles and the standard deviations show a much lower efficiency in comparison with the VECM estimation, this is probably caused also by ignoring the identification problem.

3.2 Average aggregation

After skip aggregation, we now move on to temporal aggregation through averaging. Applying Marcellino's procedure derives the following aggregated process:

$$\begin{bmatrix}\Delta P_{1t}\\ \Delta P_{2t}\end{bmatrix} = \begin{bmatrix}-0.66\\ 0.04\end{bmatrix}\left[P_{1t-1} - 1.01P_{2t-1} - 10.39\right] + \begin{bmatrix}0.03 & -0.06\\ 0.09 & 0.02\end{bmatrix}\begin{bmatrix}\Delta P_{1t-1}\\ \Delta P_{2t-1}\end{bmatrix}$$
$$+ \begin{bmatrix}u_{1t}\\ u_{2t}\end{bmatrix} + \begin{bmatrix}0.50 & -0.11\\ 0.29 & 0.27\end{bmatrix}\begin{bmatrix}u_{1t-1}\\ u_{2t-1}\end{bmatrix} + \begin{bmatrix}0.02 & -0.01\\ 0.02 & 0.00\end{bmatrix}\begin{bmatrix}u_{1t-2}\\ u_{2t-2}\end{bmatrix} \quad \Sigma = \begin{bmatrix}42.04 & 34.92\\ 34.92 & 78.07\end{bmatrix} \tag{19}$$

As he concludes, the difference between the aggregation schemes lies solely in their MA dynamics. Especially, the elements on the main diagonal of M_1 reach greater values compared to skip aggregation. Here Working's effect becomes visible.

When we use AIC to select the number of lags, we register a tendency to use more lags as for the skip aggregated data sets. 32 times three, 52 times four, 13 times five, and even once seven and once eight lags respectively are suggested. Therefore, we use four lags in our estimations. As visible in table 3, α_1 is this time on average underestimated, when the MA dynamic is not accounted for, and $\Gamma_{1,11}$ and $\Gamma_{1,22}$ take over most of effects caused by serial correlation of the error terms. Nevertheless, this can be different for longer aggregation period lengths and a different non-aggregated process. In the EC-VARMA estimation, we exclude the second MA lag, as it very close to zero. We can still reproduce most parameter values through the estimation. The deviations from the true values are stronger due to this exclusion. As for skip aggregation, the standard deviations of the parameter are higher for the EC-VARMA model. Therefore, the estimation can be seen as less efficient. Irrespectively of aggregation scheme or the used estimation procedure, the long-run parameters are correctly and efficiently estimated.

		VECM					EC-VARMA				
	true value	mean	standard deviation	5%-quantile	95%-quantile	p-value	mean	standard deviation	5%-quantile	95%-quantile	p-value
β_2	1.01	1.01	0.00	1.01	1.01	0.00	1.01	0.00	1.01	1.01	0.00
β_0	10.39	10.37	0.16	10.09	10.61	0.05	10.37	0.16	10.09	10.62	0.15
α_1	-0.66	-0.50	0.01	-0.52	-0.48	0.41	-0.65	0.05	-0.72	-0.58	0.98
α_2	0.04	0.16	0.02	0.13	0.18	0.00	0.04	0.07	-0.06	0.14	0.81
$\Gamma_{1,11}$	0.03	0.15	0.01	0.13	0.17	0.66	0.03	0.04	-0.04	0.10	0.46
$\Gamma_{1,21}$	0.09	0.24	0.02	0.21	0.26	0.83	0.10	0.06	0.00	0.20	0.70
$\Gamma_{1,12}$	-0.06	0.01	0.01	-0.01	0.03	0.30	-0.06	0.03	-0.10	-0.01	0.22
$\Gamma_{1,22}$	0.02	0.22	0.01	0.19	0.24	0.06	0.01	0.04	-0.05	0.07	0.43
$\Gamma_{2,11}$	0.00	-0.02	0.01	-0.04	-0.01	0.68	-	-	-	-	-
$\Gamma_{2,21}$	0.00	-0.06	0.01	-0.08	-0.04	0.89	-	-	-	-	-
$\Gamma_{2,12}$	0.00	0.02	0.01	0.01	0.03	0.80	-	-	-	-	-
$\Gamma_{2,22}$	0.00	-0.02	0.01	-0.04	0.00	0.50	-	-	-	-	-
$M_{1,11}$	0.28	-	-	-	-	-	0.27	0.08	0.14	0.41	0.87
$M_{1,21}$	0.26	-	-	-	-	-	0.25	0.12	0.05	0.45	0.77
$M_{1,12}$	-0.09	-	-	-	-	-	-0.09	0.04	-0.15	-0.03	0.71
$M_{1,22}$	0.08	-	-	-	-	-	0.09	0.05	0.01	0.17	0.38

Table 2: Model estimations with skip aggregated data sets, p-values of Jarque-Bera test, source: own simulation and calculation

		VECM					EC-VARMA				
	true value	mean	standard deviation	5%-quantile	95%-quantile	p-value	mean	standard deviation	5%-quantile	95%-quantile	p-value
β_2	1.01	1.01	0.00	1.01	1.01	0.00	1.01	0.00	1.01	1.01	0.00
β_0	10.39	10.38	0.16	10.14	10.60	0.01	10.38	0.16	10.11	10.62	0.02
α_1	-0.66	-0.42	0.01	-0.45	-0.40	0.47	-0.63	0.02	-0.67	-0.60	0.81
α_2	0.04	0.13	0.02	0.10	0.17	0.56	0.05	0.03	0.01	0.11	0.53
$\Gamma_{1,11}$	0.03	0.29	0.01	0.27	0.31	0.07	0.04	0.02	0.00	0.08	0.38
$\Gamma_{1,21}$	0.09	0.29	0.02	0.26	0.32	0.85	0.10	0.03	0.05	0.15	0.74
$\Gamma_{1,12}$	-0.06	0.07	0.01	0.05	0.09	0.50	-0.04	0.02	-0.07	-0.01	0.25
$\Gamma_{1,22}$	0.02	0.39	0.02	0.35	0.42	0.21	0.05	0.02	0.01	0.08	0.61
$\Gamma_{2,11}$	0.00	-0.11	0.01	-0.12	-0.08	0.92	-	-	-	-	-
$\Gamma_{2,21}$	0.00	-0.15	0.01	-0.17	-0.13	0.04	-	-	-	-	-
$\Gamma_{2,12}$	0.00	0.01	0.01	-0.01	0.03	0.31	-	-	-	-	-
$\Gamma_{2,22}$	0.00	-0.12	0.02	-0.14	-0.09	0.72	-	-	-	-	-
$\Gamma_{3,11}$	0.00	0.03	0.01	0.01	0.04	0.88	-	-	-	-	-
$\Gamma_{3,21}$	0.00	0.06	0.01	0.04	0.09	0.65	-	-	-	-	-
$\Gamma_{3,12}$	0.00	-0.01	0.01	-0.03	0.00	0.53	-	-	-	-	-
$\Gamma_{3,22}$	0.00	0.03	0.01	0.00	0.05	0.09	-	-	-	-	-
$\Gamma_{4,11}$	0.00	-0.01	0.01	-0.02	0.01	0.81	-	-	-	-	-
$\Gamma_{4,21}$	0.00	-0.02	0.01	-0.05	0.00	1.00	-	-	-	-	-
$\Gamma_{4,12}$	0.00	0.01	0.01	-0.01	0.03	0.80	-	-	-	-	-
$\Gamma_{4,22}$	0.00	-0.00	0.01	-0.03	0.02	0.68	-	-	-	-	-
$M_{1,11}$	0.50	-	-	-	-	-	0.46	0.04	0.39	0.51	0.56
$M_{1,21}$	0.29	-	-	-	-	-	0.27	0.05	0.18	0.35	0.34
$M_{1,12}$	-0.11	-	-	-	-	-	-0.10	0.02	-0.14	-0.06	0.76
$M_{1,22}$	0.27	-	-	-	-	-	0.25	0.03	0.19	0.31	0.31

Table 3: Model estimations with average aggregated data sets, p-values of Jarque-Bera test, source: own simulation and calculation

3.3 Price discovery measures

The PT measure and the IS measure depend on the estimated parameter values. As presented in the previous subsections, these values change when the data set is temporally aggregated. In table 4, we give the means and standard deviations for the PT measure. In case for the VECM estimations, we observe only a slight shift from PT^2 to PT^1. For the EC-VARMA models this shift is stronger. The standard deviations increased as well, due to less estimation efficiency.

	non-aggregated data sets		skip-aggregated data sets		average-aggregated data sets	
	PT^1	PT^2	PT^1	PT^2	PT^1	PT^2
	VECM					
mean	0.68	0.32	0.76	0.24	0.76	0.24
standard deviation	0.01	0.01	0.02	0.02	0.03	0.03
	EC-VARMA					
mean	0.69	0.31	0.94	0.06	0.92	0.08
standard deviation	0.01	0.01	0.09	0.09	0.04	0.04

Table 4: PT measures for VECMs and EC-VARMA models, source: own calculation

	non-aggregated data sets		skip-aggregated data sets		average-aggregated data sets	
	IS^1	IS^2	IS^1	IS^2	IS^1	IS^2
	VECM					
mean	0.33	0.67	0.52	0.48	0.53	0.47
standard deviation	0.02	0.02	0.02	0.02	0.02	0.02
mean (r. o.)	0.18	0.82	0.03	0.97	0.03	0.97
standard deviation (r. o.)	0.02	0.02	0.01	0.01	0.01	0.01
	EC-VARMA					
mean	0.34	0.66	0.52	0.48	0.53	0.47
standard deviation	0.02	0.02	0.02	0.02	0.02	0.02
mean (r. o.)	0.24	0.76	0.04	0.96	0.07	0.93
standard deviation (r. o.)	0.01	0.01	0.02	0.02	0.01	0.01

Table 5: IS measures for VECMs and EC-VARMA models for both orderings of the variables (r. o. = reverse ordering), source: own calculation

For the IS measure, we give the means and standard deviations in table 5. As this measure depends on the ordering of the price variables, we also added means and standard deviations for reverse ordering (r. o.). We observe the biggest differences between the two models for the non-aggregated data sets. When the data is temporally aggregated, we observe nearly identical shifts for the VECM and the EC-VARMA models. These shifts are also unclear in direction because they happen in opposite directions depending on the ordering of the variables.

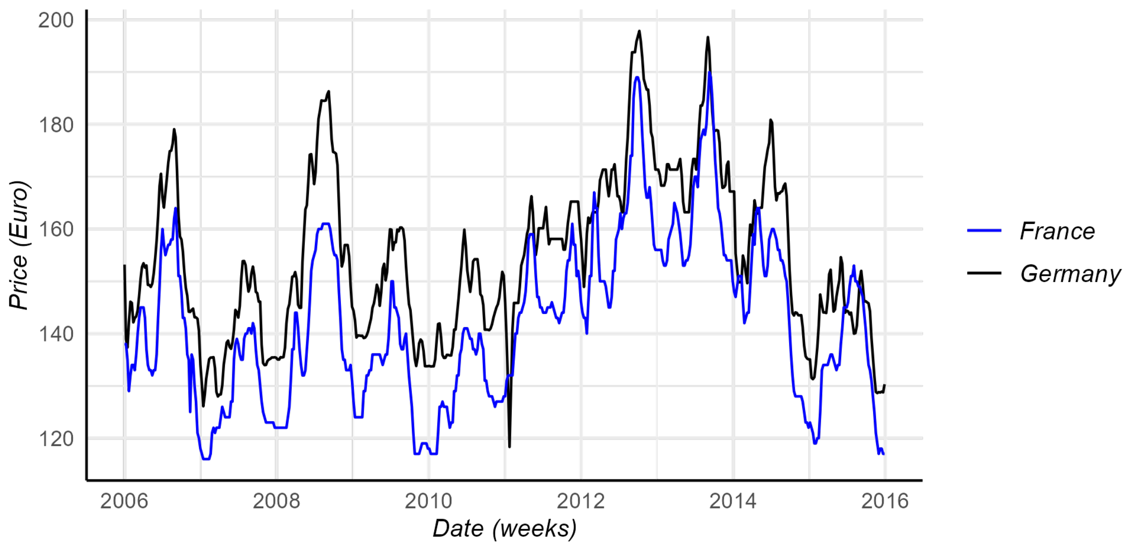

Figure 1: Weekly French and German hog prices (class E), source: own illustration, data from Agra Europe Weekly.

4 Empirical Example: Price transmission between the French and German hog markets

After the Monte-Carlo Simulations, we look at French and German hog prices from 2006 until 2015, whether we can find effects of temporal aggregation. From the EU Access2Markets database, we know that France imported in seven of those ten years no hogs from Germany and at max hogs in value of 250,000 euros in 2006. In the other direction, French exports increased from 16,3 million euros to 31,2 million euros in the first six years of our sample. Afterwards they fell back to 22,1 million in 2015. Both countries are important producers of hogs and are part of European single market. Therefore, they trade a lot with their other neighboring countries. We want to point out, that the German and French price are probably contemporaneous aggregates of regional prices, which can also lead to a MA dynamic. Figure 1 shows the ten years of weekly French and German hog prices. The corresponding sample size is 522.

In the following, price 1 is the German price and price 2 the French one. Using the ADF test on both prices, we conclude that both prices are integrated. For the non-aggregated data set, AIC, BIC and HQ suggest a VECM with 2 lags. The model specification also requires a constant in the long-run price relationship. The Johansen test concludes that the prices are stationary. Athanasopoulos et al. (2016) demonstrated that the Johansen test has a tendency to suggest a higher cointegration rank than the underlying process actually possesses. Their identification method based on canonical correlations concludes that the prices are cointegrated. For the specification of the EC-VARMA model, we use a canonical correlation test table (table 6). Because our system consist of two time series variables, we must find two scalar components. In yellow, we marked the smallest AR and MA order combinations that fulfill this condition. Suggested combinations are a VECM with 2 lags or a EC-VARMA model with one AR and one MA lag. Note that the first AR order from the table is used for the error correction parameter. An EC-VARMA model with five MA lags is also possible, but it would have more free parameters than the other specifications. Therefore, we focus on the other two.

The parameter values of the VECM and the EC-VARMA with non-aggregated prices are given in table 7. The error variance-covariance matrices for the VECM $\mathbf{\Sigma} = \begin{bmatrix} 5.21 & 0.85 \\ 0.85 & 4.83 \end{bmatrix}$ and for the EC-VARMA $\mathbf{\Sigma} =$

		MA order q					
		0	1	2	3	4	5
AR order p	0	0	0	0	0	0	0
	1	0	0	1	1	1	2
	2	0	2	2	3	3	3
	3	2	2	3	2	5	5
	4	2	3	3	4	5	7
	5	1	1	4	5	6	8

Table 6: Scalar components found by canonical correlation tests for the non-aggregated sample

$\begin{bmatrix} 5.69 & 0.91 \\ 0.91 & 4.85 \end{bmatrix}$ are very similar. For the prediction of the temporal aggregation effects, we use the EC-VARMA estimation because it leads to aggregated models with less lags.

	VECM			EC-VARMA		
	estimate	standard deviation	p-value	estimate	standard deviation	p-value
β_2	-1.03	-	-	-1.01	-	-
β_0	-8.65	-	-	-10.39	-	-
α_1	-0.08	0.02	0.00	-0.11	0.02	0.00
α_2	0.09	0.02	0.00	0.05	0.01	0.00
$\Gamma_{1,11}$	0.72	0.04	0.00	0.44	0.08	0.00
$\Gamma_{1,21}$	0.29	0.04	0.00	0.22	0.06	0.00
$\Gamma_{1,12}$	0.06	0.04	0.22	-0.13	0.10	0.16
$\Gamma_{1,22}$	0.29	0.04	0.00	0.56	0.06	0.00
$\Gamma_{2,11}$	-0.28	0.05	0.00	-	-	-
$\Gamma_{2,21}$	0.00	0.04	0.95	-	-	-
$\Gamma_{2,12}$	-0.01	0.04	0.79	-	-	-
$\Gamma_{2,22}$	0.16	0.04	0.00	-	-	-
$M_{1,11}$	-	-	-	0.26	0.09	0.00
$M_{1,21}$	-	-	-	0.09	0.08	0.28
$M_{1,12}$	-	-	-	0.15	0.11	0.17
$M_{1,22}$	-	-	-	-0.30	0.08	0.00

Table 7: Model estimations with non-aggregated data set, p-values of t-test, source: own calculation

Plugging in the estimated process in the framework of Marcellino, we predict the following skip aggregated process:

$$\begin{bmatrix} \Delta P_{1t} \\ \Delta P_{2t} \end{bmatrix} = \begin{bmatrix} -0.66 \\ 0.03 \end{bmatrix} [P_{1t-1} - 1.01 P_{2t-1} - 10.39] + \begin{bmatrix} -0.02 & 0.06 \\ -0.09 & -0.01 \end{bmatrix} \begin{bmatrix} \Delta P_{1t-1} \\ \Delta P_{2t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} + \begin{bmatrix} 0.15 & -0.08 \\ 0.14 & 0.12 \end{bmatrix} \begin{bmatrix} u_{1t-1} \\ u_{2t-1} \end{bmatrix} \mathbf{\Sigma} = \begin{bmatrix} 21.75 & 9.36 \\ 9.36 & 79.77 \end{bmatrix} \quad (20)$$

This aggregated process can only be seen as an approximation because the number of weekly data points per month varies between four and five. We use the aggregation period length of four, because it is the closer to the actual aggregation period length.

For the skip aggregated data set with the first prices of each month, AIC suggests a VECM with two lags. In table 10, we give its respective estimated parameter values. The decrease in sample size leads to a rise of the standard deviations and the longer time periods between price observations lead to which is slightly bigger than predicted. This under prediction is due to the mentioned aggregation period length issue. From the two adjustment parameters only, one stays statistically significant and not the one, which is closer to its predicted value.

We also estimate two EC-VARMA models. For the SCM specification, we look at test table 8. There are three combinations of AR and MA order with minimal lags: a VECM with two lags, an EC-VARMA model with just one MA lag and a model without error correction but with 5 MA lags. We choose the second option. Because, the canonical correlation test also finds scalar components with either just one AR or one MA lag, we must set further parameter restrictions. We choose the scalar component with one MA lag as our second scalar component. The model is therefore restricted to 0 for α_2. Additionally, the matrix before the dependent variables A_0 cannot be an identity matrix anymore because the scalar component methodology allows to add multiples of scalar components with more lags to ones with less. $A_{0,21}$ is therefore part of the estimated variables. In table 10, we also give the "normalized" parameter values, i.e., the ones we receive when we multiply the model with the inverse of A_0.

		MA order q					
		0	1	2	3	4	5
AR order p	0	0	1	1	1	1	2
	1	1	2	3	3	3	3
	2	1	2	4	5	5	5
	3	2	4	6	6	7	7
	4	2	4	6	7	9	9
	5	2	4	6	8	9	10

Table 8: Scalar components found by canonical correlation tests for the skip aggregated sample

As third option, we present the aggregated model in FMA form. The prediction of the parameters change, we derive the FMA form by multiplying our predicted model with the adjoint of $M(L)$. When we use the information criterion by Kascha and Trenkler (2015), a model with one AR and one MA lag is suggested.

Comparing the model estimates, we observe similar values for the cointegration parameters. The alphas are reversed in which one is bigger in contrast to the prediction for the VECM. The EC-VARMA SCM model has the biggest share of statistical significant parameters, this is probably due to very limited number of free parameter. All three models have similar variance-covariances with $\mathbf{\Sigma} = \begin{bmatrix} 52.97 & 39.97 \\ 39.97 & 59.64 \end{bmatrix}$ for the VECM, $\mathbf{\Sigma} = \begin{bmatrix} 56.8 & 43.66 \\ 43.66 & 65.38 \end{bmatrix}$ for the EC-VARMA SCM and $\mathbf{\Sigma} = \begin{bmatrix} 57.33 & 42.12 \\ 42.12 & 62.68 \end{bmatrix}$ for the EC-VARMA FMA model.

As seen in the previous chapter, the error correction and AR parts are identical between the aggregation schemes and only the MA dynamic should differ. For average aggregation, we predict the time series process of equation 21.

$$\begin{bmatrix} \Delta P_{1t} \\ \Delta P_{2t} \end{bmatrix} = \begin{bmatrix} -0.66 \\ 0.03 \end{bmatrix} [P_{1t-1} - 1.01 P_{2t-1} - 10.39] + \begin{bmatrix} -0.02 & 0.06 \\ -0.09 & -0.01 \end{bmatrix} \begin{bmatrix} \Delta P_{1t-1} \\ \Delta P_{2t-1} \end{bmatrix}$$
$$+ \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} + \begin{bmatrix} 0.37 & -0.08 \\ 0.19 & 0.32 \end{bmatrix} \begin{bmatrix} u_{1t-1} \\ u_{2t-1} \end{bmatrix} + \begin{bmatrix} 0 & -0.01 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_{1t-2} \\ u_{2t-2} \end{bmatrix} \mathbf{\Sigma} = \begin{bmatrix} 14.56 & 6.52 \\ 6.52 & 56.83 \end{bmatrix} \quad (21)$$

In table 11, we present similarly to skip aggregation our results for the three different estimations. For the VECM and the EC-VARMA FMA model, the respective information criteria suggest the same number of lags. For the EC-VARMA SCM mode, we look at test table 9. This time, the three possible models are the VECM with two lags, an EC-VARMA model with one AR and one MA lag and an EC-VARMA model with two MA lags. Each model has a second scalar component with just the error correction parameter. $A_{0,21}$ is therefore a free parameter, that must be estimated.

		MA order q					
		0	1	2	3	4	5
AR order p	0	0	0	0	1	1	1
	1	1	1	3	3	3	3
	2	1	3	4	5	5	4
	3	2	4	5	6	6	6
	4	2	4	6	6	7	8
	5	2	4	6	7	8	9

Table 9: Scalar components found by canonical correlation tests for the average aggregated sample

For all three estimations, α_1 is bigger than α_2 as we predicted, but we observe for all estimations an relative increase in α_2 in comparison to α_1. β_0 is estimated closest to our prediction by the EC-VARMA SCM model, for the VECM β_0 is estimated nearly double as big. For all estimations, we estimate very few statistically significant parameter. The variance-covariance matrices are with $\mathbf{\Sigma} = \begin{bmatrix} 40.85 & 29.44 \\ 29.44 & 40.09 \end{bmatrix}$ for the VECM, $\mathbf{\Sigma} = \begin{bmatrix} 44.14 & -44.52 \\ -44.52 & 64.9 \end{bmatrix}$ for the EC-VARMA SCM and $\mathbf{\Sigma} = \begin{bmatrix} 44.16 & 30.54 \\ 30.54 & 41.03 \end{bmatrix}$ for the EC-VARMA FMA model. The EC-VARMA SCM strongly deviates from the other estimations, because the covariance of its error terms is negative.

Overall, we find that the long-run parameters such as β_2 are not affected by temporal aggregation. This holds mostly for a constant in the long-run equilibrium as well. We also find a hardly deniable MA dynamic for the aggregated price data sets. The EC-VARMA FMA specification can include the effects of the MA dynamic but in contrast to an EC-VARMA SCM model it cannot give the true aggregated parameter. Because it is difficult to estimate EC-VARMA models with small data sets, it can be helpful to apply different specifications.

		VECM			EC-VARMA SCM				EC-VARMA FMA			
	prediction	estimate	standard deviation	p-value	estimate	standard deviation	p-value	normalized	prediction	estimate	standard deviation	p-value
$A_{0,21}$	-	-	-	-	-0.03	0.07	0.70	-	-	-	-	-
β_2	-1.01	-1.00	-	-	-1.01	-	-	-1.01	-1.01	-1.01	-	-
β_0	-10.41	-12.40	-	-	-10.85	-	-	-10.85	-10.41	-11.05	-	-
α_1	-0.66	-0.23	0.16	0.16	-0.94	0.22	0.00	-0.94	-0.60	-0.64	0.19	0.00
α_2	0.03	0.35	0.17	0.04	-	-	-	0.03	0.28	0.39	0.20	0.05
$\Gamma_{1,11}$	0.02	0.08	0.15	0.59	-	-	-	-	0.05	0.08	0.21	0.71
$\Gamma_{1,21}$	0.09	0.27	0.16	0.08	-	-	-	-	0.21	0.25	0.17	0.15
$\Gamma_{1,12}$	-0.06	0.19	0.15	0.22	-	-	-	-	-0.14	-0.24	0.13	0.08
$\Gamma_{1,22}$	0.01	0.11	0.16	0.48	-	-	-	-	-0.05	-0.27	0.17	0.11
$\Gamma_{2,11}$	-	-0.29	0.13	0.03	-	-	-	-	0.01	-	-	-
$\Gamma_{2,21}$	-	-0.25	0.14	0.07	-	-	-	-	0.02	-	-	-
$\Gamma_{2,12}$	-	0.29	0.12	0.02	-	-	-	-	0.01	-	-	-
$\Gamma_{2,22}$	-	0.07	0.13	0.59	-	-	-	-	0.02	-	-	-
$M_{1,11}$	0.31	-	-	-	0.88	0.25	0.00	0.88	0.22	0.42	0.21	0.05
$M_{1,21}$	0.36	-	-	-	0.69	0.14	0.00	0.67	0.00	-	-	-
$M_{1,12}$	-0.15	-	-	-	-0.60	0.24	0.02	-0.60	0.00	-	-	-
$M_{1,22}$	-0.09	-	-	-	-0.26	0.14	0.06	-0.24	0.22	0.42	0.21	0.05
$M_{2,11}$	-	-	-	-	-	-	-	-	0.03	-	-	-
$M_{2,21}$	-	-	-	-	-	-	-	-	0.00	-	-	-
$M_{2,12}$	-	-	-	-	-	-	-	-	0.00	-	-	-
$M_{2,22}$	-	-	-	-	-	-	-	-	0.03	-	-	-

Table 10: Model estimations with skip aggregated data set, p-values of t-test, source: own calculation

		VECM			EC-VARMA SCM				EC-VARMA FMA			
	prediction	estimate	standard deviation	p-value	estimate	standard deviation	p-value	normalized	prediction	estimate	standard deviation	p-value
$A_{0,21}$	-	-	-	-	1.70	0.06	0.00	-	-	-	-	-
β_2	-1.01	-0.95	-	-	-1.02	-	-	-1.02	-1.01	-1.02	-	-
β_0	-10.41	-19.81	-	-	-9.92	-	-	-9.92	-10.41	-8.92	-	-
α_1	-0.66	-0.30	0.15	0.04	-0.37	0.36	0.31	-0.37	-0.71	-0.55	0.19	0.00
α_2	0.03	0.18	0.15	0.23	0.82	0.13	0.00	0.19	0.32	0.34	0.17	0.06
$\Gamma_{1,11}$	0.02	0.12	0.15	0.42	-	-	-	-	-0.03	-0.25	0.15	0.10
$\Gamma_{1,21}$	0.09	0.35	0.15	0.02	-	-	-	-	0.19	0.14	0.13	0.26
$\Gamma_{1,12}$	-0.06	0.24	0.15	0.11	-	-	-	-	-0.30	0.03	0.14	0.83
$\Gamma_{1,22}$	0.01	0.25	0.15	0.09	-	-	-	-	-0.23	-0.10	0.16	0.51
$\Gamma_{2,11}$	-	-0.11	0.13	0.40	-	-	-	-	0.03	-	-	-
$\Gamma_{2,21}$	-	-0.18	0.13	0.17	-	-	-	-	0.05	-	-	-
$\Gamma_{2,12}$	-	0.09	0.13	0.49	-	-	-	-	-0.00	-	-	-
$\Gamma_{2,22}$	-	-0.13	0.13	0.32	-	-	-	-	0.03	-	-	-
$M_{1,11}$	0.54	-	-	-	0.50	0.13	0.00	0.50	0.65	0.59	0.13	0.00
$M_{1,21}$	0.38	-	-	-	-	-	-	0.85	0.00	-	-	-
$M_{1,12}$	-0.19	-	-	-	0.19	0.27	0.47	0.19	0.00	-	-	-
$M_{1,22}$	0.10	-	-	-	-	-	-	0.33	0.65	0.59	0.13	0.00
$M_{2,11}$	0.02	-	-	-	-0.03	0.35	0.93	-0.03	0.13	-	-	-
$M_{2,21}$	0.03	-	-	-	-	-	-	-0.05	0.00	-	-	-
$M_{2,12}$	-0.01	-	-	-	0.00	0.21	1.00	0.00	0.00	-	-	-
$M_{2,22}$	-0.01	-	-	-	-	-	-	0.00	0.13	-	-	-

Table 11: Model estimations with average aggregated data set, p-values of t-test, source: own calculation

For the VECM estimations and the EC-VARMA SCM models, we calculated the PT and IS measures and present them in table 12. The magnitudes of the individual PT measures change when the data is temporally aggregated, still the price which corrects the most stays the same. Only for the VECM with average-aggregated data, we observe a switch. Additionally the VECM and the EC-VARMA model come initially to different results which price corrects more.

The IS measure has similar results for the non-aggregated data set for both models. When we skip aggregate the data, we observe a stronger shift to the German price for the VECM than for the EC-VARMA model. For average aggregation, we get completely opposite results. Price transmission measures depend on the estimated parameters. These change with temporal aggregation, therefore the price transmission measures will change as well. In the worst case, we find measure values for aggregated data, which stand in opposition to ones of the non-aggregated data.

	non-aggregated data set		skip-aggregated data set		average-aggregated data set	
	PT^1	PT^2	PT^1	PT^2	PT^1	PT^2
VECM	0.47	0.53	0.40	0.60	0.63	0.37
EC-VARMA SCM	0.69	0.31	0.97	0.03	0.78	0.22
	IS^1	IS^2	IS^1	IS^2	IS^1	IS^2
VECM	0.65	0.35	0.90	0.10	0.79	0.21
VECM (r. o.)	0.52	0.48	0.80	0.20	0.93	0.07
EC-VARMA SCM	0.55	0.45	0.73	0.27	0.01	0.99
EC-VARMA SCM (r. o.)	0.68	0.32	0.68	0.32	0.29	0.71

Table 12: PT and IS measures for the VECM and EC-VARMA SCM models (r. o. = reverse ordering), source: own calculation

5 Conclusions

Temporal aggregation creates additional challenges to correctly model price transmission. The gap between the frequency on with market participants act and the frequency scientist analyze their actions leads to internal dependencies in the time series models. In theory, the aggregated data stays cointegrated and long-run price equilibriums are not affected. The short-run parameters just change.

For the time series model, the induced internal dependencies must be modelled with MA lags to receive interpretable parameter values. When we simply estimate VECM with aggregated data, we unconsciously assume, that no information is transmitted between markets between two price observations. Still, switching to EC-VARMA models create a new issue.

The identification problem complicates now the estimation, we must set restrictions before the model can be estimated. If we use the FMA form, the restrictions will modify all relevant parameters, resulting in a correct estimation but its estimates are not easily interpretable. For EC-VARMA in SCM form, we need sufficient sample sizes, to find all scalar components. In our empirical example, ten years of monthly data were only enough to find the approximations of aggregated model. EC-VARMA estimations can give more statistically significant parameters than the ones of VECMs. Larger samples might also mitigate the greater parameter standard deviations, we observed in our Monte-Carlo experiments. The FMA form is relatively simple to specify and to estimate, but only the SCM form allows to get insights on the parameters of the unrestricted model, on parameters that are interpretable.

Nevertheless, it is advisable to use high frequency data, because internal dependency structures can be avoided. High frequency is hereby a relative term, i.e., if most price information is cycled on weekly basis, e.g., by a newspaper, weekly data can be high frequency data.

We only cover the temporal aggregation cases where the data is aggregated synchronously. When aggregation schemes are mixed, or the data is asynchronously aggregated, additional problems may arise. Here, a skewed internal dependency structure might strongly affect estimations of the long-run price equilibrium, even though it should be unaffected in theory.

Price transmission measures, such as the PT and IS measure, directly depend on the estimated model parameters. If they change either through the temporal aggregation of the data or the chosen model to estimate them, they cannot be interpreted without taking these conditions into account.

References

Amemiya, Takeshi and Roland Y. Wu (1972). "The effect of aggregation on prediction in the autoregressive model". In: *Journal of the American Statistical Association* 67(339), pp. 628–632.

Amikuzuno, Joseph (2010). *Spatial price transmission analysis in agricultural markets: Does the data frequency improve our estimation?* Conference paper. Meeting of the African Association of Agricultural Economists*AAEA* Cape Town South Africa.

Athanasopoulos, George, Donald S. Poskitt, Farshid Vahid, and Wenying Yao (2016). "Determination of Long-run and Short-run Dynamics in EC-VARMA Models via Canonical Correlations". In: *Journal of Applied Econometrics* 31(6), pp. 1100–1119.

Bartel, Holger and Helmut Lütkepohl (1998). "Estimating the Kronecker indices of cointegrated echelon-form VARMA models". In: *The Econometrics Journal* 1(1), pp. 76–99.

Chambers, Marcus J. (2003). "The asymptotic efficiency of cointegration estimators under temporal aggregation". In: *Econometric Theory* 19(1), pp. 49–77.

Chambers, Marcus J. (2011). "Cointegration and sampling frequency". In: *The Econometrics Journal* 14(2), pp. 156–185.

Cramon-Taubadel, Stephan von and Barry K. Goodwin (2021). "Price transmission in agricultural markets". In: *Annual Review of Resource Economics* 13, pp. 65–84.

Engle, Robert F. and Clive W. J. Granger (1987). "Co-integration and error correction: representation, estimation, and testing". In: *Econometrica: journal of the Econometric Society*, pp. 251–276.

Gonzalo, Jesus and Clive W. J. Granger (1995). "Estimation of common long-memory components in cointegrated systems". In: *Journal of Business & Economic Statistics* 13(1), pp. 27–35.

Granger, Clive W. J. (1988). *Aggregation of time series variables-a survey.* Discussion Paper / Institute for Empirical Macroeconomics 1. Federal Reserve Bank of Minneapolis.

Hasbrouck, Joel (1995). "One security, many markets: Determining the contributions to price discovery". In: *The journal of Finance* 50(4), pp. 1175–1199.

Johansen, Soren and Katarina Juselius (1990). "Maximum likelihood estimation and inference on cointegration—with appucations to the demand for money". In: *Oxford Bulletin of Economics and statistics* 52(2), pp. 169–210.

Kascha, Christian and Carsten Trenkler (2015). "Simple identification and specification of cointegrated VARMA models". In: *Journal of Applied Econometrics* 30(4), pp. 675–702.

Lewis, Richard and Gregory C. Reinsel (1985). "Prediction of multivariate time series by autoregressive model fitting". In: *Journal of multivariate analysis* 16(3), pp. 393–411.

Lütkepohl, Helmut (1987). *Forecasting aggregated vector ARMA processes.* Vol. 284. Springer Science & Business Media.

Marcellino, Massimiliano (1999). "Some consequences of temporal aggregation in empirical analysis". In: *Journal of Business & Economic Statistics* 17(1), pp. 129–136.

Mengel, Caroline and Stephan von Cramon-Taubadel (2014). *Distance and border effects on price transmission: A meta-analysis.* Discussion paper. GlobalFood Discussion Papers.

Phillips, Peter C. B. (1991). "Error correction and long-run equilibrium in continuous time". In: *Econometrica: Journal of the Econometric Society*, pp. 967–980.

Phillips, Peter C. B. and Bruce E. Hansen (1990). “Statistical inference in instrumental variables regression with I (1) processes”. In: *The Review of Economic Studies* 57(1), pp. 99–125.
Pons, Gabriel and Andreu Sansó (2005). “Estimation of cointegrating vectors with time series measured at different periodicity”. In: *Econometric Theory* 21(4), pp. 735–756.
Poskitt, Donald S. and Helmut Lütkepohl (1995). *Consistent specification of cointegrated autoregressive moving-average systems.* SFB 373 DiscussionPapers.
Saikkonen, Pentti and Helmut Lütkepohl (1996). “Infinite-order cointegrated vector autoregressive processes”. In: *Econometric theory* 12(5), pp. 814–844.
Tiao, George C. and Ruey S. Tsay (1989). “Model specification in multivariate time series”. In: *Journal of the Royal Statistical Society: Series B (Methodological)* 51(2), pp. 157–195.
Working, Holbrook (1960). “Note on the correlation of first differences of averages in a random chain”. In: *Econometrica: Journal of the Econometric Society*, pp. 916–918.
Wu, Xuedong, Jeffrey H. Dorfman, Berna Karali, et al. (2015). *The Impact of Data Frequency On Stationarity Tests Of Commodity Futures Prices.* Conference poster. 2015 AAEA & WAEA Joint Annual Meeting, July 26-28, San Francisco, California.
Yap, Sook Fwe and Gregory C. Reinsel (1995). “Estimation and testing for unit roots in a partially nonstationary vector autoregressive moving average model”. In: *Journal of the American Statistical Association* 90(429), pp. 253–267.

Appendix

Numeric example for temporal aggregation

In this appendix, we give an example for the procedure of Marcellino (1999). We use here same the EC-VARMA model as in the Monte Carlo experiment:

$$\begin{bmatrix}\Delta P_{1t}\\ \Delta P_{2t}\end{bmatrix} = \begin{bmatrix}-0.11\\ 0.05\end{bmatrix}\left[P_{1t-1} - 1.01P_{2t-1} - 10.39\right] + \begin{bmatrix}0.44 & -0.13\\ 0.22 & 0.56\end{bmatrix}\begin{bmatrix}\Delta P_{1t-1}\\ \Delta P_{2t-1}\end{bmatrix} + \begin{bmatrix}u_{1t}\\ u_{2t}\end{bmatrix} + \begin{bmatrix}0.26 & 0.15\\ 0.09 & -0.3\end{bmatrix}\begin{bmatrix}u_{1t-1}\\ u_{2t-1}\end{bmatrix} \tag{1}$$

with error means $E\begin{bmatrix}u_{1t}\\ u_{2t}\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}$ and variance-covariance matrix $\Sigma = \begin{bmatrix}5.69 & 0.91\\ 0.91 & 4.85\end{bmatrix}$.

At first we transfer the EC-VARMA into its level-VARMA representation (we ignore for now the constant):

$$\begin{bmatrix}P_{1t}\\ P_{2t}\end{bmatrix} = \begin{bmatrix}1.33 & -0.02\\ 0.27 & 1.51\end{bmatrix}\begin{bmatrix}P_{1t-1}\\ P_{2t-1}\end{bmatrix} + \begin{bmatrix}-0.44 & 0.13\\ -0.22 & -0.56\end{bmatrix}\begin{bmatrix}P_{1t-2}\\ P_{2t-2}\end{bmatrix} + \begin{bmatrix}u_{1t}\\ u_{2t}\end{bmatrix} + \begin{bmatrix}0.26 & 0.15\\ 0.09 & 0.3\end{bmatrix}\begin{bmatrix}u_{1t-1}\\ u_{2t-1}\end{bmatrix} \tag{2}$$

Now the model can be expressed with the lag operator L by the following equation:

$$\left(\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} - \begin{bmatrix}1.33 & -0.02\\ 0.27 & 1.51\end{bmatrix}L - \begin{bmatrix}-0.44 & 0.13\\ -0.22 & -0.56\end{bmatrix}L^2\right)\begin{bmatrix}P_{1t}\\ P_{2t}\end{bmatrix} = \left(\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} - \begin{bmatrix}-0.26 & -0.15\\ -0.09 & -0.3\end{bmatrix}L\right)\begin{bmatrix}u_{1t}\\ u_{2t}\end{bmatrix} \tag{3}$$

We call the matrices before the lag operators A_1 to A_2 respectively. They constitute together with the identity matrix in the beginning the AR lag matrix polynomial $A(L)$:

$$A(L) = I - A_1L - A_2L^2 \tag{4}$$

To temporally aggregate to the monthly model, we pre-multiply a specific lag matrix polynomial $B(L)$ to both sides of the equation. The resulting aggregated AR lag matrix polynomial $A^*(L)$ only depends on each forth price observation:

$$A^*(L) = B(L)A(L) = I - A_4^*L^4 - A_8^*L^8 \tag{5}$$

Through re-indexing, we move the length of one period from one week to one month:

$$A^*(L^*) = I - A^*_{1*}L^* - A^*_{2*}L^{*2} \tag{6}$$

To derive the coefficients of B(L), we need two matrices:

$$A_v = \underbrace{\left[\begin{array}{cc|cc|ccc} 1.33 & -0.02 & -0.44 & 0.13 & 0 & ... & 0 \\ 0.27 & 1.51 & -0.22 & -0.56 & 0 & ... & 0 \end{array}\right]}_{2\times 2(2*4)} \tag{7}$$

$$A_m = \underbrace{\left[\begin{array}{cc|cc|cc|cc|c|cc} -1 & 0 & 1.33 & -0.02 & -0.44 & 0.13 & 0 & 0 & ... & 0 & 0 \\ 0 & -1 & 0.27 & 1.51 & -0.22 & -0.56 & 0 & 0 & ... & 0 & 0 \\ \hline 0 & 0 & -1 & 0 & 1.33 & -0.02 & -0.44 & 0.13 & \ddots & 0 & 0 \\ 0 & 0 & 0 & -1 & 0.27 & 1.51 & -0.22 & -0.56 & \ddots & 0 & 0 \\ \hline \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \hline 0 & 0 & ... & ... & 0 & 0 & -1 & 0 & \dots & -0.44 & 0.13 \\ 0 & 0 & ... & ... & 0 & 0 & 0 & -1 & \dots & -0.22 & -0.56 \end{array}\right]}_{2(2*4-2)\times 2(2*4)} \tag{8}$$

From 7 and 8, we determine $A_{v,-l}$ and $A_{m,-l}$ by deleting each fourth block column, i.e., we will delete the fourth one, the eighth one and the twelfth one. The coefficients matrix $B = -(B_1, B_2, ..., B_9)$ is then determined by:

$$B = A_{v,-l}A^{-1}_{m,-l} = \left[\begin{array}{cc|cc|cc|cc|cc|cc} -1.33 & 0.02 & -1.32 & -0.08 & -1.2 & -0.27 & -0.7 & 0.09 & -0.32 & 0.14 & -0.1 & 0.08 \\ -0.27 & -1.51 & -0.55 & -1.71 & -0.74 & -1.77 & -0.71 & -0.79 & -0.44 & -0.28 & -0.18 & -0.07 \end{array}\right] \tag{9}$$

These coefficient matrices, we insert in $B(L) = I + B_1L + B_2L^2 + ... + B_6L^6$ (notice the sign change before the B-matrices) and determine the aggregated autoregressive polynomial:

$$A^*(L) = B(L)A(L) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} -0.37 & -0.61 \\ -0.13 & -0.98 \end{bmatrix} L^4 + \begin{bmatrix} 0.03 & -0.06 \\ 0.09 & 0.02 \end{bmatrix} L^8 \tag{10}$$

expressed with the aggregated lag operator:

$$A^*(L^*) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} -0.37 & -0.61 \\ -0.13 & -0.98 \end{bmatrix} L^{*1} + \begin{bmatrix} 0.03 & -0.06 \\ 0.09 & 0.02 \end{bmatrix} L^{*2} \tag{11}$$

We can now transform the pure autoregressive representation to its error correction form:

$$\Pi^* = -\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.37 & 0.61 \\ 0.13 & 0.98 \end{bmatrix} - \begin{bmatrix} -0.03 & 0.06 \\ -0.09 & -0.02 \end{bmatrix}\right) = \begin{bmatrix} -0.66 & 0.66 \\ 0.04 & -0.04 \end{bmatrix} = \begin{bmatrix} -0.66 \\ 0.04 \end{bmatrix} \begin{bmatrix} 1 & -1.01 \end{bmatrix} \tag{12}$$

and

$$\Gamma^*_1 = -\begin{bmatrix} -0.03 & 0.06 \\ -0.09 & -0.02 \end{bmatrix} = \begin{bmatrix} 0.03 & -0.06 \\ 0.09 & 0.02 \end{bmatrix} \tag{13}$$

In the beginning, we ignored the constant. As it is part of the cointegration relationship, it doesn't change. As for the parameter before P_{2t-1}, the alphas absorb all the changes due to temporal aggregation.

So far, we do not differentiate between skip and average aggregation because their Π^* and Γ^*_1 are identical (average aggregation can be seen as the average of all possible skip aggregations). The aggregation schemes differ only in their MA dynamic, which we derive in the following, starting with skip aggregation: Because we premultiplied $B(L)$ to $A(L)P_t$, we also must premultiply it to $M(L)u_t$:

$$B(L)M(L)u_t = \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} + \begin{bmatrix} 1.59 & 0.13 \\ 0.36 & 1.81 \end{bmatrix} \begin{bmatrix} u_{1t-1} \\ u_{2t-1} \end{bmatrix} + ... + \begin{bmatrix} 0.02 & -0.01 \\ 0.05 & 0.05 \end{bmatrix} \begin{bmatrix} u_{1t-7} \\ u_{2t-7} \end{bmatrix} \tag{14}$$

Of equation 14, we derive the variance-covariance matrix Σ_0:

$$\Sigma_0 = B(1)\Sigma B(1)' \tag{15}$$

$$\Sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 5.69 & 0.91 \\ 0.91 & 4.85 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1.59 & 0.13 \\ 0.36 & 1.81 \end{bmatrix} \begin{bmatrix} 5.69 & 0.91 \\ 0.91 & 4.85 \end{bmatrix} \begin{bmatrix} 1.59 & 0.13 \\ 0.36 & 1.81 \end{bmatrix}' + \\ ... + \begin{bmatrix} 0.02 & -0.01 \\ 0.05 & 0.05 \end{bmatrix} \begin{bmatrix} 5.69 & 0.91 \\ 0.91 & 4.85 \end{bmatrix} \begin{bmatrix} 0.02 & -0.01 \\ 0.05 & 0.05 \end{bmatrix}' = \begin{bmatrix} 61.95 & 51.42 \\ 51.42 & 115.59 \end{bmatrix} \tag{16}$$

Additionally, we need the first and second autocovariance matrix Σ_1. For this, we shift the first $B(L)$ in equation 15 by 4 or 8 lags respectively:

$$\Sigma_1 = \begin{bmatrix} 1.59 & 0.13 \\ 0.36 & 1.81 \end{bmatrix} \begin{bmatrix} 5.69 & 0.91 \\ 0.91 & 4.85 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1.67 & 0.27 \\ 0.75 & 2.21 \end{bmatrix} \begin{bmatrix} 5.69 & 0.91 \\ 0.91 & 4.85 \end{bmatrix} \begin{bmatrix} 1.59 & 0.13 \\ 0.36 & 1.81 \end{bmatrix}' + \\ ... + \begin{bmatrix} 0.02 & -0.01 \\ 0.05 & 0.05 \end{bmatrix} \begin{bmatrix} 5.69 & 0.91 \\ 0.91 & 4.85 \end{bmatrix} \begin{bmatrix} 0.17 & -0.07 \\ 0.32 & 0.22 \end{bmatrix}' = \begin{bmatrix} 12.05 & 3.36 \\ 19.43 & 21.68 \end{bmatrix} \tag{17}$$

With Σ_0 and Σ_1, we set up the equation system described by 18 and 19, where Σ^* is the aggregated error variance-covariance matrix and the M_i^* the MA coefficient matrices. M_0^* is set to the identity matrix.

$$\Sigma_0 = \sum_{i=0}^{1} M_i^* \Sigma^* M_i^{*\prime} \tag{18}$$

$$\Sigma_1 = M_1^* \Sigma^* M_0^{*\prime} \tag{19}$$

To solve this equation system, we use the R-package "nleqslv". It gives the following solutions:

$$\Sigma^* = \begin{bmatrix} 58.88 & 47.99 \\ 47.99 & 108.69 \end{bmatrix} M_1^* = \begin{bmatrix} 0.28 & -0.09 \\ 0.26 & 0.08 \end{bmatrix}$$

After deriving the MA-dynamic for skip aggregation, we move on to average aggregation. Its derivation is similar, we just must adapt $B(L)$ by premultiplying additionally the polynomial $1/4 + 1/4L + 1/4L^2 + 1/4L^3$:

$$(1/4 + 1/4L + 1/4L^2 + 1/4L^3)B(L)u_t = \begin{bmatrix} 0.25 & 0.00 \\ 0.00 & 0.25 \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} + \begin{bmatrix} 0.65 & 0.03 \\ 0.09 & 0.7 \end{bmatrix} \begin{bmatrix} u_{1t-1} \\ u_{2t-1} \end{bmatrix} + \\ \begin{bmatrix} 1.06 & 0.1 \\ 0.28 & 1.25 \end{bmatrix} \begin{bmatrix} u_{1t-2} \\ u_{2t-2} \end{bmatrix} + ... + \begin{bmatrix} 0.00 & 0.00 \\ 0.01 & 0.01 \end{bmatrix} \begin{bmatrix} u_{1t-10} \\ u_{2t-10} \end{bmatrix} \tag{20}$$

Its variance-covariance and autocorrelations matrices are:

$$\Sigma_0 = \begin{bmatrix} 49.55 & 42.30 \\ 42.30 & 93.16 \end{bmatrix} \Sigma_1 = \begin{bmatrix} 17.24 & 8.84 \\ 22.23 & 32.11 \end{bmatrix} \Sigma_2 = \begin{bmatrix} 0.42 & -0.02 \\ 0.92 & 0.78 \end{bmatrix}$$

Because we have three autocovariance matrices, we also need three lags in the aggregated MA-dynamic. The respective equation system (21, 22, 23) contains therefore also four equations. M_0^* is set to the identity matrix.

$$\Sigma_0 = \sum_{i=0}^{2} M_i^* \Sigma^* M_i^{*\prime} \tag{21}$$

$$\Sigma_1 = M_1^* \Sigma^* M_0^{*\prime} + M_2^* \Sigma^* M_1^{*\prime} \tag{22}$$

$$\Sigma_2 = M_2^* \Sigma^* M_0^{*\prime} \tag{23}$$

Its solutions are:

$$\Sigma^* = \begin{bmatrix} 42.04 & 34.92 \\ 34.92 & 78.07 \end{bmatrix} M_1^* = \begin{bmatrix} 0.50 & -0.11 \\ 0.29 & 0.27 \end{bmatrix} M_2^* = \begin{bmatrix} 0.02 & -0.01 \\ 0.02 & 0.00 \end{bmatrix}$$

In summary, we derived theoretically for skip aggregation the following model:

$$\begin{bmatrix} \Delta P_{1t} \\ \Delta P_{2t} \end{bmatrix} = \begin{bmatrix} -0.66 \\ 0.04 \end{bmatrix} \left[P_{1t-1} - 1.01 P_{2t-1} - 10.41 \right] + \begin{bmatrix} 0.03 & -0.06 \\ 0.09 & 0.02 \end{bmatrix} \begin{bmatrix} \Delta P_{1t-1} \\ \Delta P_{2t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} + \begin{bmatrix} 0.28 & -0.09 \\ 0.26 & 0.08 \end{bmatrix} \begin{bmatrix} u_{1t-1} \\ u_{2t-1} \end{bmatrix} +, \quad \Sigma^* = \begin{bmatrix} 58.88 & 47.99 \\ 47.99 & 108.69 \end{bmatrix} \tag{24}$$

and for average aggregation this model:

$$\begin{bmatrix} \Delta P_{1t} \\ \Delta P_{2t} \end{bmatrix} = \begin{bmatrix} -0.66 \\ 0.04 \end{bmatrix} \left[P_{1t-1} - 1.01 P_{2t-1} - 10.41 \right] + \begin{bmatrix} 0.03 & -0.06 \\ 0.09 & 0.02 \end{bmatrix} \begin{bmatrix} \Delta P_{1t-1} \\ \Delta P_{2t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} + \begin{bmatrix} 0.50 & -0.11 \\ 0.29 & 0.27 \end{bmatrix} \begin{bmatrix} u_{1t-1} \\ u_{2t-1} \end{bmatrix} + \begin{bmatrix} 0.02 & -0.01 \\ 0.02 & 0.00 \end{bmatrix} \begin{bmatrix} u_{1t-2} \\ u_{2t-2} \end{bmatrix}, \quad \Sigma^* = \begin{bmatrix} 42.04 & 34.92 \\ 34.92 & 78.07 \end{bmatrix} \tag{25}$$