Optimal Hedging with the Vector Autoregressive Model

Lukasz Gatarek[1]
Søren Johansen[2]

[1] *Erasmus School of Economics, Erasmus University Rotterdam, Tinbergen Institute, the Netherlands;*
[2] *University of Copenhagen, CREATES, Denmark.*

Optimal hedging with the vector autoregressive model

Lukasz T. Gatarek[1][2] & Søren Johansen[3]

9 Feb 2014

Abstract

We derive the optimal hedging ratios for a portfolio of assets driven by a Cointegrated Vector Autoregressive model with general cointegration rank. Our hedge is optimal in the sense of minimum variance portfolio.

We consider a model that allows for the hedges to be cointegrated with the hedged asset and among themselves. We find that the minimum variance hedge for assets driven by the CVAR, depends strongly on the portfolio holding period. The hedge is defined as a function of correlation and cointegration parameters. For short holding periods the correlation impact is predominant. For long horizons, the hedge ratio should overweight the cointegration parameters rather then short-run correlation information. In the infinite horizon, the hedge ratios shall be equal to the cointegrating vector. The hedge ratios for any intermediate portfolio holding period should be based on the weighted average of correlation and cointegration parameters.

The results are general and can be applied for any portfolio of assets that can be modeled by the CVAR of any rank and order.

Keywords: hedging, cointegration, minimum variance portfolio

JEL Classification: C22, C58, G11

[2]Econometric Institute and Tinbergen Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands. E-mail: gatarek@tlen.pl

[3]Department of Economics, University of Copenhagen and CREATES, Department of Economics and Business, Aarhus University, DK-8000 Aarhus C. E-mail: soren.johansen@econ.ku.dk.

Introduction

The idea of minimum variance portfolio dates back to [Markowitz, 1952]. It is defined as a portfolio of individually risky assets that, when taken together, result in the lowest possible risk level for the rate of expected return. Such a portfolio hedges each investment with an offsetting investment; the individual investor's choice on how much to offset investments depends on the level of risk and expected return he/she is willing to accept. The investments in a minimum variance portfolio are individually riskier than the portfolio as a whole. The name of the term comes from how it is mathematically expressed in Markowitz Portfolio Theory, in which volatility is used as a replacement for risk, and in which less variation in volatility correlates to less risk in an investment. Since the seminal paper of [Markowitz, 1952], the notion of minimum variance portfolio and minimum variance hedging has been explored and extended heavily in both financial and econometric literature, see [Grinold and Kahn, 1999]. However, the common denominator of those methods remain the same. They either aim at minimizing volatility of a portfolio itself or volatility of some function of a portfolio. This function often represents the evolution of the portfolio over time. This is also the purpose of the hedging problem we define.

In general, the hedging methods can be divided in two classes: static and dynamic methods. The static hedging techniques assume that the hedged portfolio is selected given information available in period t, and remains unchanged during the entire holding period $t, \ldots, t + h$. On the contrary, the dynamic hedging methods allow for rebalancing of the portfolio during the holding period.

Our method is static. We find the optimal hedging ratios for a portfolio of assets driven by a Cointegrated Vector Autoregressive model (CVAR). We start with a simple process, which relates the hedged asset to hedges via cointegration relation. The hedges are exogenous and are modeled via random walks. In a general specification of the model, the hedges can also be related to the hedged asset via cointegration relation.

The general results that we find, define the optimal hedging ratios as a function of correlation and cointegration parameters in the model. We find that a minimum variance portfolio held for one period should be based on the hedge ratios driven only by correlation. In the infinite horizon, the hedge ratios will be equal to a cointegrating vector. The hedge ratios for any intermediate portfolio holding period should be based on the weighted average of the correlation and cointegration parameters. Our result are general and can be applied to a CVAR model of any rank and order.

1 Exogenous case

1.1 Regression estimation of a cointegration relation

Single regression equation estimation of cointegration relation has been a topic of research for more than two decades. There are many methods of correcting the finite-sample biases in static regressions. [Phillips, 1988] and [Phillips, 1990] have argued that the performance of estimators of cointegrating vectors based on individual regressions is adversely affected by the existence of first-order biases. These biases have no effect on the consistency of the estimators, but result in the asymptotic distributions having non-zero means.

Such biases might play an important role in finite samples. Let us consider a simple bivariate cointegration model with an endogenous variable y_{1t} cointegrated with a weakly

exogenous variable y_{2t}

$$y_{1t} = \beta y_{2t} + u_{1t}$$
$$y_{2t} = y_{2,t-1} + u_{2t}$$

(1.1)

where where u_t are independent identically distributed (i.i.d.) $N(0, \Phi)$ random errors with mean zero and variance Φ, where

$$\Phi = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix}.$$

When disturbances u_{it}, $i = 1, 2$ are autocorrelated or intercorrelated, a static regression of y_{1t} on y_{2t} provides an estimate which can be biased in fairly large samples. This is due to disregarding any information about the process generating y_{2t}. Under i.i.d. Gaussian errors, the parameters of the process can be estimated via the full-system maximum likelihood estimation of cointegrated system, or by maximum likelihood of the regression model with parameters (β, Φ). This amounts to a single-equation regression of y_{1t} corrected for y_{2t-1} and y_{2t}. This method is referred to as a single-equation dynamic regression. The long-run estimates obtained from the properly specified dynamic equation are equivalent, asymptotically, to the full system estimates.

For the system in (1.1), the definition of dynamic regression is simple. In what follows we summarize this procedure. First we substitute in (1.1) using the second equation to obtain

$$y_{1t} = \beta y_{2,t-1} + \beta u_{2t} + u_{1t} = \beta y_{2,t-1} + v_{1t},$$

where $v_{1t} = \beta u_{2t} + u_{1t}$, and we define $\beta^{Corr} = \phi_{22}^{-1} \xi_{21} = (\phi_{21} + \beta \phi_{22}) \phi_{22}^{-1}$. For $v_{2t} = u_{2t}$ we then have v_t i.i.d. $N(0, \Xi)$ where

$$\Xi = \begin{pmatrix} \xi_{11} & \xi_{12} \\ \xi_{12} & \xi_{22} \end{pmatrix} = \begin{pmatrix} \phi_{11} + \beta^2 \phi_{22} + 2\beta \phi_{12} & \phi_{12} + \beta \phi_{22} \\ \phi_{21} + \beta \phi_{22} & \phi_{22} \end{pmatrix}.$$

From the multivariate Normal theory, the following relation holds

$$E(v_{1t}|u_{2t}) = \frac{Cov(v_{1t}, u_{2t})}{Var(u_{2t})} u_{2t} = \xi_{21} \phi_{22}^{-1} u_{2t} = \beta^{Corr} u_{2t},$$

such that

$$E(y_{1t}|y_{2t}, y_{2,t-1}) = E(\beta y_{2t} + v_{1t}|u_{2t}) = \beta y_{2t} + \beta^{Corr} u_{2t} = \beta y_{2t} + \beta^{Corr} \Delta y_{2t}$$

results in a regression form of the cointegration model

$$y_{1t} = \beta y_{2,t-1} + \beta^{Corr} \Delta y_{2t} + \eta_{1t}.$$

(1.2)

Here $\beta^{Corr} = \phi_{22}^{-1} \xi_{21}$ collects the information with respect to the correlation between v_{1t} and u_{2t} and corrects the estimate from a single equation by information about the correlation structure among random errors. We will refer to it as a β-correction parameter. By construction $\eta_t = y_{1t} - E(y_{1t}|y_{2t}, y_{2,t-1})$ is an i.i.d. sequence independent of the sequence u_{2t}.

The single equation in (1.2) represents unbiased estimation of the cointegration parameter β. We can prove that the correction works by analyzing covariance of η_t and u_{2t}. By construction of η_t

$$\eta_t = u_{1t} - \phi_{21}\phi_{22}^{-1}u_{2t},$$

we find no covariance between η_t and u_{2t}

$$Cov(\eta_t, u_{2t}) = E((u_{1t} - \phi_{12}\phi_{22}^{-1}u_{2t}), u_{2t}) = \phi_{12} - \phi_{12}\phi_{22}^{-1}\phi_{22} = 0,$$

what implies that running dynamic regression (with y_{2t} and $y_{2,t-1}$) leads to no loss of information from ϕ_{12}. The variance of η_t is given by

$$Var(\eta_t) = E(u_{1t} - \phi_{12}\phi_{22}^{-1}u_{2t})^2 = \phi_{11} - \phi_{12}\phi_{22}^{-1}\phi_{21}.$$

Several features are now evident. By construction η_t is uncorrelated with u_{2t}. The process u_{2t} is serially uncorrelated with past u_{2t} and with past u_{1t}. It follows that η_t and u_{2t} are uncorrelated at all lags. The covariance matrix of $(v_{1t}, u_{2t})'$ is diagonal, and the estimation of a single dynamic equation provides a fully efficient and unbiased estimate of the cointegration parameter β. Traditionally β^{Corr} is disregarded after running the regression in (1.2). Its main and only role is to provide an unbiased estimator of β. The information that is conveyed by this parameter, however, might be useful on its own. In what follows we will present an example of how it can be used in construction of hedged portfolios for assets modeled according to the specification in (1.1).

1.1.1   Hedging under full mean reversion

Model (1.1) can be interpreted as a realistic representation of the data generating process for prices of pairs of assets in many financial markets. In that case the process y_{1t} is endogenous and its realization cointegrates with the path of y_{2t} with cointegrating parameter β. The process y_{2t} is defined as an exogenous random walk. Apart from cointegration between y_{1t} and y_{2t}, the process is also characterized by intercorrelation in the disturbances. Despite widely used volatility models in financial econometric theory, cointegrated systems with independent and identically distributed errors remain a valid representation of various relations in financial market, such as daily measurement of financial instruments with limited liquidity. An illiquid asset can be defined as an asset which cannot readily be sold at its face value, see [Vayanos, 2004]. The price formation process in such markets underlie different principles than the highly volatile intra-day, high frequency trading environments. For illiquid markets, the volatility clustering patterns etc., are not reflected in the price quotations. That concerns to an even higher extent quoted financial contracts, whose maturity is distant in the future, for instance in 2 or 3 years. Such families of future contracts, each with different maturity, but referring to the same underlying asset, can be modeled by common stochastic trends. Apart from that, the amount of correlation among such assets is naturally very high. This is related to the sparsity of observations (due to illiquidity), what implies a small amount of noise at the level of daily measurements. In such systems, hedging of a particular maturity with closer or further maturities, has remained a financial industry standard for decades. The informational content, filtered by price quotations for different maturities, remains unchanged across maturity, only the amount and speed of adjustment of distinct maturities to the common market trend is different. Due to limited volatility in such

markets, such systems can be well represented by the specification in (1.1), with reservation, however, that the strong assumption about the full reversion in this model should be relaxed. The simplistic specification in model (1.1) implies full mean reversion from period to period. This follows from an alternative representation of the model in (1.1) given by

$$\Delta y_{1t} = \alpha y_{1,t-1} + \beta y_{2t} + u_{1t}$$
$$\Delta y_{2t} = u_{2t},$$

(1.3)

where $\alpha = -1$ and $\Delta y_{1t} = y_{1t} - y_{1,t-1}$. This representation of the system in (1.1) is referred to as an error correction form, where α can be interpreted as mean reversion or speed of adjustment. Estimation of the parameter α constitutes an important part of econometric modeling in cointegrated systems. For expository reasons we start with the rather restrictive specification of full reversion, $\alpha = -1$, which will be replaced by a more general model below. We derive first an optimal hedge for a model with full mean reversion.

As shown above, the regression estimation of model (1.1) requires a correlation correction represented by the parameter β^{Corr}. The information contained in this parameter can be informative for a hedging procedure based on this process. Our goal consists of finding a minimum variance portfolio. In what follows, we define this hedge in detail given the data generating process for assets specified by the equations (1.1).

We define the hedging parameter β_h as the amount invested in the hedge y_{2t} in order to hedge asset y_{1t} and thus we define the portfolio in time period t

$$s_t = y_{1t} - \beta_h y_{2t}$$

(1.4)

which consists of assets y_{1t} and y_{2t} weighted with $(1, -\beta_h)'$. The subscript h indicates that the main determinant of the hedge we define is the portfolio holding period denoted by h.

In financial market one can consider a long and a short position in a given security. The long position means the holder of the position owns the security and will profit if the price of the security goes up. The short position is defined as the sale of a borrowed security, with the expectation that the asset will fall in value. Then, the investor must eventually return the borrowed security by buying it back from the market. Because it can be purchased cheaper then at the time of borrowing, the difference in price results in profit for the investor.

In portfolio hedging, traditionally a long position in asset y_{1t} is hedged with short a position in another asset y_{2t} and vice versa. Thus the sign in front of the weight β_h symbolizes the market convention regarding hedging practice.

Over time the portfolio value can change. Let $s_{t+h} - s_t$ denote the change in portfolio market value, induced by the risk drivers we wish to hedge against. Provided a hedge y_{2t} is available, the hedging problem involves computing the vector of portfolio weights $(1, -\beta_h)'$ that is optimal according to a particular method. The choice of method depends on the objective of the portfolio manager. Although those objectives can be various, in most cases they aim at keeping the risk caused by the unhedged part of y_{1t} minimal over time. The market value of the portfolio defines the exposure of the portfolio, which can be interpreted as an unhedged part of y_{1t} in period t. This invokes an alternative interpretation of the portfolio value as a spread

$$s_t = y_{1t} - \beta_h y_{2t}.$$

(1.5)

Further we define a hedging error after h observations as

$$e_t(h) = s_{t+h} - s_t = y_{1,t+h} - y_{1t} - \beta_h(y_{2,t+h} - y_{2t}). \tag{1.6}$$

The hedging problem we consider aims at minimizing the conditional variance of the hedging error for a holding period h given the past information $\mathcal{I}_t = \sigma(y_s, s \leq t\}$.

The optimization problem is defined as

$$\min_{\beta_h} \; Var(e_t(h)|\mathcal{I}_t) = \min_{\beta_h} \; Var(s_{t+h}|\mathcal{I}_t). \tag{1.7}$$

The optimal portfolio is selected in period t and it is held up to period $t + h$. It is a static hedge, as the portfolio is not rebalanced during periods $t, \ldots, t + h$.

We present an optimal hedge which explores both the long term cointegration parameter β, but also the correlation between the random errors in y_{1t} and y_{2t}.

Theorem 1.1 *Let $y_t, t = 1, \ldots, T$, be bivariate and given by*

$$y_{1t} = \beta y_{2t} + u_{1t},$$
$$y_{2t} = y_{2,t-1} + u_{2t},$$

where u_i are i.i.d. $(0, \Phi)$. Then the optimal hedge coefficient for the portfolio $s_t = y_{1t} - \beta_h y_{2t}$ at time horizon h is given by

$$\beta_h^* = \beta + h^{-1}\phi_{22}^{-1}\phi_{21} = h^{-1}(\beta(h-1) + \beta^{Corr}), \tag{1.8}$$

and the minimal variance is

$$Var(y_{1,t+h} - \beta_h^* y_{2,t+h}|\mathcal{I}_t) = \phi_{11} - h^{-1}\phi_{12}\phi_{22}^{-1}\phi_{21}.$$

Proof: See Appendix.

The hedge defined in (1.8) is a weighted average of the correlation correction, β^{Corr}, and the cointegration parameter with weights: $1/h$ and $(h-1)/h$. The resulting formula complies with the stylized facts about the short- and long-term hedging in the sense that for a short period, $h = 1$, we hedge fully based on the correlation

$$\beta_h^\star = \beta^{Corr},$$

whereas for a long period, when $h \to \infty$, we hedge fully based on cointegration

$$\beta_h^\star = \beta.$$

1.2   Multivariate Specification

1.2.1   Regression estimation of cointegration relation

We define the model with n assets, from which $n - 1$ are hedges. Let $y_t = (y_{1t}, y_{2t}')' \in \mathbb{R}^{1+(n-1)}, t = 1, \ldots, T$ be given by

$$y_{1t} = \beta' y_{2t} + u_{1t}$$
$$y_{2t} = y_{2t-1} + u_{2t}, \tag{1.9}$$

6

where u_i are i.i.d. $(0, \Phi)$. The main difference between the univariate case (with one hedge modeled as a random walk y_{2t}) and the multivariate case, is the possibility of correlation between the innovations of the random walks $y_{2t}, y_{3t}, \ldots, y_{nt}$. Two hedges modeled by correlated random walks are substitutes. In the extreme scenario, if two potential hedges are fully correlated, then having one of them to hedge y_{1t} is enough for a portfolio. The optimal hedges that we derive for the portfolio based on assets modeled according to (1.9) takes into account not only the ratio between the cointegration and correlation parameters but also deeply explores the correlation in order to account for optimal amount of hedges in a portfolio. By comparing the variance of the optimal portfolio we can see what we gain by including more hedges.

To derive the composition of the hedging portfolio, we first need to derive the β-correction parameters, representing the presence of correlation in the errors in the multivariate specification. As we allow for $n - 1$ cointegrating parameters β_i, $i \in 2, \ldots, n$, we need $n - 1$ correction parameters. To derive them we follow the same procedure as in the single hedge scenario. We insert the random walks in y_2 into first equation in (1.9). The resulting equation for y_{1t} can be specified as

$$
\begin{aligned}
y_{1t} &= \beta' y_{2t} + \beta' u_{2t} + u_{1t}, \\
&= \beta' y_{2t} + v_{1t}
\end{aligned}
$$

where $v_{1t} = \beta' u_{2t} + u_{1t}$. Similarly to the univariate case, in order to define the correction parameters, we explore the covariance between v_{1t} and u_{2t}. We obtain

$$
Var \begin{pmatrix} v_{1t} \\ u_{2t} \end{pmatrix} = \begin{pmatrix} \beta' \Phi_{22} \beta + \beta' \Phi_{21} + \Phi_{12} \beta & \Phi_{12} + \beta' \Phi_{22} \\ \Phi_{21} + \Phi_{22} \beta & \Phi_{22} \end{pmatrix}
$$

Based on properties of the multivariate Normal distribution we obtain

$$
E(v_{1t} | u_{2t}) = (\beta + \Phi_{22}^{-1} \Phi_{21}) u_{2t}
$$

so

$$
y_{1t} = \beta' y_{2t} + \beta^{Corr'} \Delta y_{2t} + \eta_t,
$$

where

$$
\beta^{Corr} = \beta + \Phi_{22}^{-1} \Phi_{21},
$$

and $\{\eta_t\}$ is i.i.d. and uncorrelated with $\{u_{2t}\}$.

1.2.2 Hedging with correlated assets

The vector containing the $n - 1$ individual hedges is defined as

$$
\beta_h = (\beta_{h,2}, \ldots, \beta_{h,n})'.
$$

Then the portfolio is defined as

$$
s_t = y_{1t} - \beta_h' y_{2t}. \tag{1.10}
$$

We interpret $\beta_{h,i}$, $i \in 2, \ldots, n$ as the amount of asset y_{it}, necessary to hedge y_{1t} in the portfolio s_t. Further, the unhedged part of y_{1t} is again equal to the value (exposure) of the portfolio

$$
s_t = y_{1t} - \beta_h' y_{2t},
$$

and the hedging error over h periods is defined as in (1.6). The optimal hedge that minimizes the variance of the hedging error over h periods follow.

Theorem 1.2 *Let $y_t = (y_{1t}, y_{2t}')' \in \mathbb{R}^{1+(n-1)}, t = 1, \ldots, T$ be given by*

$$y_{1t} = \beta' y_{2t} + u_{1t},$$
$$y_{2t} = y_{2t-1} + u_{2t},$$

where u_i are i.i.d. $(0, \Phi)$. Then the optimal hedge coefficient for the portfolio $s_t = y_{1t} - \beta_h' y_{2t}$ at time horizon h is given by

$$\beta_h^* = \beta + h^{-1} \Phi_{22}^{-1} \Phi_{21} = h^{-1}(\beta(h-1) + \beta^{Corr}),$$

and the minimal variance is

$$Var(y_{1,t+h} - \beta_h^* y_{2,t+h} | \mathcal{I}_t) = \Phi_{11} - h^{-1} \Phi_{12} \Phi_{22}^{-1} \Phi_{21}.$$

Proof: See Appendix.

It is seen that the variance of the optimal portfolio increases with the horizon h, from the conditional variance of u_{1t} given u_{2t}, $\Phi_{11} - \Phi_{12}\Phi_{22}^{-1}\Phi_{21}$ for $h = 1$, to the variance of the cointegrating relation Φ_{11} for $h \to \infty$. It is also seen that including more correlated hedges, the varian of u_{1t} given u_{2t} will decrease, but the limit remains the same.

2 Full cointegration model

The analysis under exogenous hedges is now generalized to the full cointegration model, see [Johansen 1996]. Thus we consider a model that allows the hedges to be cointegrated with y_{1t} and among themselves such that the number of cointegrating relations could be more than one. The assumption of full reversion is dropped and general adjustment coefficients are allowed.

Theorem 2.1 *Let $(y_{1t}, y_{2t}')' \in \mathbb{R}^{1+(n-1)}, t = 1, \ldots, T$ be given by*

$$\Delta y_t = \alpha \gamma' y_{t-1} + u_t,$$

where u_t are i.i.d. $(0, \Xi)$ and α and γ are $r \times n$ matrices, and the eigenvalues of $\rho = I_r + \gamma'\alpha$ are inside the unit circle. Then the optimal hedge for the portfolio $s_t = y_{1t} - \beta_h' y_{2t} = (1, -\beta_h')' y_{t+h}$ at time horizon h is given by

$$\beta_h^* = \Sigma_{h22}^{-1} \Sigma_{h21},$$

where

$$\Sigma_h = Var(y_{t+h} | \mathcal{I}_t) = \begin{pmatrix} \Sigma_{h11} & \Sigma_{h12} \\ \Sigma_{h21} & \Sigma_{h22} \end{pmatrix}$$

is given by

$$\Sigma_h = hC\Xi C' - \alpha(\gamma'\alpha)^{-2}(I_r - \rho^h)\gamma'\Xi C' - C\Xi\gamma(I_r - \rho'^h)(\alpha'\gamma)^{-2}\alpha' \qquad (2.1)$$

$$+ \alpha(\gamma'\alpha)^{-1}[\sum_{i=0}^{h-1} \rho^i \gamma'\Xi\gamma\rho'^i](\alpha'\gamma)^{-1}\alpha',$$

and $C = \gamma_\perp(\alpha_\perp'\gamma_\perp)^{-1}\alpha_\perp'$. The minimal variance is

$$Var(y_{1,t+h} - \beta_h^* y_{2,t+h} | \mathcal{I}_t) = \Sigma_{h11} - \Sigma_{h12}\Sigma_{h22}^{-1}\Sigma_{h21}.$$

Proof: See Appendix.

Note that for $h = 1$, we get $I_r - \rho = -\gamma'\alpha$ and $C + \alpha(\gamma'\alpha)^{-1}\gamma' = I_n$ and therefore

$$\Sigma_1 = C\Xi C' + \alpha(\gamma'\alpha)^{-1}\gamma'\Xi C' + C\Xi\gamma(\alpha'\gamma)^{-1}\alpha' + \alpha(\gamma'\alpha)^{-1}\gamma'\Xi\gamma(\alpha'\gamma)^{-1}\alpha' = \Xi,$$

and for $h \to \infty$ we find $h^{-1}\Sigma_h = C\Xi C'$, the long run-variance of the process y_t.

2.1 Some special cases

The results in Theorem 2.1 cover the cases considered so far, in particular we focus on a model with one cointegrating relation and $n - 1$ exogenous hedges.

We consider the model (1.3) written in error correction form

$$\Delta y_{1t} = \alpha_1(y_{1,t-1} + \alpha_1^{-1}\beta'y_{2t-1}) + \beta'u_{2t} + u_{1t}$$
$$\Delta y_{2t} = u_{2t}$$

where u_t are i.i.d.$(0, \Phi)$. We find the covariance matrix

$$\Xi = Var\left(\begin{array}{c} \beta'u_{2t} + u_{1t} \\ u_{2t} \end{array}\right) = \left(\begin{array}{cc} \Phi_{11} + \beta'\Phi_{21} + \Phi_{12}\beta + \beta'\Phi_{22}\beta & \beta'\Phi_{22} + \Phi_{12} \\ \Phi_{21} + \Phi_{22}\beta & \Phi_{22} \end{array}\right),$$

and the parameters $\alpha = (\alpha_1, 0, \ldots, 0)'$, $\gamma = (1, \alpha_1^{-1}\beta')'$, $\gamma'_\perp = (-\alpha_1^{-1}\beta, I_{n-1})$ and $\alpha'_\perp = (0, I_{n-1})$, but also

$$C = I_n - \alpha(\gamma'\alpha)^{-1}\gamma' = \left(\begin{array}{cc} 0 & -\alpha_1^{-1}\beta' \\ 0 & I_{n-1} \end{array}\right).$$

Inserting into the general expression (2.1) we first find that because $\alpha'_\perp = (0, I_{n-1})$ we get

$$\Sigma_{h22} = h\alpha'_\perp \Xi \alpha_\perp = h\Phi_{22}.$$

Next we see that $\alpha'_\perp C = \alpha'_\perp = (0, I_{n-1})$, $1 + \rho = 1 + \gamma'\alpha = 1 + \alpha_1$, $C'e_1 = (0, -\alpha_1^{-1}\beta')$ and we find

$$\Sigma_{h21} = \alpha'_\perp \Sigma_h e_i = h\alpha'_\perp \Xi(0, -\alpha_1^{-1}\beta') - \alpha'_\perp \Xi\gamma(1 - (1 + \alpha_1)^h)\alpha_1^{-1}$$
$$= -h\Phi_{22}\beta\alpha_1^{-1} - (\Phi_{21} + \Phi_{22}\beta + \alpha_1^{-1}\Phi_{22}\beta)(I_r - (1 + \alpha_1)^h)\alpha_1^{-1}$$

and therefore

$$\beta_h^* = -\beta\alpha_1^{-1} + h^{-1}(\Phi_{22}^{-1}\Phi_{21} + \beta(1 + \alpha_1^{-1}))((1 + \alpha_1)^h - 1)\alpha_1^{-1}. \tag{2.2}$$

In particular we can take $\alpha_1 = -1$, see (1.1), and find the full mean reversion case which gives

$$\beta_h^* = \beta + h^{-1}\Phi_{22}^{-1}\Phi_{21},$$

see Theorem 1.2.

3 Numerical illustrations

3.1 Full mean reversion

To present how the developed methodology works in practice, we present the optimal hedge implied by the resulting formulas for a selection of model specifications. It is shown how the optimal hedging formula balances the correlation and cointegration parameters depending on the protfolio holding period h. As shown in formula (1.8), for infinitely held portfolios, the optimal hedge is given by the cointegration parameter. This parameter results in a stationary portfolio s_t with a constant variance. Any other selection of the hedging parameter leads to a portfolio which is nonstationary in variance. That is also true about the optimal hedge we derive for any holding horizon h different from $h = \infty$. Still, we intend to construct the portfolio according to the optimal ratio β_h^* implied by the formula, because for a given holding period h, it results in lower variance than the portfolio based on cointegration hedge and held h time periods. It is crucial for the methodology to know *a priori* the holding horizon for this portfolio.

To start with, we show the role of strength of correlation for the construction of the optimal hedge. We start with case of no correlation in random errors of process in (1.1) and then we increase the correlation subsequently

- No covariance of the innovations
  We select the following parameters for model (1.1)

$$\beta = 0.4$$
$$\Phi = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.2 \end{pmatrix}. \tag{3.1}$$

These parameters imply the following β^{Corr} parameter

$$\beta^{Corr} = \xi_{21}\phi_{22}^{-1} = (\phi_{21} + \beta\phi_{22})\phi_{22}^{-1} = \beta = 0.4, \tag{3.2}$$

which, according to (1.8), defines the optimal hedge ratio for a portfolio s_t held up to period $t + 1$, that is, for $h = 1$. In this degenerate case, the optimal hedge for $h = 1$, i.e. β^{Corr}, is equal to β. As a consequence

$$\beta_h^\star = \frac{\beta^{Corr} + \beta(h - 1)}{h} = \beta. \tag{3.3}$$

To understand why the relation $\beta^{Corr} = \beta$ holds under no covariance in innovations we need to refer to the regression representation of cointegration

$$y_{1t} = \beta y_{2t-1} + \beta^{Corr}\Delta y_{2t} + \eta_t \tag{3.4}$$

and the definition of process η_t

$$\eta_t = u_{1t} - \phi_{21}\phi_{22}^{-1}u_{2t}|_{\phi_{21}=0} = u_{1t}. \tag{3.5}$$

After substituting for η_t in (3.4) we obtain

$$y_{1t} = \beta y_{2t-1} + \beta^{Corr}y_{2t} - \beta^{Corr}y_{2,t-1} + u_{1t}. \tag{3.6}$$

It follows immediately that only if $\beta^{Corr} = \beta$ we obtain the original equation for y_{1t} in model (1.1), i.e.

$$y_{1t} = \beta y_{2t} + u_{1t}. \tag{3.7}$$

- Low covariance of the innovations
  In this scenario we keep the cointegration relation at the same level as in the previous example, but we allow for some correlation in the innovations. The following parameters are selected for model (1.1)

$$\beta = 0.4$$
$$\Phi = \begin{pmatrix} 0.1 & 0.01 \\ 0.01 & 0.2 \end{pmatrix}. \tag{3.8}$$

These parameters imply the following β^{Corr}

$$\beta^{Corr} = \xi_{21}\phi_{22}^{-1} = (\phi_{21} + \beta\phi_{22})\phi_{22}^{-1} = (0.01 + 0.4 \times 0.2)0.2^{-1} = 0.45. \tag{3.9}$$

Thus the optimal hedge for $h = 1$ equals to $\beta^{Corr} = 0.45$ and then it smoothly drops to asymptotically reach the level of the cointegration hedge $\beta = 0.4$. Any intermediate portfolio holding horizon h requires a weighted average of those two parameters, according to equation (1.8).

- High covariance in the innovations
  Finally we consider scenario of high correlation, implied by the covariance of 0.12. The following parameters are selected for model (1.1)

$$\beta = 0.4$$
$$\Phi = \begin{pmatrix} 0.1 & 0.12 \\ 0.12 & 0.2 \end{pmatrix}. \tag{3.10}$$

In that case an optimal hedge required for $h = 1$ rises as much as to 1, higher then in the previous scenario, where $\beta^{Corr} = 0.45$. Thus, the higher correlation automatically translates into a higher hedge for the optimal portfolio for $h = 1$

$$\beta^{Corr} = \xi_{21}\phi_{22}^{-1} = (\phi_{21} + \beta\phi_{22})\phi_{22}^{-1} = (0.12 + 0.4 \times 0.2)0.2^{-1} = 1. \tag{3.11}$$

Due to unchanged cointegration parameter, the optimal hedge for an infinitely held portfolio remains the same, i.e. $\beta = 0.4$.

3.2 Limited reversion

- No covariance in the innovations
  Similarly to case of full mean reversion, we start with the scenario of no correlation. We consider the model in (1.3). Mean reversion is set equal to $\alpha = -0.2$. Cointegration level and covariance matrix are unchanged from the previous example

$$\alpha = -0.2$$
$$\beta = 0.4$$
$$\Phi = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.2 \end{pmatrix} \tag{3.12}$$

11

which results in the same level of β^{Corr}, as well. The limiting level of the hedge for $h \to \infty$ is immediately found from equation (2.2)

$$\beta_h(h \to \infty) = -\frac{\beta}{\alpha} = -\frac{0.4}{-0.2} = 2. \tag{3.13}$$

- High covariance in the innovations
  For the high correlation scenario under limited mean reversion, we consider the following parameters

$$\begin{aligned} \alpha &= -0.2 \\ \beta &= 0.4 \\ \Phi &= \begin{pmatrix} 0.1 & 0.12 \\ 0.12 & 0.2 \end{pmatrix}. \end{aligned} \tag{3.14}$$

Results of this scenario are presented in Figure 1.

Finally, in Figure 2, we present the volatility of the portfolio s_t as a function of the portfolio holding period h, for three different type of hedges: correlation hedge β^{Corr}, cointegration hedge β, and the optimal hedge β^* respectively.

As expected, the optimal hedge portfolio results in lowest variance, irrespective of the holding period h, see Figure 2. Furthermore, the correlation hedge applied for short holding periods leads to lower volatility of portfolio compared to the cointegration hedge. However, for portfolios held for a long horizon, the cointegration hedge clearly outperforms the former. Asymptotically the volatility of the optimal portfolio becomes identical with volatility of portfolio constructed with the cointegration based hedge. It is remarkable how non-optimal are the portfolios constructed with the correlation hedge β^{Corr}, if they are held for long horizon h. The variance of such portfolios is growing with h and for long horizons it highly surpasses the variance not only of the optimal hedge, but also the variance corresponding to the cointegration hedge.

4 Summary

We derive the optimal hedging ratios for a portfolio of assets driven by a Cointegrated Vector Autoregressive model with general cointegration rank. Our hedge is optimal in the sense of minimum variance portfolio. We start with the exogenous case, in which the hedged asset depends on hedges via a cointegration relation, and the hedges are exogenous, modeled by random walks. Then we consider a model that allows for the hedges to be cointegrated with the hedged asset and among themselves. We find that the minimum variance hedge for assets driven by the CVAR, depends strongly on the portfolio holding period. The hedge is defined as a function of correlation and cointegration parameters. For short holding periods the correlation impact is predominant. For long horizons, the hedge ratio should overweight the cointegration parameters rather then short-run correlation information. In the infinite horizon, the hedge ratios shall be equal to the cointegrating vector. The hedge ratios for any intermediate portfolio holding period should be based on the weighted average of correlation and cointegration parameters.

Our results are general and can be applied for any portfolio of assets that can be modeled by the CVAR of any rank and order. The further research aims at a dynamic version of the developed methodology. In that case the static hedge kept for the entire portfolio holding horizon shall be replaced by a hedge that is dynamically rebalanced during this period.

References

[Grinold and Kahn, 1999] Grinold, R., Kahn, R., (1999), *Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Controlling Risk*, McGraw-Hill, New York

[Johansen 1996] *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models.* Oxford University Press, Oxford (2. ed.)

[Markowitz, 1952] Markowitz, H., (1952), Portfolio selection, *Journal of Finance*, 77-90, 7(1).

[Phillips, 1988] Phillips, P.C.B., (1988), Reflections on econometric methodology, *Economic Record*, 64, 344-359.

[Phillips, 1990] Phillips, P.C.B., (1990), Optimal inference in co-integrated systems, *Econometrica*, 59, 282-306.

[Vayanos, 2004] Vayanos, D., (2004), Flight to Quality, Flight to Liquidity, and the Pricing of Risk, *NBER Working Paper*, 10327.

Appendix

We give the proofs of Theorems 1.1, 1.2, and 2.1, but first we give a simple lemma on the minimization of a variance which will be used in the proofs.

Lemma 1 *Let $y_t = (y_{1t}, y_{2t}) \in \mathbb{R}^n$ be a random variable with $y_{1t} \in \mathbb{R}$, and $y_{2t} \in \mathbb{R}^{n-1}$ and variance matrix*

$$\Sigma = \left(\begin{array}{cc} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{array} \right) > 0.$$

Then

$$\min_{\beta \in \mathbb{R}^{n-1}} Var(y_{1t} - \beta' y_{2t}) = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$$

is attained for

$$\beta^* = \Sigma_{22}^{-1} \Sigma_{21}.$$

Theorem 1.1 *Let $y_t, t = 1, \ldots, T$, be bivariate and given by*

$$y_{1t} = \beta y_{2t} + u_{1t},$$
$$y_{2t} = y_{2,t-1} + u_{2t},$$

where u_i are i.i.d. $(0, \Phi)$. Then the optimal hedge coefficient for the portfolio $s_t = y_{1t} - \beta_h y_{2t}$ at time horizon h is given by

$$\beta_h^* = \beta + h^{-1} \phi_{22}^{-1} \phi_{21} = h^{-1}(\beta(h-1) + \beta^{Corr}),$$

and the minimal variance is

$$Var(y_{1,t+h} - \beta_h^* y_{2,t+h} | \mathcal{I}_t) = \phi_{11} - h^{-1} \phi_{12} \phi_{22}^{-1} \phi_{21}.$$

Proof of Theorem 1.1. The portfolio $s_t = y_{1t} - \beta_h y_{2t}$ at horizon h is

$$
\begin{aligned}
s_{t+h} &= y_{1t+h} - \beta_h y_{2t+h} \\
&= u_{1t+h} - (\beta_h - \beta) y_{2t+h},
\end{aligned}
$$

using the first equation $y_{1,t+h} = \beta y_{2,t+h} + u_{1,t+h}$. Then

$$Var\left(\left(\begin{array}{c} u_{1t+h} \\ y_{2t+h} \end{array} \right) | \mathcal{I}_t \right) = \left(\begin{array}{cc} \phi_{11} & \phi_{12} \\ \phi_{21} & h\phi_{22} \end{array} \right),$$

and the optimal choice of β_h for minimizing $Var(s_{t+h} | \mathcal{I}_t))$ is found from Lemma 1 as

$$\beta_h^* - \beta = h^{-1} \phi_{22}^{-1} \phi_{21}$$

and the minimal variance is

$$Var(y_{1,t+h} - \beta_h^* y_{2,t+h} | \mathcal{I}_t) = \phi_{11} - h^{-1} \phi_{12} \phi_{22}^{-1} \phi_{21}.$$

Using $\beta^{Corr} = \phi_{22}^{-1} \xi_{21} = \phi_{21} \phi_{22}^{-1} + \beta$ the optimal hedge coefficient can be expressed as

$$\beta_h^* = \beta + h^{-1} \phi_{22}^{-1} \phi_{21} = \beta + h^{-1}(\beta^{Corr} - \beta) = h^{-1}(\beta(h-1) + \beta^{Corr}).$$

∎

Theorem 1.2 *Let $y_t = (y_{1t}, y'_{2t})' \in \mathbb{R}^{1+(n-1)}, t = 1, \ldots, T$ be given by*

$$y_{1t} = \beta' y_{2t} + u_{1t},$$
$$y_{2t} = y_{2t-1} + u_{2t},$$

where u_i are i.i.d. $(0, \Phi)$. Then the optimal hedge coefficient for the portfolio $s_t = y_{1t} - \beta_h^{*'} y_{2t}$ at time horizon h is given by

$$\beta_h^* = \beta + h^{-1} \Phi_{22}^{-1} \Phi_{21} = h^{-1}(\beta(h-1) + \beta^{Corr}),$$

and the minimal variance is

$$Var(y_{1,t+h} - \beta_h^* y_{2,t+h} | \mathcal{I}_t) = \Phi_{11} - h^{-1} \Phi_{12} \Phi_{22}^{-1} \Phi_{21}.$$

Proof of Theorem 1.2. The portfolio $s_t = y_{1t} - \beta_h' y_{2t}$ at horizon h is

$$s_{t+h} = y_{1t+h} - \beta_h' y_{2t+h} = u_{1t+h} - (\beta_h - \beta)' y_{2t+h},$$

using the first equation $y_{1,t+h} = \beta' y_{2,t+h} + u_{1,t+h}$. Then

$$Var\left(\begin{pmatrix} u_{1t+h} \\ y_{2t+h} \end{pmatrix} | \mathcal{I}_t \right) = \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & h\Phi_{22} \end{pmatrix},$$

and the optimal choice of β_h for minimizing $Var(s_{t+h} | \mathcal{I}_t)$) is found from Lemma 1 as

$$\beta_h^* - \beta = h^{-1} \Phi_{22}^{-1} \Phi_{21}$$

and the minimal variance is

$$Var(y_{1,t+h} - \beta_h^{*'} y_{2,t+h} | \mathcal{I}_t) = \Phi_{11} - h^{-1} \Phi_{12} \Phi_{22}^{-1} \Phi_{21}.$$

Using $\beta^{Corr} = \Phi_{22}^{-1} \xi_{21} = \Phi_{22}^{-1} \Phi_{21} + \beta$ the optimal hedge coefficient can be expressed as

$$\beta_h^* = \beta + h^{-1} \Phi_{22}^{-1} \Phi_{21} = \beta + h^{-1}(\beta^{Corr} - \beta) = h^{-1}(\beta(h-1) + \beta^{Corr}).$$

∎

Theorem 2.1 *Let $(y_{1t}, y'_{2t})' \in \mathbb{R}^{1+(n-1)}, t = 1, \ldots, T$ be given by*

$$\Delta y_t = \alpha \gamma' y_{t-1} + u_t,$$

where u_t are i.i.d.$(0, \Xi)$ and α and γ are $n \times r$ matrices, and the eigenvalues of $\rho = I_r + \gamma' \alpha$ are inside the unit circle. Then the optimal hedge for the portfolio $s_t = y_{1t} - \beta_h' y_{2t} = (1, -\beta_h')' y_{t+h}$ at time horizon h is given by

$$\beta_h^* = \Sigma_{h22}^{-1} \Sigma_{h21},$$

where

$$\Sigma_h = Var(y_{t+h} | \mathcal{I}_t) = \begin{pmatrix} \Sigma_{h11} & \Sigma_{h12} \\ \Sigma_{h21} & \Sigma_{h22} \end{pmatrix}$$

is given by

$$\Sigma_h = hC\Xi C' - \alpha(\gamma'\alpha)^{-2}(I_r - \rho^h)\gamma'\Xi C' - C\Xi\gamma(I_r - \rho'^h)(\alpha'\gamma)^{-2}\alpha'$$

$$+ \alpha(\gamma'\alpha)^{-1}[\sum_{i=0}^{h-1}\rho^i\gamma'\Xi\gamma\beta\rho'^i](\alpha'\gamma)^{-1}\alpha',$$

and $C = \gamma_\perp(\alpha'_\perp\gamma_\perp)^{-1}\alpha'_\perp$. The minimal variance is

$$Var(y_{1,t+h} - \beta_h^* y_{2,t+h}|\mathcal{I}_t) = \Sigma_{h11} - \Sigma_{h12}\Sigma_{h22}^{-1}\Sigma_{h21}.$$

Proof of Theorem 2.1. The portfolio is $s_{t+h} = (1, -\beta_h')'y_{t+h}$ and

$$\Sigma_h = Var(y_{t+h}|\mathcal{I}_t) = \begin{pmatrix} \Sigma_{h11} & \Sigma_{h12} \\ \Sigma_{h21} & \Sigma_{h22} \end{pmatrix}.$$

The result in Lemma 1 gives the minimizing value of β_h as

$$\beta_h^* = \Sigma_{h22}^{-1}\Sigma_{h21},$$

and the minimal variance

$$Var(y_{1,t+h} - \beta_h^* y_{2,t+h}|\mathcal{I}_t) = \Sigma_{h11} - \Sigma_{h12}\Sigma_{h22}^{-1}\Sigma_{h21}.$$

Thus we only need to find Σ_h. From the equations we find that the cointegrating relation $\gamma'y_t$ is an $AR(1)$ process with autoregressive parameter $\rho = I_r + \gamma'\alpha$,

$$\gamma'y_t = \rho\gamma'y_{t-1} + \gamma'u_t.$$

Under the assumption that the eigenvalues of ρ are inside the unit circle, the stationary solution is $\gamma'y_t = \sum_{i=0}^{\infty}\rho^i\gamma'u_{t-i}$, such that

$$\gamma'y_{t+h} = \sum_{i=0}^{h-1}\rho^i\gamma'u_{t+h-i} + \sum_{i=h}^{\infty}\rho^i\gamma'u_{t+h-i}$$

$$= \sum_{i=t+1}^{t+h}\rho^{t+h-i}\gamma'u_i + \rho^h\sum_{i=0}^{\infty}\rho^i\gamma'u_{t-i} = \sum_{i=t+1}^{t+h}\rho^{t+h-i}\gamma'u_i + \rho^h\gamma'y_t.$$

Because $\alpha'_\perp y_t$ is a random walk we find

$$\alpha'_\perp y_{t+h} = \sum_{i=t+1}^{t+h}\alpha'_\perp u_i + \alpha'_\perp y_t.$$

Using the

$$I_n = \gamma_\perp(\alpha'_\perp\gamma_\perp)^{-1}\alpha'_\perp + \alpha(\gamma'\alpha)^{-1}\gamma' = C + \alpha(\gamma'\alpha)^{-1}\gamma',$$

we then find

$$y_{t+h} = (C + \alpha(\gamma'\alpha)^{-1}\gamma')y_{t+h}$$

$$= \sum_{i=t+1}^{t+h}(Cu_i + \alpha(\gamma'\alpha)^{-1}\rho^{t+h-i}\gamma'u_i) + (C + \alpha(\gamma'\alpha)^{-1}\rho^h\gamma')y_t = z_{1t} + z_{2t},$$

17

where z_{1t} is independent of the past $\mathcal{I}_t$ and therefore z_{2t}. Therefore we get, using $\sum_{i=0}^{h-1} \rho^i = (I_r - \rho)^{-1}(I_r - \rho^h) = -(\gamma'\alpha)^{-1}(I_r - \rho^h)$, that

$$\Sigma_h = Var(z_{1t}) = \sum_{i=0}^{h-1}[C + \alpha(\gamma'\alpha)^{-1}\rho^i\gamma']\Xi[C' + \gamma\rho'^i(\alpha'\gamma)^{-1}\alpha']$$

$$= hC\Xi C' - \alpha(\gamma'\alpha)^{-2}(I_r - \rho^h)\gamma'\Xi C' - C\Xi\gamma(I_r - \rho'^h)(\alpha'\gamma)^{-2}\alpha'$$

$$+ \alpha(\gamma'\alpha)^{-1}\sum_{i=0}^{h-1}\rho^i\gamma'\Xi\gamma\rho'^i(\alpha'\gamma)^{-1}\alpha'.$$
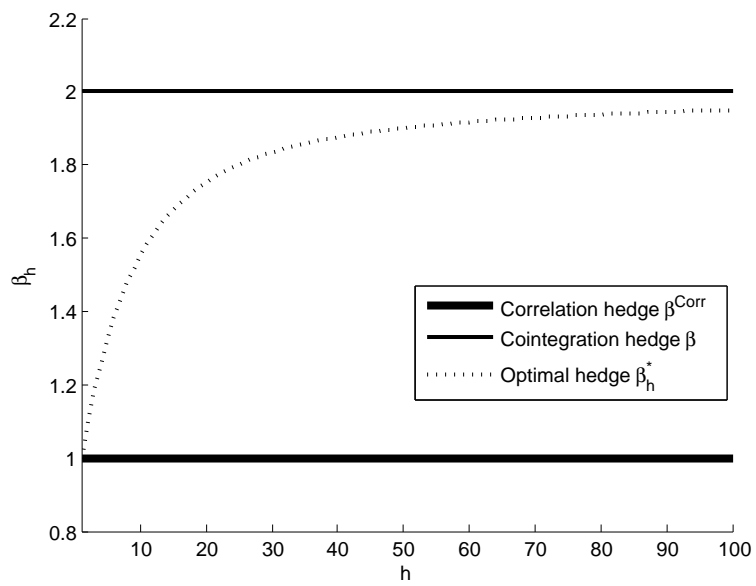
■

Figure 1: Optimal hedge coefficient $\beta_h^\star$ under limited mean reversion and high covariance.
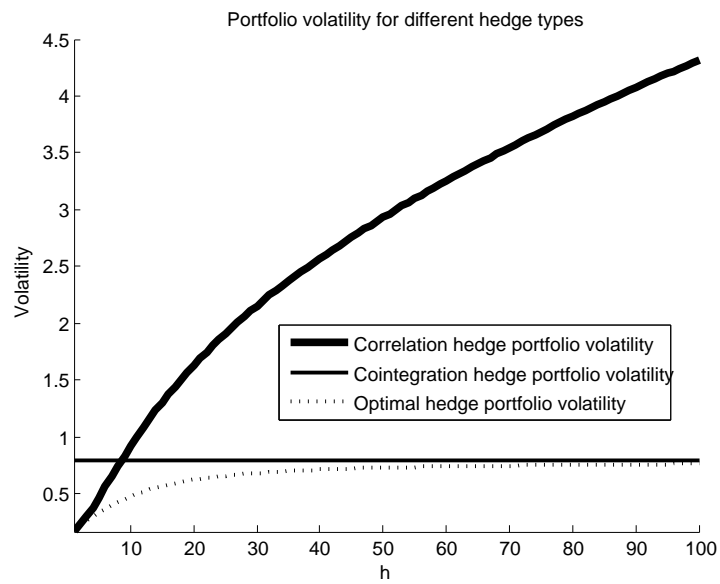
Figure 2: Volatility of various portfolios under limited mean reversion and high covariance for the correlation hedge β^{Corr} (thick line), the cointegration hedge β (thin line) and for the optimal hedge $\beta_h^\star$ (dotted line) for different holding periods.