

Lang, Harald

**Working Paper**

## Herman Wold on Optimal Properties of Exponentially Weighted Forecasts

IUI Working Paper, No. 179

**Provided in Cooperation with:**

Research Institute of Industrial Economics (IFN), Stockholm

*Suggested Citation:* Lang, Harald (1987) : Herman Wold on Optimal Properties of Exponentially Weighted Forecasts, IUI Working Paper, No. 179, The Research Institute of Industrial Economics (IUI), Stockholm

This Version is available at:

<https://hdl.handle.net/10419/94942>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

A list of Working Papers on  
the last pages

No. 179, 1987

**HERMAN WOLD ON OPTIMAL PROPERTIES OF  
EXPONENTIALLY WEIGHED FORECASTS**

by  
Harald Lang

This is a preliminary paper.  
Comments are welcome.

Oct., 1987

Herman Wold on Optimal Properties of  
Exponentially Weighed Forecasts

by  
Harald Lang

Herman Wold's doctoral thesis "A Study in the Analysis of Stationary Time Series" (Wold (1938)) is very often referred to in the econometrics and time series literature, and the result referred to is most often the "canonical representation" of a stationary time series that appears in chapters 1 and 2. But there are two more chapters that seem to be less well known. As an example, I will in this note point out that John Muth's result in "Optimal Properties of Exponentially Weighed Forecasts" (Muth (1960)) is already contained as a special case in Wold (1938).

When Milton Friedman presented his famous "Permanent Income Hypothesis" (Friedman(1957)) he used an exponentially weighed sum of current and lagged observations on actual income as an identifier of permanent income, and in Muth (1960) John Muth investigated the optimality of this measure. The setup was as follows: Assume that log of actual income  $y_t$  is the sum of three terms

$$y_t = \alpha + w_t + v_t \quad (1)$$

where  $\alpha$  is a constant,  $w_t$  is (log of) permanent income and  $v_t$  is transitory income. The term  $v_t$  is assumed to be "white noise" and  $w_t$  is assumed to be a "random walk". The increment  $dw_t$  of  $w_t$  and  $v_t$  may be contemporaneously, but not intertemporally correlated. For notational simplicity we assume that  $dw_t$  and  $v_t$  are uncorrelated. Under these assumptions Muth shows that the optimal forecast of  $y_{t+1}$  (which is the same as the optimal measure of  $w_t$ ) is the exponentially weighed sum

$$y_{t+1}^* = (1-\beta) \sum_{i=0}^{\infty} \beta^i y_{t-i} \quad (2)$$

where  $\beta$  is given by

$$\beta = 1 + \frac{\sigma_{dw}^2}{2\sigma_v^2} - \frac{\sigma_{dw}}{\sigma_v} \left[ 1 + \frac{\sigma_{dw}^2}{4\sigma_v^2} \right]^{1/2} \quad (3)$$

In chapters 3 and 4 Wold (op. cit.) studies the following question: let  $\theta_t$  be a stationary stochastic process with zero mean. Under what conditions is there an auto-regressive representation of  $\theta$ :

$$\theta_t = \sum_{i=1}^{\infty} a_i \theta_{t-i} + \epsilon_t \quad (A)$$

(where  $\epsilon_t$  is white noise) and when is there a moving average representation:

$$\theta_t = \sum_{i=0}^{\infty} m_i \epsilon_{t-i} \quad (M)$$

Observe that when (A) and (M) exist

$$\theta_t^* = \sum_{i=1}^{\infty} a_i \theta_{t-i}$$

is the optimal forecast of  $\theta_t$ , given  $\theta_{t-i}$ ,  $i > 0$ .

Let us introduce some notation: call the function

$$A(z) = 1 - \sum_{i=1}^{\infty} a_i z^i$$

( $z$  being a complex number) the characteristic function of the auto-regressive representation (A), if it exists, and the function

$$M(z) = \sum_{i=0}^{\infty} m_i z^i$$

the characteristic function of the moving average representation (M). We also introduce the characteristic function of the correlogram of  $\theta$  according to

$$R(z) = \sum_{i=-\infty}^{+\infty} r_i z^i$$

where  $r_i = \text{corr}(\theta_t, \theta_{t-i})$ .

Some of the results that Wold derived in chapters 3 and 4 may now be summarised as the following theorem:

Theorem: Assume that  $r_i = 0$  for  $i > N$  and that  $R(z)$  has no complex roots on the unit circle. Then the autoregressive representation (A) as well as the moving average representation (M) exist. The characteristic functions are obtained as follows: Let  $\alpha_1, \dots, \alpha_N$  be the zeroes of  $R(z)$  lying inside the unit circle (if  $r_N \neq 0$  it is easy to see that there are exactly  $N$  of them). Then

$$M(z) = (1 - \alpha_1 z) \cdot (1 - \alpha_2 z) \cdot \dots \cdot (1 - \alpha_N z)$$

and

$$A(z) = \frac{1}{M(z)} \quad (4)$$

Let us now see how Muth's result is a special case of this theorem.

Taking first differences of (1) gives

$$dy_t = dw_t + v_t - v_{t-1}$$

We see that the autocorrelation coefficients  $r_i$  of  $dy$  are

$$r_0 = 1, \quad r_1 = r_{-1} = \frac{-\sigma_v^2}{\sigma_{dw}^2 + 2\sigma_v^2}, \quad r_n = 0 \text{ for } |n| > 1$$

Hence  $R(z) = r_{-1}z^{-1} + 1 + r_1z$  has two roots:

$$z_{1,2} = -\frac{1}{2r_1} \pm \left[ \frac{1}{4r_1^2} - 1 \right]^{1/2}$$

none of which lies on the unit circle. Hence, by the theorem above, (A) and (M) exist and since the root lying inside the unit circle is

$$\beta = -\frac{1}{2r_1} - \left[ \frac{1}{4r_1^2} - 1 \right]^{1/2} \quad (5)$$

we have  $M(z) = 1 - \beta z$  and

$$A(z) = \frac{1}{1 - \beta z} = \sum_{i=0}^{\infty} \beta^i z^i$$

This means that

$$dy_t = - \sum_1^{\infty} \beta^i dy_{t-i} + \epsilon_t$$

which implies

$$y_{t+1} = (1-\beta) \sum_0^{\infty} \beta^i y_{t-i} + \epsilon_t$$

and hence

$$y_{t+1}^* = (1-\beta) \sum_0^{\infty} \beta^i y_{t-i}$$

which is Muth's result (2). If we plug the expression for  $r_1$  into (5), we obtain Muth's expression (3).

The reader who searches Wold's work for this specific result is apt to be disappointed. For Wold's result is never expressed as a single theorem; it is developed at various points in chapters 3 and 4. Wold does not, moreover, use the "characteristic functions" of (A) and (M). He writes down the recursive equations corresponding to e.g. eq. (4). But the result (and much more) is all there.

#### REFERENCES

- Friedman M. (1957): A Theory of the Consumption Function, Princeton University Press, Princeton.
- Muth J.F. (1960): Optimal Properties of Exponentially Weighed Forecasts, J.A.S.A. vol.55, no.290.
- Wold H. (1938): A Study in the Analysis of Stationary Time Series, Almqvist & Wiksell, Uppsala.