DISCUSSION PAPERS IN ECONOMICS

Working Paper No. 00-06

Minimum Dispersion and Unbiasedness:
'Best' Linear Predictors for
Stationary ARMA *a*-Stable Processes

Jonathan B. Hill
Department of Economics, University of Colorado at Boulder
Boulder, Colorado

September 2000

Center for Economic Analysis
Department of Economics

University of Colorado at Boulder
Boulder, Colorado 80309

: ” ”
A A α-

by

Jonathan B. Hill*

Dept. of Economics, University of Colorado at Boulder

Job Market Paper 1

The present paper is currently under review for pu lication in

Stochastic Processes and their Applications.

Abstract

Many volatile financial time series have been assumed to be driven by a distribution with an infinite population variance beginning with the seminal observations by Mandlebrot (1963) and Fama (1965). Although a flourish of statistical and econometric research

e Words and Phrases: Best Linear Predictors, stable-law, infinite variance, covariation

AM 1991 ubject Classifications. Pri ar : 60 G25; **econdar :** 62 10.

* *Dept. of Economics, Campus Box 256, University of Co orado at Bou der, Bou der Co o. 80309-0256*, jonathan.hill@colorado.edu

in the fields of estimation and inference for processes with infinite variance has been the result, few attempts have been made to charact

1. Introduction

In the present paper, we are interested in methods of optimal linear forecasts of the causal-invertible $A \quad MA$ process $X_n - \phi_1 X_{n-1} - ... - \phi_p X_{n-p} = \epsilon_n + \theta_1 \epsilon_{n-1} + ... + \theta_q \epsilon_{n-q}$ where the innovations ϵ_t are *iid* non-normal α-stable. The stable-laws[1] have been the focus of a growing body of empirical and theoretical research in financial and macro-economics since the now classical studies by Mandlebrot (1965) and Fama (1963) on the distributional behavior of common asset prices. In particular, in those and subsequent studies (e.g. Cheng and   achev, 1995, Jansen and de Vries, 1991, and McCulloch, 1984, 1987, 1996, 1997) the empirical distributional characteristics of asset returns, asset prices, option prices, and forward and spot exchange rates have in many cases displayed the characteristic of having "heavy distribution tails"; indeed, there exists evidence for a substantially high probability of large deviations. Importantly, the odds of large positive or negative swings are typically too great to be modelled by a normal distribution, and without further behavioral information (e.g. structural breaks[2]), choice of the assumed underlying

[1] Recall that any stable random variable with α

distribution would be necessarily arbitrary.

Stable random variables, however, have the property of infinite population variance (i.e. heavy tails) for all stable-laws except the normals, and are conveniently distributionally "stable" under addition: the sum of stables is itself a stable random variable. This property of stable summation is particularly attractive for modelling the behavioral properties of volatile low frequency (e.g. weekly/mothly) asset returns which are constructed as sums (products) of high frequency returns (e.g. daily/hourly) . For a general survey on the development of pricing models for stably-distributed assets, see McCulloch (1996). See, also, Gamrowski and achev (1999), and the citations therein, for the recent development of a stable-law *capi al asse pricing model*.

Although research in statistical and eco ing out-of-sample values of the infinite variance process X based on the available information $X_1, ..., X_n$. In particular, we develop methods which dramatically simplify existing solution mechanics for deriving the optimal linear forecast, and provide prediction which uniformly out-performs existing computationally burdensome techniques.

If we denote by L the lag operator, we have

$$\phi_p(L)X_t = \theta_q(L)\epsilon_t, \tag{1}$$

where we restrict the polynomials $\phi_p(L) = 1 + z\phi_1 +...+ z^p\phi_p$ and $\theta_p(L) = 1 + z\theta_1 +...+ z^q\theta_q$ to have no common roots, and to satisfy

$$\phi_p(L)\theta_p(L) \neq 0 \quad \forall z \in R \text{ such that } |z| < 1. \tag{2}$$

In particular, we assume $\phi_p(L)$ has no roots outside the unit circle. It follows (see Brockthe *mmse* predictor is optimally linear and identically the conditional expectations. The linear space of the stable-laws with $1 \leq \alpha < 2$, however, is a Banach space, and whenever $\alpha < 1$, only a metric space. Indeed, when $\alpha < 2$, in most cases we must choose between minimizing an acceptable L_ρ-metric, or employing the traditional conditional expectations form, $E\left[X_{n+k}|X_1, ..., X_n\right]$

standard c.f. manipulation dictates that

$$X_n \overset{D}{=} \epsilon_1 \left(\sum_{i=1}^{\infty} |\pi_i|^\alpha \right)^{1/\alpha} \qquad C_{X_n}^\alpha = C_\epsilon^\alpha \sum_{i=1}^{\infty} |\pi_i|^\alpha < \infty \tag{6}$$

where the finiteness of the resulting scale follows necessarily from (3). Thus, minimization of the prediction error dispersion is identical to the problem of finding a vector **a** which minimizes the stable-scale associated with the prediction error, provided the error is stably distributed. In order to see that the prediction error $\hat{e} \equiv \hat{X}_{n+k} - \sum_{i=1}^{n} a_i X_{n+1-i}$ is governed by a stable-law $\forall n \geq 1$, observe that

$$\begin{aligned} X_{n+k} - \sum_{i=1}^{n} a_i X_{n+1-i} &= \sum_{i=0}^{\infty} \pi_i \epsilon_{n+k-i} - \sum_{j=1}^{n} a_j \sum_{i=0}^{\infty} \pi_i \epsilon_{n+1-j-i} \\ &= \sum_{i=0}^{k-1} \pi_i \epsilon_{n+k-i} + \sum_{i=1}^{n} \left(\pi_{n+k-i} - a_1 \pi_{n-i} - \ldots - a_{n-i+1} \pi_0 \right) \epsilon_i \\ &\quad + \sum_{i=0}^{\infty} \left(\pi_{n+k+i} - a_1 \pi_{n+i} - \ldots - a_n \pi_i \right) \epsilon_{-i} \\ &= \sum_{i=0}^{\infty} \lambda_i \epsilon_{n+k-i} \overset{d}{=} \epsilon_1 \left(\sum_{i=0}^{\infty} |\lambda_i|^\alpha \right)^{1/\alpha}, \end{aligned} \tag{7}$$

where $\sum_{i=0}^{\infty} |\lambda_i|^\alpha < \infty$ follows from implies a computationally extensive solution technique, and when $\alpha < 1$ the solution set is typically not unique. See Cline and Brockwell (1985: Theorem 3.2 and Lemma 4.1).

$+ \ b^{<\alpha-1>} \left[z_3, bz_2\right]_\alpha$ by inspection of the stable characteristic function and spectrum, Cf. Miller (1977) of linear unbiased predictors.

The rest of the paper is organized as follows. In Section 2, we derive the relationship between the conditional expectations, and the unbiased and minimum dispersion linear predictors. e subsequently construct optimal projection functions for asymptotic and truncated predictors, and explicitly derive optimal solutions according to the covariation-orthogonality criterion for A ( ) and A $MA(1,1)$ processes in Sections 3 and 4. Section 5 follows with derivations for higher order moving average and A MA processes, including the characterization of a fast numeircal algorithm for solving the best linear unbiased predictor based on resursive prediction residuals. Section 6 concludes with a numerical comparison of the minimum dispersion and unbiased predictors.

2. Conditional Expectations and Covariation-Orthogonal Linear Predictors Our investigation begins with an articulation of the relationship between multivariate conditional expectations, the COLP, and predictor unbiasedness when the innovations sequence ϵ_t is governed by an *iid* stable-law. Further details can be found in Cambanis and u (1992) and Samorodnitsky and Taqqu (1994). Throughout, we assume $\alpha \in (0,2)$, $\alpha \neq 1$.

Lemma 1 (Existence) *Le* $X_1, ..., X_n; X_{n+k}$ *be join ly* α

which is, by definition, (iii). e conclude this proof by establishing the identity between (i) and (v). Now, consider any linear predictor, say

$$\hat{X}_{n+k} = \sum_{i=1}^{n} b_i X_{n+1-i}, \tag{17}$$

and observe that we have the tautological regression form

$$X_{n+k} = \sum_{i=1}^{n} b_i X_{n+1-i} + \tilde{e} \tag{18}$$

where $\tilde{e} = X_{n+k} - \sum_{i=1}^{n} b_i X_{n+1-i}$ is governed by a stable-law $\forall b_i \in (-\infty, \infty)$, Cf. (7). Therefore, by (3) and (5), is $(E[\tilde{e}|X_1, ..., X_n])$ is identically

$$\sum_{i=0}^{n-1} \left| \sum_{j=0}^{i} \pi_{i-j} \left(a_{j+1} - b_{j+1}\right) \right|^{\alpha} + \sum_{i=0}^{\infty} \left| \pi_{n+i} \left(a_1 - b_1\right) + ... + \pi_i \left(a_n - b_n\right) \right|^{\alpha},$$

which is trivially minimized by setting $b_j = a_j \ \forall j = 1..n$. ■

emark 1: The identity of (i) - (v) holds for any (possibly maximally) skewed stable $n+1$ vector $(\mathbf{X}, Y)$. Indeed, in general we do not even require the stipulations detailed in (1) - (3)

long to the domain of attraction of a stable-law are not guaranteed to generate a linear conditional expectations. In particular, condition (iii) will not be appropriate, and conditions (i) and (iv) are certainly not necessarily implied. See Cioczek and Taqqu (1994).

e leave for Sections 3 and 4 details on the uniqueness of the solution set **a** in Lemma 1.

3. Linear Prediction with (P) Processes e begin by establishing a general result which characterizes the unique linear projection $\hat{X}_{n+k} = P(X_{n+k}, \mathbf{X})$ based on *covaria ion-or hogonali y.* The following lemma establishes concretely that the COLP is the MDLP as $n \to \infty$ for any stationary $A \;\; MA$ α-stable processes. e subsequently derive the COLP for any finite order autoregressive process, and conclude the section by proving the COLP is the best linear unbiased estimator for any stationary $A \;\; MA$ α-stable processes and any $n \geq 1$. In order to exploit the property of quasi-linearity in the second argument of the covariation, we require the assumption that in all cases the vector $(X_1, ..., X_n; Y)$ ˜ $S\alpha S$. It should be pointed out that in the joint $S\alpha S$ case, (13) is trivially satisfied (Cambanis and u, 1992), hence it suffices to consider condition (iii) of Lemma 1.

Lemma 2 (Uniqueness) *For any A MA(, q) process X_n assume (1 -(3 are rue, and deno e by $\hat{S}$ he class of random variables of he form*

$$\sum_{j=n+1}^{\infty} \rho_j \epsilon_j + \sum_{j=1}^{\infty} \nu_j X_{n+1-j} \tag{19}$$

where $\sum_{i=1}^{\infty} |\rho_j|^{\delta} < \infty$ and $\sum_{i=1}^{\infty} |\nu_j|^{\delta} < \infty$ for some $\delta < \min(1, \alpha)$. Then, for any $Y \in \hat{S}$, he se

$$P_n^Y = \left\{ \sum_{i=1}^{n} a_i X_{n+1-i} : E\left[Y - \sum_{i=1}^{n} a_i X_{n+1-i} | X_n, ..., X_1 \right] = 0 \right\}$$

consis s of exac ly one elemen , $\hat{Y}_n = \sum_{i=1}^{n} \nu_i X_{n+1-i}$, *as* $n \to \infty$. *Moreover, he mapping* $Y \to \hat{Y}_n$ *is linear on* $\hat{S}$. *Fur her, as* $n \to \infty$ *he random variable* $\hat{Y}_n$ *is he unique elemen of he se*

$$\acute{P}_n^Y = \left\{ \sum_{i=1}^{n} a_i X_{n+1-i} : \ \textit{is} \ \left(Y - \sum_{i=1}^{n} a_i X_{n+1-i} \right) \ \textit{is minimized} \right\}.$$

Proof. For any $Y \in \hat{S}$,

$$\begin{aligned} Y - \sum_{i=1}^{n} a_i X_{n+1-i} &= \sum_{j=n+1}^{\infty} \rho_j \epsilon_j + \sum_{j=1}^{n} \epsilon_{n+1-j} \left[\sum_{i=1}^{j} (\nu_i - a_i)\, \pi_{j-i} \right] \\ &\quad + \sum_{j=n+1}^{\infty} \epsilon_{n+1-j} \left[\sum_{i=1}^{n} (\nu_j - a_j)\, \pi_{j-i} + \sum_{i=n+1}^{j} \nu_j \pi_{j-i} \right], \end{aligned} \tag{20}$$

consequently, by identities (ii) and (iii) of Lemma 1, $\forall t = 1..n$, it suffices to solve

$$\left[Y - \sum_{i=1}^{n} a_i X_{n+1-i}, X_{n+1-t} \right]_{\alpha} = \sum_{j=0}^{n-t} \lambda_{j+t} \pi_j^{<\alpha-1>} + \sum_{j=n+1}^{\infty} \varphi_{j+t} \pi_j^{<\alpha-1>} = 0 \tag{21}$$

where

$$\lambda_j = \sum_{i=1}^{j} (\nu_i - a_i)\, \pi_{j-i} \quad \varphi_j = \sum_{i=1}^{n} (\nu_i - a_i)\, \pi_{j-i} + \sum_{i=n+1}^{j} \nu_i \pi_{j-i}. \tag{22}$$

In the limit, the sufficiency of $a_j = \nu_j$, $j = 1..n$, is trivial. Now, by (21) for $t = n$, $\lim_{n\to\infty} \left(\sum_{i=1}^{n} (\nu_i - a_i)\, \pi_{n-i} \right) \pi_0^{<\alpha-1>} = 0$, and when $t = n - 1$,

$$\lim_{n\to\infty} \left(\sum_{i=1}^{n-1} (\nu_i - a_i)\, \pi_{n-1-i} \right) \pi_0^{<\alpha-1>} + \left(\sum_{i=1}^{n} (\nu_i - a_i)\, \pi_{n-i} \right) \pi_1^{<\alpha-1>} = 0.$$

For non-degenerate sequences $\{\pi_j\}_{j\geq 1}$, we conclude necessarily that $\nu_n = a_n$ as $n \to \infty$. It follows recursively that as $n \to \infty$ equality in (21) holds *if and only if* $a_j = \nu_j$, $j = 1..n$, hence the unique element of the set P_n^Y as $n \to \infty$ is $\hat{Y}_n = \sum_{i=1}^{n} \nu_i X_{n+1-i}$, as claimed.

The linearity of the mapping $Y \to \hat{Y}_n$ follows easily by observing that for any $Y = \sum_{i=1}^{m} b_i Y_i$, $Y_i \in \hat{S}$, $i = 1..n$, then

$$Y = \sum_{j=n+1}^{\infty} \tilde{\rho}_j \epsilon_j + \sum_{j=1}^{\infty} \tilde{\nu}_j X_{n+1-j},$$

Corollary 3 *For any A MA(, q) process X_n assume (1 -(3 are rue, and deno e by $\hat{S}$ he class of random variables defined in Lemma 2. Moreover, deno e by E_0 he even   ha $E[Y - \sum_{i=1}^{n} a_i X_{n+1-i} | X_1, ..., X_n] = 0$. Consequen ly, for any $Y \in \hat{S}$, he se*

$$\tilde{P}_n^Y = \left\{ \sum_{i=1}^{n} a_i X_{n+1-i} : \ is \ \left(Y - \sum_{i=1}^{n} a_i X_{n+1-i} | E_0 \right) \ is\ minimized \right\}$$

consis s of exac ly one elemen , $\hat{Y}_n = \sum_{i=1}^{n} \tilde{\nu}_i X_{n+1-i}$, for any $n \geq 1$ where $\tilde{\nu}_i$ solves he following implici sys em of n-equa ions,

$$\sum_{i=1}^{n} a_i \lambda_{i,t} = \sum_{i=1}^{n} \lambda_{i,t} \tilde{\nu}_{i,t}, \tag{24}$$

where

$$\begin{aligned} \lambda_{i,t} &= \pi_i \sum_{j=1-t}^{n-t} \pi_j^{<\alpha-1>} + \sum_{j=n+1}^{\infty} \pi_j^{<\alpha-1>} \pi_{j-i} \qquad (25) \\ \tilde{\nu}_{i,t} &= \nu_i + \frac{\omega_{n,t}}{n \lambda_{i,t}} \\ \omega_{n,t} &= \sum_{i=0}^{t} \nu_{n+1+i} \left(\sum_{j=0}^{\infty} \pi_{j+t-1} \pi_{n+1+j}^{<\alpha-1>} \right) + \sum_{i=t+1}^{\infty} \nu_{n+1+i} \left(\sum_{j=0}^{\infty} \pi_j \pi_{n+1+j}^{<\alpha-1>} \right). \end{aligned}$$

Moreover, he mapping $Y \to \hat{Y}_n$ is linear on $\hat{S}$. Finally, $\lim_{n\to\infty} \tilde{\nu}_i = \lim_{n\to\infty} \tilde{\nu}_{i,t} = \nu_i$ for every $i = 1, 2, ...$

emark 1: The set $\tilde{P}_n^Y$ defines the unique truncated covariation orthogonal projection of $\hat{S} \to \overline{s\ an}(X_1, ...X_n)$, the space of all linear combinations of the available data.

For large $n < \infty$, (21) and (22) demonstrate that the truncated COLP will be approximately the MDLP as n grows large. Conversely, for large n the MDLP will be roughly unbiased. However, for general A ( ) processes, the following result establishes that the truncated COLP is identically the MDLP provided $n \geq$.

Theorem 4 *Le $X_1, ..., X_n; X_{n+k}$ be join ly α-s able wi h $\alpha \in (0, 2)$, and assume (1 -(3 holds wi h $q = 0$, $n \geq$. Then, provided $k = 1$, (i - (v of Lemma 1 each provide a_j*

$= \phi_j$, $j = 1..$. *Moreover,* $\forall k \geq 1$ *we ob ain he recursive rela ionship*

$$\hat{X}_{n+k} = \phi_1 \hat{X}_{n+k-1} + ... + \phi_p \hat{X}_{n+k-p} \tag{26}$$

where op imally $\hat{X}_j = X_j \ \forall j \leq n$. *Fur her,* $\hat{X}_{n+k}$ *ob ains he minimum level of error dispersion.*

Proof. (26) is immediate by Lemma 2 and Corollary 3. For the resulting minimized level of dispersion, see Cline and Bockwell *where we define*

$$= \frac{|\theta + \phi|^{\alpha}}{1 - |\phi|^{\alpha}} \quad \eta = \theta \left[(\theta + \phi)^{<\alpha-1>} - \phi^{k-1} \right]. \tag{29}$$

The corresponding predic ion error is

$$1 + \quad \left(1 - |\phi|^{\alpha\ k-1)}\right) +$$

$$|\theta^{n}|^{\alpha} \Big| \quad \phi^{k-1}$$

for $t = n - 1$

$$\left[X_{n+k} - \sum_{i=1}^{n} a_i X_{n+1-i}, X_{n+1-t}\right]_\alpha = \theta \; _{n-2} + \left[1 + (\theta + \phi)^{<\alpha-1>}\theta\right] \; _{n-1} + \left[(\theta + \phi)^{<\alpha-1>} + \phi^{<\alpha-1>}\left(\frac{|\theta + \phi|^\alpha}{1 - |\phi|^\alpha}\right)\right] \; _{n} = 0, \tag{33}$$

and when $t = n$,

$$\left[X_{n+k} - \sum_{i=1}^{n} a_i X_{n+1-i}, X_{n+1-t}\right]_\alpha = \theta \; _{n-1} + \left[1 + \frac{|\theta + \phi|^\alpha}{1 - |\phi|^\alpha}\right] \; _{n} = 0. \tag{34}$$

This system of n-equations offers a recursive solution in the following stepwise manner. From (34),

$$_{n} = -\left(\frac{\theta}{1 + \;}\right) \; _{n-1}, \tag{35}$$

which is substituted into (33) providing

$$_{n-1} = -\theta \; _{n-2}\left[\frac{1 + \;}{1 + \;\;(1 + \eta)}\right]. \tag{36}$$

Subsequently, employing (32) recursively, we deduce that

$$_{j} = -\theta \; _{j-1}\left[\frac{1 + \;\; \sum_{i=1}^{n-j} \eta^{i-1}}{1 + \;\; \sum_{i=1}^{n} \eta^{i-1}}\right]. \tag{37}$$

Observing that $\; _0 = -\phi^{k-1}$

Subsequently, from (38) we deduce

$$
\begin{aligned}
{}_j + \theta \ {}_{j-1} &= \phi^{k-1}(-\theta)^j \left(\frac{(1-\eta)\eta^{n-j}}{1-\eta + \ (1-\eta^n)} \right) \\
| \ {}_j + \theta \ {}_{j-1}| &= |\theta^n|^\alpha \left| \phi^{k-1}\theta^{j-n}\eta^{n-j} \left(\frac{(1-\eta)}{1-\eta + \ (1-\eta^n)} \right) \right|^\alpha \\
&= |\theta^n|^\alpha |\tilde{\eta}^j|^\alpha \left| \frac{\phi^{k-1} \ (1-\eta)}{1-\eta + \ (1-\eta^n)} \right|^\alpha ,
\end{aligned}
$$

hence

$$
\begin{aligned}
& is \ (X_{n+k} - \mathbf{a}'X) \\
= & \ 1 + \ \left(1 - |\phi|^{\alpha \ k-1)}\right) + |\theta^n|^\alpha \left| \frac{\phi^{k-1} \ (1-\eta)}{1-\eta + \ (1-\eta^n)} \right|^\alpha \left[\sum_{j=0}^{n-1} |\tilde{\eta}^j|^\alpha + 1 \right] \\
\to & \ 1 + \ \left(1 - |\phi|^{\alpha \ k-1)}\right)
\end{aligned}
$$

as $n \to \infty$, the minimum level of dispersion follows from (28). Moreover,

$$\hat{Y} = \left(b_1 + b_2\phi + ... + b_k\phi^{k-1}\right)\hat{X}_{n+1},$$

which follows by recursively solving (32) - (34) with $\ _0 = -\left(b_1 + b_2\phi + ... + b_k\phi^{k-1}\right)$. This linear structure is identical to the truncated MDLP linearity property except, of course, for the unbiased one step-ahead predictor $\hat{X}_{n+1}$ itself. It follows trivially that any linear combination of biased minimum dispersion linear predictors will be biased.

emark 4: By the definition of covariation, when $\alpha = 2$, the COLP is identically the least squares predictor. See Brockwell and Davis (1983) and Cline and Brockwell (1985) for accounts of optimal recursive methods for deriving linear predictors in the finite variance case..

The following corollary is an immediate result of Theorems 4 and 5.

Corollary 6 *Le* $(X_1, ..., X_n; X_{n+k})$ *be join ly* $S\alpha S$, *and assume (3 holds. Provided* X_n *is A MA*$(1,0)$ *and* $|\phi| < 1$, $a_1 = \phi^k$ *and* $a_i = 0\ \forall i > 1$. *In par icular, he COLP minimizes he predic ion error dispersion. Moreover, whenever* X_n *is A MA*$(0,1)$ *wi h* $k = 1$,

$$a_j = -(\theta^{j-1})\left[\frac{1 - |\theta|^{\alpha\ n+1-j)}}{1 - |\theta|^{\alpha\ n+1)}}\right] \quad j = 1..n. \tag{39}$$

emark 1: Observe that in the A (1) case, the optimal k-ahead COLP is merely $\hat{X}_{n+k} = \phi^k X_n$, and in the $MA(1)$ for any $k \geq 2$, $\hat{X}_{n+k} = 0$.

Example e consider two A $MA(1,1)$ cases. Let $X_t = .3X_{t-1} + .8\epsilon_{t-1} + \epsilon_t$, and put $C_\epsilon = 1$, $\alpha = 1.75$, $n = 3$ and $k = 1$. As a consequence of Theorem 5 and Cline and Brockwell (1985: Theorem 3.2) the COLP and MDLP are respectively

$$\hat{X}_4^C = .9922X_3 - .6164X_2 + .2542X_1$$

$$\hat{X}_4^M = .8959X_3 - .5435X_2 + .2339X_1.$$

The resulting levels of prediction error dispersion and $is\ \left(E\left[\hat{e}|X_1,...,X_n\right]\right)$ are

$$
\begin{aligned}
is\ \left(\hat{e}^C\right) &= .16252 & \quad is\ \left(E\left[\hat{e}^C|X_1,...,X_n\right]\right) = 0.0 \\
is\ \left(\hat{e}^M\right) &= .15046 & \quad is\ \left(E\left[\hat{e}^M|X_1,...,X_n\right]\right) = .70313.
\end{aligned}
$$

Finally, let $X_t = .9X_{t-1} - .25\epsilon_{t-1} + \epsilon_t$, and put $C_\epsilon = 1$, $\alpha = 1.2$, $n = 5$ and $k = 5$. Then,

$$
\begin{aligned}
\hat{X}_{10}^C &= .42740X_5 + .10625X_4 + .026414X_3 + .0065739X_2 + .0021326X_1 \\
\hat{X}_{10}^M &= .42647X_5 + .10662X_4 + .026654X_3 + .0066641X_2 + .0023058X_1,
\end{aligned}
$$

where $\theta_j = 0$, $j < 0$, $j > n$, $\theta_0 = 1$ and $a_j = 0$, $j \leq 0$, $j > n$. The above implicit linear system of equations can be immediately solved upon application of Cr mer's ule, which completes the proof. ■

emark 1: The solution set $\mathbf{a}$ is unique $\forall \alpha \in (0, 2)$. Observe that $\forall \alpha < 1$ the MDLP reduces the choice set of $\mathbf{a}$ to $\binom{n+q}{q}$ possibilities as functions of the solutions to the polynomial $z^q + \theta_1 z^{q-1} + ... + \theta^q = 0$, Cf. Cline and Brockwell (1985: p. 293). Thus, the computational burden will be comparatively extensive for higher order moving average processes for large n.

Finally, consider the general $A \;\; MA(\;\; , q)$ processed denoted in (1) - (3), and define the following sequence

$$\varphi_1 = 1, \quad \varphi_i = \sum_{j=1}^{i} \psi_{j-i+1} \varphi_{i-1} \quad i = 2..k, \tag{44}$$

where the coefficients $\{\psi_j\}_{j \geq 1}$ are defined in (3). Then, by (3), for any $k \geq 1$ we deduce

$$X_{n+k} = \sum_{j=0}^{k-1} \rho_j \epsilon_{n+k-j} + \sum_{j=1}^{\infty} \nu_j X_{n+1-j} \tag{45}$$

where $\rho_j = \sum_{i=1}^{k} \varphi_i$, $j = 1..k$ and $\nu_j = \sum_{i=1}^{k} \psi_{j+k-i} \varphi_i$, $j = 1, 2, ...$ Consequent to Lemma 2, $X_{n+k} \in P_n^Y$ and $a_j = \nu_j$, $j = 1..n$, is tha

Corollary 9 *For any A MA(, q) process X_n such ha (1 -(3 hold, for any $k \geq 1$, $n \leq \infty$ and $\alpha \in (0, 2)$, $\alpha \neq 1$, here exis s a unique CO P coefficien se* **a** *such ha (i - (v of emma 1 are sa is ed. In par icular, a_j solves*

$$\sum_{i=1}^{n} a_i \lambda_{i,t} = \sum_{i=1}^{n} \nu_i \lambda_{i,t} + \sum_{i=0}^{t} \nu_{n+1+i} \left(\sum_{j=0}^{\infty} \pi_{j+t-1} \pi_{n+1+j}^{<\alpha-1>} \right) + \sum_{i=t+1}^{\infty} \nu_{n+1+i} \left(\sum_{j=0}^{\infty} \pi_j \pi_{n+1+j}^{<\alpha-1>} \right), \tag{47}$$

where $\forall t = 1..n$,

$$\lambda_{i,t} = \pi_i \sum_{j=1-t}^{n-t} \pi_j^{<\alpha-1>} + \sum_{j=n+1}^{\infty} \pi_j^{<\alpha-1>} \pi_{j-i} \tag{48}$$

$$\nu_j = \sum_{i=1}^{k} \psi_{j+k-i} \varphi_i$$

and $\pi_j = 0$, $j < 0$.

Proof. Subsequent to remark 1 of Lemma 2, $\forall t = 1..n$, $[X_{n+k} - \mathbf{a}'\mathbf{X} | X_{n+1-t}]_\alpha = 0$ implies (47), and (48) is immediate due to (44) and (45). The unique coefficient set **a** follows from traditional methods for solving linear systems of equations. ■

emark 1: e may infer from (47) and (48) that for large n, the truncated predictor $X^*_{n+k} = \sum_{j=1}^{k-1} \psi_j X^*_{n+k-j} + \sum_{j=k}^{n+k-1} \psi_j X_{n+k-j}$ will be close to optimal. Of course, conversely the truncated MDLP will be approximately unbiased for large n.

Solving the implied n-equation system in (47) will be difficult even for low order processses. However, existing algorithmic techniques that employ recursive prediction residuals can be easily extended to the stable laws. For the following derivations, we require some compact notation, in addition to stable-representations detailed in previous sections. Define the recursive prediction residual $e_k = X_{k+1} - \hat{X}_{k+1}$, $k = 0...n-1$, where $\hat{X}_1 = 0$ by convention. Additionally, for each horizon $h \in \mathfrak{N}$, define the real-valued sequence $\{\theta_{n,i}^{h)}\}_{i=1}^{n}$

which proves the result. ■

emark 1: The substantive distinction between between the classical "residuals" Hilbert-space algorithm and the above stable-law procedure lies in the non-symmetry of the α-orthogonality condition, $[e_i, e_j]_\alpha$. Indeed, by construction, $[X_{i+1} - \hat{X}_{i+1}, X_{i+1-j}]_\alpha$ $= 0$ for every $j = 1, 2, ...i,$

and arbitrary $n \in \{1, 10, 50, 100\}$ we plot the differential

$$is\ \ (X_{n+k} - \mathbf{a}_m'\mathbf{X}) - \ \ is\ \ (X_{n+k} - \mathbf{a}_c'\mathbf{X}) \tag{53}$$

over the computed grid, and report the extrema

$$\min_{\alpha,\theta)\subseteq[1.01,1.99]\times[.01,.99]} \{\ \ is\ \ (X_{n+k} - \mathbf{a}_m'\mathbf{X}) - \ \ is\ \ (X_{n+k} - \mathbf{a}_c'\mathbf{X})\} \tag{54}$$

$$\max_{\alpha,\theta)\subseteq[1.01,1.99]\times[.01,.99]} \{\ \ is\ \ (X_{n+k} - \mathbf{a}_m'\mathbf{X}) - \ \ is\ \ (X_{n+k} - \mathbf{a}_c'\mathbf{X})\}.$$

Additionally, in order to demonstrate the bias in the canonical expected prediction error of the MDLP, we plot the dispersion of the conditional expectations of the MDLP prediction error,

$$\tilde{\ }_{MDLP} = \sum_{j=1}^{n} |z_j + \theta z_{j-1}|^{\alpha} + \left(\frac{|\theta + \phi|^{\alpha}}{1 - |\phi|^{\alpha}}\right) z_n, \tag{55}$$

where $z_j = \sum_{i=0}^{j} \phi^{j-i}[(a_{c,i} - \phi a_{c,i-1}) - (a_{m,i} - \phi a_{m,i-1})]$. Of course, by Corollary 4 the COLP is unbiased, hence $\tilde{\ }_{COLP} = 0$, thus $\tilde{\ }_{MDLP}$ also serves as the differential. Finally, for each $n \in [1, 100]$ and $\phi \in \{0, .7\}$ we calculate the number of grid points for which

$$|\ \ is\ \ (X_{n+k} - \mathbf{a}$$

the resulting prediction error bias, a problem classically associated with efficient, biased predictors/estimators.

In the $MA(1)$

of the criterion function, the unbiased predictor is universally easier to compute than existing methods. The result is a dominant method for deriving optimal linear forecasts of highly volatile time-series.

References

Brockwell, P.J. and .A. Davis, 1983, ecursive Prediction and Exact Likelihood Determination for Gaussian Processes, Technical eport 65, Dept. of Statistics, Colorado State University.

Brockwell, P.J. and .A. Davis, 1987, Time Series: Theory and Methods (Springer: New York).

Cambanis, S. and I. Fakhre-Zakeri, On Prediction of Heavy-Tailed Autoregressive Sequences: Forward Versus everse Time, Theory and Probability Applications 39, 217-233.

Cambanis, S. and G. Miller, 1981, Linear Problems in th Order and Stable Processes, SIAM Journal of Mathematics 41, 43-69.

Cambanis, S and A. . Soltani, 1982, Prediction of Stable Processes: Spectral and Moving Average epresentations, Tech. eport 11, Center for Stochastic Processes (Univ. of N. Carolina, Chapel Hill).

Cambanis, S. and . u, 1992, Multiple egression on Stable Vectors, Journal of Multivariate Analysis 41, 243-272.

Caner, Mehmet, 1998, Tests for Cointegration with Infinite Variance Errors, Journal of Econometrics 86, 155-175.

Cheng, B.N. and S.T. achev, 1995, Multivariate Stable Futures Prices, Mathematical Finance 5, 133-153.

Cioczek, . and M. Taqqu, 1994, Does Asymptotic Linearity of the egression

Knight, K., 1993, Estimation in Dynamic Linear egression Models with Infinite Variance Errors, Econometric Theory 9, 570-588.

Kokoszka, Piotr and M.S. Taqqu, 1996, Infinite Variance Stable Moving Averages with Long Memory, Journal of Econometrics 73, 79-99.

Mandlebrot, B., 1963, The Variation of Certain Specualtive Prices, Journal of Business 36, 394-419.

unde, alf, 1997, The Asymptotic Null Distribution of the Box-Pierce Q-statistic for andom Variable with Infinite Variance: An Application to German Stock e- turns, Journal of Econometrics 78, 205-216.

Samorodnitsky, G. and M.S. Taqqu, 1989, Conditional Moments of Stable andom Variables, Stochastic Processes.

Samorodnitsky, G. and M.S. Taqqu, 1994, Stable Non-Gaussian andom Processes (Chapman and Hall, New York).

Soltani, A. and . Moeanaddin, 1994, On Dispersion of Stable andom Variables and its Applications in the Prediction of Multivariate Stable Processes, Journal of Applied Probability 31, 691-699.

Stuck, B. ., 1978, Minimum Dispersion Linear Filtering of Scalar Symmetric Stable Processes, IEEE Trans. Aut. Cont. 23, 507 - 509.

Zolotarev, V.M., 1971, One Dimension Stable Distributions (American Mathematical Society, Providence).