Systems Analysis of the Labor Market

Working Paper No. 13

Two Stage Least Squares and Nonlinear Systems

By
Harry H. Kelejian

February 1969
Two Stage Least Squares and Nonlinear Systems

1. Introduction

In a recent article, Eisenpress and Greenstadt (1966) concluded that the two stage least squares (henceforth 2SLS) procedure may not be applicable to a system of nonlinear structural equations. The reason for this, they suggested, is that the reduced-form equations corresponding to such a system will not, in general, be linear in the structural disturbances. The equations considered by Eisenpress and Greenstadt were nonlinear in both the structural parameters and the endogenous variables. However, their conclusion suggests that any system of equations for which the corresponding reduced-form equations are nonlinear in the structural disturbances does not lend itself to estimation by the 2SLS procedure.

The purpose of this paper is to demonstrate that structural equations which are linear in the parameters but contain regressors which are nonlinear functions of endogenous and predetermined variables can be consistently estimated by the 2SLS procedure. In particular, it is shown that the consistency of the 2SLS estimates does not depend upon the condition that the reduced-form equations be linear in the structural disturbances. Further, it is shown that if the functional forms

* This article is based upon a research study prepared under a contract with the Office of Manpower Policy, Evaluation, and Research, U.S. Department of Labor, under the authority of the Manpower Development and Training Act, and jointly sponsored by the Departments of Commerce and Health, Education, and Welfare and the National Science Foundation. Researchers undertaking such projects are encouraged to express their own judgment freely. Therefore, points of view or opinion stated in this document do not necessarily represent the official position or policy of these agencies.

** I would like to thank Stephen Goldfeld of Princeton and James Ramsey of Michigan State for helpful comments on an earlier draft. They are not, of course, responsible for any shortcomings of this paper.

1 For a discussion of some of the properties relating to systems containing regressors which are nonlinear functions of endogenous variables see Fisher (1965, 1966, pp. 127-162).
of the reduced-form equations are not known and, therefore, approximated by
polynomials, the polynomials must be of the same degree if the 2SLS estimates
are to be consistent. Finally, it is shown that 2SLS instruments lack a certain
invariance property with respect to nonlinear transformations.

2. The Model

Consider the first structural equation of an \( n \) equation system

\[
y_{1t} = X(t)B_1 + F(t)B_2 + \epsilon_{1t}, \; t = 1, \ldots, T,
\]

where \( y_{1t} \) is the \( t \)th observation on the dependent variable; \( X(t) \) is a \( 1 \times n \)
vector of observations at time \( t \) on \( n \) predetermined regressors: \( X(t) = (x_{1t}, \ldots, x_{nt}) \);
\( B_1 \) is an \( n \times 1 \) vector of parameters; \( F(t) \) is a \( 1 \times k \) vector of observations at
time \( t \) on \( k \) functions. Each of these functions is assumed to depend upon at least
one endogenous variable of the system and an arbitrary number of predetermined
variables—e.g., \( F(t) = (f_{1t}, \ldots, f_{kt}) \), where, in general, \( f_{it} = f_{i}(y_{2t}, \ldots, y_{nt}, x_{1t}, \ldots, x_{nt}, z_{1t}, \ldots, z_{pt}) \) where \( y_{2t}, \ldots, y_{nt} \) and \( z_{1t}, \ldots, z_{pt} \) are, respectively, the \( t \)th
observations on the remaining dependent and predetermined variables of the system.
\( B_2 \) is a \( k \times 1 \) vector of parameters, and \( \epsilon_{1t} \) is the \( t \)th disturbance term.

We do not assume that the dependent variables of the system appear in
the same functional form in all of the equations. For example, one of the
regressors in (1) may be \( (y_{2t} y_{3t}) \) while \( \exp(y_{2t} + y_{3t}) \) may be a regressor in
other structural equations. All functions appearing in the system that contain
at least one of the dependent variables are considered as endogenous variables.

The disturbance term and the predetermined variables are assumed to be
generated by stochastic processes with finite moments. We assume \( \mathbb{E}[\epsilon_{1t} | X(t), Z(t)] = 0 \), \( \mathbb{E}[\epsilon_{1t}^2 | X(t), Z(t)] = \sigma_1^2 \) and \( \mathbb{E}[\epsilon_{1t} \epsilon_{1s} | Z(t)] = 0 \), \( t \neq s \), where \( Z(t) = (z_{1t}, \ldots, z_{pt}) \).

Further, we assume that the processes generating \( X(t), Z(t), \) and \( \epsilon_{1t} \) are such
that the sample moments converge in probability to the population moments. We also assume that the functions $f_{it}$, $i = 1, \ldots, k$, have finite range. Finally, we assume that equation (1) is identified with respect to zero restrictions. The problem, of course, is that the elements of $F(t)$ are expected to be correlated with the disturbance term.

3. The Reduced Form Equations

Consider the functions $f_{it}$, $i = 1, \ldots, k$, appearing as regressors in (1). Each of these functions can be considered as a random variable. Further, the assumptions outlined above imply that each of these functions has finite first and second moments. Therefore, because the mathematical expectation of one variable conditional upon a set of other variables is, in general, a function of those conditioning variables, we have

$$E[f_{it} | X(t), Z(t)] = h_{it}, \ i = 1, \ldots, k; \ t = 1, \ldots, T,$$

where $h_{it}$ is a function of the elements of $X(t)$ and $Z(t)$ — e.g. $h_{it} = h_{i}(X(t), Z(t))$. It is clear from (2) that each function $f_{it}$ can be expressed as

$$f_{it} = h_{it} + v_{it}$$

where $v_{it}$ is a stochastic element such that $E[v_{it} | X(t), Z(t)] = 0$—see Wold (1960, pp. 445-6). The equations given in (3) are defined as the reduced-form equations for the functions $f_{it}$, $i = 1, \ldots, k$ in terms of the elements of $X(t)$ and $Z(t)$.

It should be evident that the equations given in (3) do not represent the solution of the system for the functions $f_{it}$ in terms of the elements of $X(t)$ and $Z(t)$, and linear combinations of the structural disturbance terms. For instance,  

---

assume that the solution of the system for $f_{jt}$ in terms of the elements of $X(t)$ and $Z(t)$, and the structural disturbances is given by

$$f_{jt} = g_j(X(t), Z(t), e(t)),$$

where $e(t) = (e_{1t}, \ldots, e_{mt})$ where $e_{2t}, \ldots, e_{mt}$ are the disturbance terms in the remaining $m-1$ structural equations. Because the dependent variables do not appear in the same functional form in all equations, the functions $g_j$ will, in general, be nonlinear. Then the $j$th equation in (3), $i = j$, is obtained by assuming

$$E[g_j(X(t), Z(t), e(t))|X(t), Z(t)] = h_{jt} = g_j(X(t), Z(t), 0).$$

Clearly, an additive linear function of the disturbance terms is not the only function of the disturbance terms which has a mean conditional upon given values of the elements of $X(t)$ and $Z(t)$.

An example may help to clarify and extend the above argument. Consider the explicit but simplified version of the original system:

$$y_{1t} = b_{1t}X_t + \epsilon_{1t}$$

$$y_{2t} = b_{2t}Z_t + b_3 \exp(y_{1t}) + \epsilon_{2t},$$

where the disturbances $\epsilon_{it}$ ($i = 1, 2$) are normally distributed with means zero, variances $\sigma_i^2$, and covariance $\sigma_{12} \neq 0$. Assume that each $\epsilon_{it}$ is not autocorrelated and, further, is independent of $X_t$ and $Z_t$.

The solution of (6) and (7) for $y_{2t}$ in terms of the predetermined variables and the disturbances is

$$y_{2t} = b_{2t}Z_t + b_3 \exp(b_{1t}X_t) \exp(\epsilon_{1t}) + \epsilon_{2t}.$$  

Then, because $E[\exp(\epsilon_{1t})|X_t, Z_t] = \exp(\sigma_1^2/2)$ we note that $\exp(\epsilon_{1t})$ can be expressed as
(9) \( \exp(\varepsilon_{1t}) = \exp(\sigma_1^2/2) + \varepsilon_{3t} \),

where \( E[\varepsilon_{3t}|x_t, z_t] = 0 \). Substituting (9) into (8) we obtain the reduced-form equation for \( y_{2t} \)

(10) \( y_{2t} = b_2 x_t + b_4 \exp(b_1 x_t) + v_t \),

where \( b_4 = b_3 \exp(c_1^2/2) \), and the reduced-form disturbance \( v_t = \varepsilon_{2t} + b_3 \exp(b_1 x_t) \varepsilon_{3t} \). It is clear that \( E[v_t|x_t, z_t] = 0 \).

In comparing (10) with (8), we see that the deterministic part of the reduced-form equation obtained by setting \( v_t = 0 \) cannot be derived from (8) by setting the structural disturbances \( \varepsilon_{1t} \) and \( \varepsilon_{2t} \) equal to zero. In brief, the reduced-form equation for \( y_{2t} \) is not a solution of the system. A related point is that the reduced-form disturbance, \( v_t \), is not a linear function of the structural disturbances. Indeed, it is clear that \( v_t \) is heteroskedastic with respect to \( x_t \) even though the structural disturbances are homoskedastic. Finally, it should be noted that, in general, if the structural disturbances of an econometric model are assumed to be uncorrelated rather than independently distributed over time, the reduced-form disturbances will be autocorrelated since non-linear functions of uncorrelated variables are generally correlated. Thus, the properties of the reduced form disturbances should not be inferred from those of the structural disturbances.

Returning to the generalized system in (3) another point to note concerning the reduced-form equations is that if, say, \( f_{2t} = G(f_{1t}) \) where \( G \) is nonlinear, the reduced-form equation for \( f_{2t} \) is not given by \( f_{2t} = G(h_{1t} + v_{1t}) \). The relationships involved in this case are

(11) \( E[G(h_{1t} + v_{1t})|X(t), Z(t)] = h_{2t} + G(h_{1t}) \)

since the expected value of a function is not the function of expected values.
A final point concerning an implicit assumption underlying the reduced-form equations in (3) should be noted. Because nonlinear functions of the dependent variables appear as regressors in the structural equations, the reduced-form equations may not be unique. That is, multiple solutions may exist and so more than one value of each function such as \( f_{jt} \) may correspond to a given set of values of the elements of \( X(t) \) and \( Z(t) \). Restrictions concerning the range of the variables involved may rule out some of these solutions. However, it may be the case that the restrictions involved are not sufficient to rule out all but one solution. If this is the case, we assume that all the observations on the functions \( f_{jt} \), \( j = 1, \ldots, k, \ t = 1, \ldots, T \), have been generated by a single solution. The alternative is to assume that we know, a priori, the solutions from which the various observations have been generated. The assumption of a single generating solution may appear very restrictive. However, it should be noted that such an assumption is required in all conditional predictions of the endogenous variables of a nonlinear system.

4. Estimation by 2SLS

Substituting the matrix representation of the reduced-form equation of (3) into the structural equation (1) we obtain

\[
Y_{1t} = X(t)B_1 + H(t)B_2 + \varepsilon_{1t}, \quad t = 1, \ldots, T,
\]

where \( H(t) \) is the \( 1 \times k \) vector whose \( i^{th} \) element is \( h_{1t} \), and \( \varepsilon_{1t} = V(t)B_2 + \varepsilon_{1t} \), where \( V(t) \) is the \( 1 \times k \) vector whose \( i^{th} \) element is \( v_{1t} \). It follows from (3) and (1) that \( E[\varepsilon_{1t}|X(t), Z(t)] = 0 \). We now note a result derived in Kelejain (1969).

---

Footnote 3: For example, if \( f_{1t} = \pm ax_{1t} + v_{1t} \) we assume that all the observations on \( f_{1t} \) have been generated by either \( f_{1t} = ax_{1t} + v_{1t} \) or \( f_{1t} = -ax_{1t} + v_{1t} \). For a further discussion of this problem see Goldfeld and Quandt (1965, pp. 8-10).
In particular, if equation (1) is identified with respect to zero restrictions, the elements of \( \mathbf{X}(t) \) and \( \mathbf{H}(t) \) are \textit{linearly independent}. Thus, if \( \mathbf{H}(t) \) were known, (12) could be regarded as a linear regression model relating \( y_{1t} \) to the elements of \( \mathbf{X}(t) \) and \( \mathbf{H}(t) \). The assumptions underlying (1) then imply that the ordinary least squares estimates of \( B_1 \) and \( B_2 \) defined by (12) are consistent.

The least squares procedure, however, cannot be applied to (12) because, in general, the elements of \( \mathbf{H}(t) \) will not be known. Indeed, it should be clear from equations (2)-(5) that the elements of \( \mathbf{H}(t) \) will generally be nonlinear functions of the elements of \( \mathbf{X}(t) \) and \( \mathbf{Z}(t) \) which involve \( B_1 \) and \( B_2 \) as well as the other structural parameters of the system.

Furthermore, the difficulties involved in deriving the expected value operations described in (3) suggest that even the functional forms of the elements of \( \mathbf{H}(t) \) will not be known. We will now show, however, that a slight modification of the usual two-step procedure involved in 2SLS yields consistent estimates of the structural parameters, \( B_1 \) and \( B_2 \).

Let the \( i \)th element of \( \mathbf{H}(t) \), \( h_{it} \), be expressed as

\[
(13) \quad h_{it} = p_{d_i}^{di} + \epsilon_{d_i}^{di} \\
= M_{d_i}^{d_i} \gamma_{i}^{d_i} + \epsilon_{d_i}^{di}, \quad i = 1, \ldots, k; \quad t = 1, \ldots, T, 
\]

where \( p_{d_i}^{di} \) is a polynomial of degree \( d_i \) in the elements of \( \mathbf{X}(t) \) and \( \mathbf{Z}(t) \); \( \epsilon_{d_i}^{di} \) is the corresponding error in the approximation; \( M_{d_i}^{d_i} \) is the row vector containing the \( t \)th observation on the powers of \( \mathbf{X}(t) \) and \( \mathbf{Z}(t) \) appearing in \( p_{d_i}^{di} \), and \( \gamma_{i}^{d_i} \) is the corresponding vector of parameters. Without loss of generality, we assume that the parameters of the polynomial, \( \gamma_{i}^{d_i} \), are such that each element of \( M_{d_i}^{d_i} \)
is uncorrelated with $d_{it}^4$. Substituting (13) into (3) yields, what might be termed, the operational reduced-form equations

$$f_{it} = M_{it}^d \gamma_i^d + \psi_{it}^d, \quad i = 1, \ldots, k; \ t = 1, \ldots, T,$$

where $\psi_{it}^d = R_{it}^d + v_{it}^d$. Because the elements of $M_{it}^d$ are uncorrelated with both $R_{it}^d$ and $v_{it}^d$, the assumptions underlying the original model (1) imply that the least squares estimate of $\gamma_i^d$ obtained by regressing $f_{it}$ on the elements of $M_{it}^d$ are consistent. Let $\hat{\gamma}_{it}^d$ be the estimate of $\gamma_i^d$ at time $t$. Then the polynomials and the residuals corresponding to (14) are estimated, at time $t$, by $\hat{\gamma}_{it}^d = \hat{M}_{it}^d \hat{\gamma}_{it}^d$ and $\hat{\psi}_{it}^d = \hat{f}_{it}^d - \hat{\psi}_{it}^d$. Now let $\hat{F}(t) = (\hat{\psi}_{it}^d, \ldots, \hat{\psi}_{kt}^d)$ and $\hat{\psi}(t) = (\hat{\psi}_{it}^d, \ldots, \hat{\psi}_{kt}^d)$. Then from (14) and (1) we see that $F(t)$ can be expressed as

$$F(t) = \hat{F}(t) + \hat{\psi}(t).$$

We turn now in the usual manner to the second step of the 2SLS procedure. In particular, substituting (15) into (1) we obtain

$$\gamma_{it} = X(t)B_1 + P(t)B_2 + [\epsilon_{it} + \hat{\psi}(t)B_2], \ t = 1, \ldots, T.$$  

The assumption $E[\epsilon_{it}X(t), Z(t)] = 0$ implies that the elements of $X(t)$ are uncorrelated with $\epsilon_{it}$. Further, since $\text{Plim} \ \gamma_i^d = \gamma_i^d$, we see from (13) that the estimates of $\hat{F}(t)$ may be regarded as functions of the elements of $X(t)$ and $Z(t)$ alone. Thus, they are also uncorrelated with $\epsilon_{it}$. It is evident, then, that the estimates of $B_1$ and $B_2$ obtained by regressing $\gamma_{it}$ on the elements of $X(t)$ and $\hat{F}(t)$ are consistent if these elements are linearly independent and are uncorrelated with the elements of $\hat{\psi}(t)$.

---

Footnote: To see that this assumption is not binding, let the $T$ observations on $M_{it}^d$ be given by the matrix $M$. Let the $T$ observations on $R_{it}^d$ be given by the vector $R$. Now $1_{Tt} \lambda = \text{Plim} \ T^{-1}(M'M)^{-1} M'R$. If the remainder is not correlated with the elements of $M_{it}^d$, $\lambda \neq 0$. Then $R$ can be expressed as $R = M\lambda + Q$, where $\text{Plim} \ T^{-1}M'Q = 0$. Let the $i$th element of $Q$ be $q_{it}$, then the right-hand side of (13) can be expressed as $M_{it}^d \gamma_i^d + R_{it}^d \lambda + q_{it} = M_{it}^d \gamma_i^d + q_{it}$, Q.E.D.
Since the degree of each polynomial is at least one, i.e., \( d_i \geq 1 \), the elements of \( X(t) \) appear as regressors in each of the regression relationships given in (14). Therefore, since the regressors in a least squares regression are orthogonal to the estimated residuals, each element of \( X(t) \) is orthogonal to each element of \( \hat{y}(t) \). However, unless all the polynomials in (14) are of the same degree, i.e., \( d_i = d_o \), \( i = 1, \ldots, k \), each element of \( \hat{P}(t) \) will not be orthogonal to each element of \( \hat{y}(t) \). The reason for this is that unless \( d_i = d_o \), \( i = 1, \ldots, k \), the same set of regressors will not correspond to each of the \( k \) regression models given in (14). Therefore, if \( d_j > d_i \), the regressors corresponding to the \( i \) regression model in (14) will be a subset of the regressors corresponding to the \( j \)th regression model. Hence, the \( j \)th element of \( \hat{y}(t) \) will be orthogonal to the \( i \)th element of \( \hat{P}(t) \), but the \( i \)th element of \( \hat{y}(t) \), \( \hat{y}_{it}^{d_i} \), will not be orthogonal to the \( j \)th element of \( \hat{P}(t) \), \( \hat{P}_{jt}^{d_j} \). Furthermore, \( \hat{y}_{it}^{d_i} \) and \( \hat{P}_{jt}^{d_j} \) will, in general, be correlated unless the polynomial approximations in (13) to the reduced-form functions, \( h_{it} \), are perfect. To see this note first that since \( \hat{\gamma}_i^{d_i} \) is a consistent estimate of \( \gamma_i \), \( \hat{P}_{jt}^{d_j} \) and \( \hat{y}_{it}^{d_i} \) are consistent estimates of \( P_{jt}^{d_j} \) and \( y_{it}^{d_i} = R_{it}^{d_i} + v_{it} \). The condition \( E[v_{it} \mid X(t), Z(t)] = 0 \) implies that \( P_{jt}^{d_j} \) and \( v_{it} \) are uncorrelated; however, there are no conditions insuring a lack of correlation between \( P_{jt}^{d_j} \) and \( R_{it}^{d_i} \). Therefore, the 2SLS estimates of \( B_1 \) and \( B_2 \) obtained from (16) by regressing \( y_{it} \) on \( X(t) \) and \( \hat{P}(t) \) will not be consistent.

We therefore assume that \( d_i = d_o \), \( i = 1, \ldots, k \), where \( d_o \) is chosen to accomodate the highest order of the reduced-form functions \( h_{it} \), \( i = 1, \ldots, k \). Thus, the regressors of (16) are uncorrelated with \( \epsilon_{it} \) and are orthogonal to each element of \( \hat{y}(t) \).
We also assume that \( d_o \) is large enough so that the elements of \( X(t) \) and \( \hat{P}(t) \) are linearly independent. The existence of such a \( d_o \) follows from the assumption that the basic model (1) is identified subject to zero restrictions. That is such an assumption implies that the elements of \( X(t) \) and \( H(t) \) in (12) are linearly independent—see Kelejian (1969). Since the approximation to \( H(t) \), namely \( \hat{P}(t) \), improves as \( d_o \) increases, there clearly exists a \( d_o \) such that the elements of \( X(t) \) and \( \hat{P}(t) \) are linearly independent.

Returning again to (16) we note that subject to the above conditions, the 2SLS estimates of \( B_1 \) and \( B_2 \) obtained by regressing \( y_{1t} \) on \( X(t) \) and \( \hat{P}(t) \) are consistent. Let the \( T \) observations on \( X(t) \) and \( \hat{P}(t) \) be given by the matrices \( X \) and \( \hat{P} \). Then it is not difficult to show that the asymptotic variance-covariance matrix corresponding to the 2SLS estimates of \( B_1 \) and \( B_2 \) is

\[
\Gamma = \frac{1}{T} \sum_{t=1}^{T} \text{Plim} \left[(X \hat{P})' (X \hat{P})\right]^{-1}.
\]

(17) \[ \Gamma = \frac{1}{T} \sum_{t=1}^{T} \text{Plim} \left[(X \hat{P})' (X \hat{P})\right]^{-1}. \]

A consistent estimate of \( \Gamma \) is

\[
\hat{\Gamma} = S^2 \left[(X \hat{P})' (X \hat{P})\right]^{-1}
\]

where \( S^2 = \frac{1}{T-k-n} \sum_{t=1}^{T} (y_{1t} - X(t) \hat{B}_1 - \hat{F}(t) \hat{B}_2)^2. \)

---

5 Concerning this point, see the corresponding derivation for linear systems given by Goldberger (1964, pp. 329-333).
5. Lack of Invariance in 2SLS Instruments

Consider again the last \( k \) regressors described in the original model (1).

\[
(19) \quad f_{it} = f_i(y_{it}, \ldots, y_{mt}, z_{1t}, \ldots, z_{pt}), \quad i = 1, \ldots, k.
\]

The 2SLS procedure described above is one in which each of these regressors is considered as a new random variable. That is, the functional relationships between these regressors and the dependent variables of the system, \( y_{2t}, \ldots, y_{mt} \), are entirely ignored. For instance, in comparing (1) with (16) we see that the 2SLS instruments in \( \hat{F}(t) \) are simply estimates of the deterministic parts of the elements of \( F(t) \) - i.e., \( \hat{F}_{it} = \hat{f}_{it} \). \(^6\) If each \( f_{it} \) were considered as a function of the dependent variables, \( y_{2t}, \ldots, y_{mt} \), one might consider the 2SLS procedure using the instruments \( \hat{f}_{it} = f_i(y_{it}, \ldots, z_{pt}) \), \( i = 1, \ldots, k \), where \( y_{it} = E[y_{it} | X(t), Z(t)] \). We show now that the use of the instruments \( \hat{f}_{it} \) in the 2SLS procedure leads to inconsistent estimates of the structural parameters \( B_1 \) and \( B_2 \).

The functions in (19), \( f_i, \quad i = 1, \ldots, k \), are assumed to be nonlinear. Because the mathematical expectation of a nonlinear function is not, in general, the function of the expectations we see, recalling (2),

\[
(20) \quad E[f_{it} | X(t), Z(t)] = h_{it} \neq \hat{f}_{it}.
\]

It follows from (20) and (3) that

\[
(21) \quad f_{it} = \hat{f}_{it} + \nu_{it} + [h_{it} - \hat{f}_{it}].
\]

Now \( \hat{f}_{it} \) is a function of the elements of \( X(t) \) and \( Z(t) \) alone and is, therefore, consistent.

\(^6\) By "deterministic" we mean that part which is uncorrelated with the disturbance term \( \epsilon_{it} \). We omit reference to \( d_i \) in \( \hat{F}_{it} \) because of our assumption \( d_i = d_0 \), \( i = 1, \ldots, k \).
uncorrelated with $v_{it}$. However, in general

(22) $Ef_{it}^*[h_{it} - f_{it}^*] \not\rightarrow 0$.

Therefore, if $f_{it}^*$ or a consistent estimate of $f_{it}^*$, is used as the 2SLS instrument for $f_{it}$, the instrument will be correlated with the disturbance term (since it contains $h_{it} - f_{it}^*$) and hence the resulting estimates will be inconsistent.

---

7 The reader can convince himself of (22) by working through the example in which $f_{it} = \exp(y_{2t})$ and $y_{2t} = X(t)c_1 + Z(t)c_2 + q_t$ where $c_1$ and $c_2$ are parameter vectors and $q_t$ is a normally distributed variable which is independent of $X(t)$ and $Z(t)$, and $E_{\bar{t}}c = 0$, $E_{\bar{t}}^2 = \sigma^2$, and $E_{t}q_{s} = 0$ if $t \neq s$. 
Bibliography


