A note on some properties of STUR processes

Gawon Yoon[*]

Pusan National University

Abstract

This note presents possibly hitherto unnoticed, or only implicitly discussed, properties of the stochastic unit root process developed in Granger and Swanson (1997) and Leybourne, McCabe, and Tremayne (1996).

Keywords: stochastic unit root

---

1. Introduction

This note presents some properties of stochastic unit root [STUR] processes developed by Granger and Swanson (1997), Leybourne, McCabe, and Tremayne (1996), Leybourne, McCabe, and Mills (1996), and McCabe and Tremayne (1995). Specifically, it examines if STUR processes have stationary distributions and estimates their tail indices. It also shows that the estimation method employed by Granger and Swanson does not yield consistent estimates and that taking differences cannot make the STUR processes (weakly) stationary. These properties shed some light on the previous simulation results in Granger and Swanson (1997) and Gonzalo and Lee (1998). Some of the properties presented here are already known; however, it appears that most are hitherto unnoticed or only implicitly discussed in the literature. It is hoped that these properties will contribute to better understanding of this important and flexible class of models.

2. Main results

Among alternative formulations of STUR processes[1], the one used in Granger and Swanson (1997), GS hereafter, will be studied for its simplicity. The same notations in GS are used to avoid unnecessary confusions. GS consider the following process:

---

[1] See for instance Leybourne *et al*. (1996) and Leybourne *et al*. (1996).

$$x_t = a_t x_{t-1} + \varepsilon_t \tag{1}$$

$t = 1, 2, \ldots$ where $\varepsilon_t \sim i.i.d\left(0, \ \sigma_\varepsilon^2\right)$ and

$$a_t = \exp\left(\alpha_t\right) \tag{2}$$

$\alpha_t \sim N\left(m, \ \sigma_\alpha^2\right)$ with a power spectrum $g_\alpha\left(\omega\right)$. x_0 is assumed to be a constant. Assume also

that α_t is governed by an AR(1) process

$$\alpha_t = \mu + \rho\alpha_{t-1} + \eta_t \tag{3}$$

where $\left|\rho\right| < 1$ and $\eta_t \sim i.i.d.N\left(0, \ \sigma_\eta^2\right)$ is independent of ε_t. GS put additional condition that

$$E\left(a_t\right) = 1 \tag{4}$$

and call the process STURA.[2] Thus, x_t has a unit root only on average and can be mildly

explosive for some periods. When $E\left(a_t\right) > 1$, x_t is an explosive stochastic root process in

Granger (2000). Equation (4) implies that

$$m + \frac{1}{2}\sigma_\alpha^2 = 0. \tag{5}$$

It follows that $m < 0$, given that $\sigma_\alpha^2 > 0$.

A multiplicative stochastic process of the form $x_{t+1} = A_{t+1}x_t + b_{t+1}$ is studied in depth in

Vervaat (1979), Brandt (1986) and Horst (2001), among others. Brandt (1986) shows that under the

---

[2] Another process, STURB, assumes instead that $m + \frac{1}{2} \times \frac{1}{2\pi} g_\alpha\left(0\right) = 0$.

assumption of strict stationarity and ergodicity of $(A_{t+1}, \; b_{t+1})$, x_t has a stationary distribution if

$$-\infty \le E\left(\ln |A_1|\right) < 0 \tag{6}$$

and

$$E\left(\ln |b_1|\right)^+ < \infty \tag{7}$$

where $\omega^+ = max(\omega, \; 0)$. He also shows that the solution to the multiplicative stochastic

equation is unique. From now on, (7) is assumed to hold for the STUR process in (1). It is easy to

show that (6) is satisfied as well because $E\left(\ln(a_1)\right) = E(\alpha_1) = m < 0$. Hence, the first property

of STUR follows:

Property 1: STUR processes are strictly stationary and have unique stationary distributions.

(1) is not a nonstationary process and it is somewhat confusing to call it a stochastic *unit root*

process. It is straightforward to show that the STUR processes are not weakly stationary, using the

well-known result in Nicholls and Quinn (1982): $Var(a_t) + (Ea_t)^2 < 1$. Further, next property

shows that STUR processes do not have any finite moments as $t \to \infty$. A positive exponent, λ,

that satisfies $E\left(a_t^\lambda\right) = 1$ is called a tail index, so that the tails of the limiting distribution are

asymptotic to a power law, $Prob\left(|x_t| > x\right) \approx cx^{-\lambda}$. x and c are constants. A distribution with a

3

lower tail index has a higher probability of extreme events. Much progress has been made recently

on the tail behavior of financial data; see for instance discussions in Mikosch (2001). For the STUR

processes considered here, their tail index is 1, almost trivially by definition, see (4). Therefore, no

finite moments exist for STUR, which is next property.

Property 2: STUR processes have a tail index 1 and therefore they have no finite moments.

Hence, STUR processes are not weakly stationary. A tail index 1 corresponds to the tail behavior of

a Cauchy distribution. The STUR processes have very heavy-tails and their autocorrelation

functions have non-degenerate limit distributions, as shown in Davis and Mikosch (1998).

Therefore, the autocorrelation functions should be used with care. Further, autocorrelations are not

particularly useful for nonlinear processes as already noted in Granger and Teräsvirta (1993). As an

illustration, $T = 1000$ observations are generated and split into half.[3] Figure 1 shows the data

series, its first difference and estimated autocorrelations functions for the two subsamples T_1 and

---

[3] The following values are used in simulating a STUR process by GAUSS: $\eta_t \sim i.i.N\left(0, \quad 0.001\right)$,

$\rho = 0.6$, $\mu = -0.0002$, and $\sigma_\varepsilon^2 = 1$, with $\alpha_0 = 0$ and $x_0 = 1$. Initial 500 observations are

discarded.

4

T_2 of the same size. Clearly, the sample autocorrelations functions are very different for the two subperiods.

How does this tail index property of STUR stand with the empirical results reported in GS? GS use various monthly data series for 1955:1 ~ 1987:12 and estimate the STUR model in (1), (2), and (3) without imposing the restriction of (4). GS employ an approximate maximum likelihood estimation method. Their estimation results are reproduced here in table 1. Without imposing the condition (4), it is not difficult to show that the tail index of STUR is

$\lambda = -\dfrac{2m}{\sigma_\alpha^2} = -\dfrac{2\mu(1+\rho)}{\sigma_\eta^2} > 0$. The second column of table 1 shows that $\hat{\mu}$ is negative only for stock price index, three interest rates, and unemployment rate. Hence, $\hat{\lambda}$ is reported only for them

at the last column, with $\hat{\lambda} = -\dfrac{2\hat{\mu}(1+\hat{\rho})}{\hat{\sigma}_\eta^2}$. Stock price index has by far the highest estimated $\hat{\sigma}_\eta^2$.

Its tail index is less than 1, implying that no moments exist. For other series, their tail indices are much bigger.[4] Interestingly, according to the simulation results in Gonzalo and Lee (1998), the standard Dickey-Fuller tests have power against a STUR alternative for these four series only out of ten series listed in table 1. However, the sample size should be very large for the tests to have a

---

[4] For unemployment rate, $\hat{\sigma}_\eta^2 = 0.00001$ is used.

power.[5]

GS also report that parameters of STUR processes are "usually fairly imprecisely estimated."

[p. 48] They use an approximate maximum likelihood method. However, the method requires the

existence of moments of high enough order for the estimates to be consistent and asymptotically

normal; see Laroque and Salanié (1994). Thus, the following property follows:

Property 3: Approximate maximum likelihood estimates of STUR are not consistent, nor

asymptotically normally distributed.

Further, GS use the approximate maximum likelihood estimates to generate predictions for forecast

evaluation exercise. Thus, the forecast performance of STUR is handicapped by estimates that are

not consistent.[6] Notice also that GS use mean squared prediction errors as a forecasting evaluation

criterion. However, given that no moments exist for the STUR process, a more natural choice for

---

[5] The standard ADF tests are not useful in distinguishing the null of difference stationarity against

STUR alternatives; see McCabe and Smith (1998).

[6] GS find that the forecast performance of STUR is not dominating other methods, for instance a

time-varying parameter model estimated by the Kalman filter, at least for short-term forecasts.

forecast evaluation would be conditional median. It is known to be optimal in the sense defined in

Hall *et al.* (2002).

It follows from the property 2 that any differences of x_t, $\Delta^d x_t$, $d \in \mathbb{N}$, do not have finite

moments, either. Therefore, a stochastic unit root process x_t cannot be made into weak stationarity

by taking differences.

Property 4: STUR processes are not weakly stationary to any order of differencing.

This property was already noted by GS when they write that "STUR processes are clearly not

difference stationary, in theory." [p.45] See also Leybourne, McCabe, and Tremayne (1996), who

call such a process $I_t(1)$. Yoon (2002) dubs such a process $I(\infty)$. Gonzalo and Lee (1998) show

that standard cointegration tests tend to find cointegrating relationship when they are applied to two

independent STUR processes. Since STUR processes are not difference stationary, the standard

cointegration testing procedures are not applicable.

3. Conclusions

This note presents some properties of stochastic unit root processes. While some of the properties

7

are already known, it appears that others are hitherto unnoticed, or only implicitly discussed in the literature. Of course, there could be other interesting properties associated with STUR processes that are not mentioned here. It is hoped that these properties will contribute to the better understanding of this important class of models.

Table 1: STUR estimation results

Series	$\hat{\mu}$	$\hat{\rho}$	$\hat{\sigma}_\eta^2$	$\hat{\lambda}$
Money stock- M2	0.0026			
Stock price index	-0.0036	0.025	0.0586	0.13
Long-term interest rate	-0.0012	-0.302	0.0001	16.75
Short-term interest rate	-0.0060	-0.328	0.0004	20.16
Medium-term interest rate	-0.0040	-0.169	0.0003	22.16
Industrial production	0.0002			
Nominal inventory	0.0002			
Unemployment rate	-0.0009	-0.376	0.0000	112.32
Consumer price index	0.0052			
Producer price index	0.0048			

The monthly data series are used in Granger and Swanson (1997) for 1955:1 ~ 1987:12. ^ denotes

estimated values reproduced from Granger and Swanson, except for the last column. Granger and

Swanson use an approximate maximum likelihood estimation method. For unemployment rate,

$\hat{\sigma}_\eta^2 = 0.00001$ is used to calculate $\hat{\lambda}$. For the last three columns, results are reported only for the

series with $\hat{\mu} < 0$.

References

Brandt, Andreas, 1986, The stochastic equation $Y_{t+1} = A_t Y_t + B_t$ with stationary coefficients, Advances in Applied Probability, 18, 211-220

Davis, Richard A. and Thomas Mikosch, 1998, The sample autocorrelations of heavy-tailed processes with applications to ARCH, Annals of Statistics, 26, 2049-2080

Gonzalo, Jesus and Tae-Hwy Lee, 1998, Pitfalls in testing for long run relationships, Journal of Econometrics, 86, 129-154

Granger, Clive W.J., 2000, Overview of nonlinear time series specifications in economics, Chung-Hua Series of Lectures by Invited Eminent Economists, No. 26, The Institute of Economics, Academia Sinica, Taipei, Taiwan

Granger, Clive W.J. and Norman R. Swanson, 1997, An introduction to stochastic unit-root processes, Journal of Econometrics, 80, 35-62

Granger, C.W.J. and Timo Teräsvirta, 1993, Modeling Nonlinear Economic Relationships, Oxford University Press: Oxford

Hall, Peter, Liang Peng, and Qiwei Yao, 2002, Prediction and nonparametric estimation for time series with heavy tails, Journal of Time Series Analysis, 23, 313-331

Horst, Ulrich, 2001, The stochastic equation $Y_{t+1} = A_t Y_t + B_t$ with non-stationary coefficients,

Journal of Applied Probability, 38, 80-94

Laroque, Guy and Bernard Salanié, 1994, Estimating the canonical disequilibrium model: Asymptotic theory and finite sample properties, Journal of Econometrics, 62, 165-210

Leybourne, Stephen J., Brendan P.M. McCabe, and Andy R. Tremayne, 1996, Can economic time series be differenced to stationarity? Journal of Business and Economic Statistics, 14, 435-446

Leybourne, Stephen J., Brendan P.M. McCabe, and Terence C, Mills, 1996, Randomized unit root processes for modelling and forecasting financial time series: Theory and applications, Journal of Forecasting, 15, 253-270

McCabe, Brendan P.M. and Richard J. Smith, 1998, The power of some tests for difference Stationarity under local heteroskedastic integration, Journal of the American Statistical Association, 93, 751-761

McCabe, Brendan M.P. and Andy R. Tremayne, 1995, Testing a time series for difference stationarity, Annals of Statistics, 23, 1015-1028

Mikosch, Thomas, 2001, Modeling dependence and tails of financial time series, manuscript

Nicholls, D.F. and B.G. Quinn, 1982, Random Coefficient Autoregressive Models: An Introduction, Lecture Notes in Statistics 11, New York: Springer-Verlag

Vervaat, Wim, 1979, On a stochastic difference equation and a representation of non-negative

infinitely divisible random variables, Advances in Applied Probability, 11, 750-783

Yoon, Gawon, 2002, An introduction to $I(\infty)$ processes, manuscript

Caption for figure

Figure 1: Simulated STUR process and estimated autocorrelation functions

1000 observations are generated from a stochastic unit root process using values listed in footnote 3.

The autocorrelation functions are estimated for two equally divided subsamples, T_1 and T_2, with

500 observations each.