İSTATİSTİK: JOURNAL OF THE TURKISH STATISTICAL ASSOCIATION
Vol. 6, No. 1, January 2013, pp. 1–9
ISSN 1300-4077 | 13 | 1 | 1 | 9

İSTATİSTİK

RANDOM WALK OR MEAN REVERSION? EMPIRICAL EVIDENCE FROM THE CRUDE OIL MARKET

Delson Chikobvu*
Department of Mathematical Statistics and Actuarial Science
University of Free State, Box 339, Bloemfontein 9300, South Africa
and
Knowledge Chinhamu†
School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal
Private Bag X54001, Durban 4000, South Africa

Abstract : The paper investigates and gives an update on the empirical evidence for or against the random walk theory in the crude oil market. The paper investigates whether there are periods when crude oil prices follow the random walk process and periods when they deviate from the random walk theory (mean reverting). Various studies often give conflicting results for the same study period. Some computations conducted over the period of 2000-2005 lead to inconclusive results (Geman, 2007), suggesting that more work remains to be done in this period and beyond. It is imperative to revisit mean reversion and random walk in the context of crude oil as it has serious implication on modeling crude oil prices. In this paper, a Garch model with time-varying properties is applied to capture periods when the random walk theory may be true and periods when it may be false. This study concludes the existence of mean reversion for crude oil price over the period 1980 to 1994 and a random walk as of February 1994 to the end of the study period in 2010. Prior to February 1994 some models, like arima models, might have been valid. Beyond this period, models need to recognize the random walk in crude oil returns.

Key words: Efficient market, Garch, Crude oil price, Random walk theory, Time varying parameters.
History: Submitted: 15 May 2012; Revised: 4 July 2012; Accepted: 9 August 2012

1. Introduction

The thrust of the study is to investigate whether crude oil prices are mean reverting or follow a random walk process. Crude oil presents an interesting case because it is the fundamental driver of most economic activities in the world. Crude oil is vital in many industries and of great importance to the maintenance of an industrialized modern economy. Higher crude oil prices have a direct impact on macroeconomic variables such as inflation, Gross Domestic Product (GDP), investments, recessions, and other macroeconomic variables (Cheong, 2009). Crude oil returns are related to the global financial markets, including contracts, opinions, risk management and other related financial derivatives.

It is thus important to investigate whether oil price predictions can be done with accuracy or not. Forecasting crude oil future prices remains one of the biggest challenges facing econometricians and statisticians. Some researchers found that crude oil prices follow a random walk, implying that tomorrow's expected crude oil prices should be the same as today's value. If crude oil prices follow a random walk, then prices would be very difficult, if not impossible to predict. It is imperative to revisit mean reversion and random walk in the context of crude oil as it has serious implication on modeling crude oil prices.

* Corresponding author. E-mail address: chikobvu@ufs.ac.za (D. Chikobvu)

† E-mail address: chinhamu@ukzn.ac.za (K. Chinhamu)

This study investigates whether crude oil prices are mean reverting or follow a random walk process at all time periods. The Augmented Dickey-Fuller (ADF) tests and the Garch model with time-varying properties approach are used to investigate mean reversion and random walk process in crude oil prices. Whilst the techniques adopted in this paper may be standard in empirical finance, the approach as presented, with time varying parameters, has not been applied to crude oil to the best of our knowledge. There are still gaps and updating that needs to be done in providing the empirical evidence in the crude oil market.

The rest of this paper is organised as follows. Section 2 gives an overview of the random walk model and related literature. Section 3 describes the empirical methods used in this study. In Section 4, the data and results are reported and discussed. Section 5 provides a conclusion.

2. Literature Review

Geman (2007) used the following model to investigate the statistical properties of crude oil prices:

$$p_t = \varphi p_{t-1} + \epsilon_t$$

to check whether the coefficient φ is significantly different from 1, where p_t is the log of crude oil price. The null hypothesis is the existence of a unit root (i.e. $\varphi = 1$). A p value smaller than 0.05 would reject the null hypothesis with a confidence level higher than 0.95. The bigger the p value, the more the random walk is validated. Geman (2007) used the West Texas International (WTI) oil spot prices over the period January 1994 to 2004. The author got an ADF p value of 0.651 for spot prices of oil prices for the period January 1994 to October 2004. The result rejects the mean reversion assumption over the whole period and confirms that log crude oil price follows a random walk during this period.

However, Geman (2007) noted that a mean reversion pattern of crude oil prices prevails for a shorter period from 1994 to 2000, and it changes into random walk as of 2000. The author used three state variable models which incorporate stochastic volatility.

Bessembinder et al. (1995) analysed the relation between crude oil price levels and slope of the futures term structure defined by the difference between a long maturity future contract and its first nearby. Assuming that future prices are unbiased expectations (under the real probability measure) of future spot oil prices, an inverse relation between prices and the slope constitutes evidence that investors expect mean reversion in spot prices, as it implies lower expected future spot prices when prices rise. The authors concluded the existence of mean reversion of oil price over the period 1982-1991, however the same computations conducted over the period 2000-2005 leads to inconclusive results (Geman, 2007). Thus more work remains to be done in this period and beyond.

Bernard et al. (2008) argue that research on crude oil price dynamics for modeling and forecasting has brought out several unsettled issues. Although statistical support is claimed for various models of price paths, many of the competing models differ with respect to their fundamental properties. One such property is mean reversion. Pindyck (1999) says that unit root tests are inconclusive in the analysis of real prices observed on a yearly basis. The author concludes that due to the persistent characteristic of crude oil price, a very long and practically unavailable series is required to perform reliable tests.

3. Methodology

A simple model for log returns

We define the natural logarithmic return (simple log return) of crude oil at time t as:

$$r_t = \log(P_t/P_{t-1}) = \log(P_t) - \log(P_{t-1}) = p_t - p_{t-1}$$

where P_t is the price of crude oil at time t.

The simplest model which can be used to test for the random walk is the simple autoregressive (AR(1)) model, namely:

$$r_t = \beta_0 + \beta_1 r_{t-1} + \epsilon_t \tag{3.1}$$

where $r_t = p_t - p_{t-1}$, is the log return of crude oil price, β_0 and β_1 are the parameters that need to be estimated and $\epsilon_t \sim \text{IID}(0,\sigma^2)$, $p_t = \log(P_t)$ is the natural logarithm of the price of crude oil at time t. If the crude oil price follows a random walk, $\beta_1 = 0$ and so

$$p_t = \beta_0 + p_{t-1} + \epsilon_t \tag{3.2}$$

the random walk with drift parameter β_0.

The natural logarithmic transformation reduces the impact of heteroscedasticity that may be present when you have large data sets with high frequency. The transformation also ensures that predicted crude oil price is positive when anti-logs are taken. The model however does not cater for changing volatility.

Three versions of the random walk are distinguished by Cambell, Lo and MacKinley (1997) and also cited in Jefferis and Smith (2005: p.59) which depend on the assumptions of the error term, namely ϵ_t. Under the first model, the error terms are independently and identically distributed with a zero mean and constant variance, denoted by $\epsilon_t \sim \text{IID}(0,\sigma^2)$. In the second model, the error terms are independent but not identically distributed, which allows for unconditional heteroscedasticity in the ϵ_t or $\epsilon_t \sim \text{NID}(0,\sigma_t^2)$. The problem of heterogeneously distributed processes is relevant, since crude oil prices have been found to display heteroscedasticity. In the third random walk model, the error terms are uncorrelated and neither independent nor identically distributed as mentioned in the research of Jefferis and Smith (2005). This paper will also focus on the third model, with volatilities changing over time.

Equation (3.1) has constant parameters and the error terms are assumed to follow the usual classical assumptions. With financial markets, the assumption of constant variance may be inappropriate as empirical evidence frequently finds that returns have a variance which changes systematically. Equation (3.1) cannot readily capture gradual deviations towards/from the random walk over successive observations.

Garch approach with time varying parameters

Emerson et al. (1997) and Zalewska-Mitura and Hall (1999) have developed, using a Garch approach, a test with time-varying parameters which detects changes towards/from the random walk where the error process does not have a full set of NIID properties. The model checks for changes towards/from the random walk and allows the error process to deviate from the property of being normally independent and identically distributed. The test does 3 things: first, it checks for the random walk; second, it detects changes from/towards the random walk, and third, it will operate with a stochastic series for which the error process might not have a full set of NIID properties.

The test is based on the following set of equations to constitute the model

$$r_t = \beta_{0t} + \beta_{1t} r_{t-1} + \delta\sigma_t^2 + \mu_t \tag{3.3}$$

$$\mu_t | \psi_{t-1} \sim N(0,\sigma_t^2) \tag{3.4}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \mu_{t-1}^2 + \gamma_1 \sigma_{t-1}^2 \tag{3.5}$$

$$\beta_{0t} = \beta_{0t-1} + v_{0t}\,; \quad v_{0t} \sim N(0,\sigma_{\beta_0}^2) \tag{3.6}$$

$$\beta_{1t} = \beta_{1t-1} + v_{1t}\,; \quad v_{1t} \sim N(0,\sigma_{\beta_1}^2) \tag{3.7}$$

in which σ_t^2 is the conditional variance of the error term ϵ_t, a Garch (1,1) model. ψ_t is the information set available at time t. α_0, α_1 and γ_1 are parameters needed to model the changing volatility. This model has three important characteristics. First, the intercept, β_{0t} and slope coefficient β_{1t} can change through time. However, the special cases where either or both of these are constant are also included. Secondly, this model incorporates an error process in which the variance changes systematically over time. Thirdly, the mean of the log return depends on its conditional variance (level of risk).The basic insight is that risk-averse investors will require compensation for holding a risky asset such as crude oil.

A maximum likelihood search procedure with a standard Kalman filter is used to estimate the model with equation (3.3), the measurement equation, and the set of equations given by (3.5), (3.6) and (3.7), the state equations. The Kalman filter sequentially updates coefficient estimates and generates the set of β_{it}'s, $i = 0,1$ and $t = 1 \cdots T$ and their standard errors. If the crude oil log returns follow a random walk with no drift, then a $100(1-\alpha)\%$ confidence band for each of β_{0t} and β_{1t} should contain zero. The method will be applied to crude oil prices in this paper. The focus of this study is to find out if crude oil prices follow a random walk process or is mean reverting.

ϵ_t and v_{it} $i = 0,1$ are playing the role of both disturbances and state variables. This is somewhat unusual. The Kalman filter is being used in the context of a model with Garch errors. The Kalman filter in its present form is not operable. This is because past values of error terms are unobservable. Nevertheless we may proceed on the basis that the model can be treated as though it were conditionally Gaussian, and we will refer to the Kalman filter as being quasi-optimal(Harvey et al, 1992) (Moonis and Shar, 2002).

Extending the model

Zalewska-Mitura and Hall (1999) have an extension to the model in the previous section.

The test is based on the following set of equations:

$$r_t = \beta_{0t} + \sum_{i=1}^{p} \beta_{it} r_{t-i} + \delta\sigma_t^2 + \epsilon_t \tag{3.8}$$

$$\epsilon_t|\psi_{t-1} \sim N(0,\sigma_t^2) \tag{3.9}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + \gamma_1\sigma_{t-1}^2 \tag{3.10}$$

$$\beta_{it} = \beta_{it-1} + v_{it}\ ;\quad v_{it} \sim N(0,\sigma_{\beta_i}^2) \tag{3.11}$$

Such a model can again be modeled using the standard Kalman filter. The parameters required to estimate time-paths of β_{it}'s and δ, $\alpha_0, \alpha_1, \gamma_1$ and all p values can be found by maximizing the Likelihood Function. If the series r_t is a random walk, the $100(1-\alpha)\%$ confidence bands for each of the β_{it}'s must contain zero.

Reasons for modeling Garch effects

Like many econometric time series, crude oil prices exhibits periods of unusually large volatility followed by periods of relative tranquility. In such instances, the assumption of a constant variance (homoscedasticity) is in appropriate. The volatility of crude oil prices is not constant over time, a property called heteroscedasticity. Modelling such varying variances involves Garch modelling. A distinguishing feature of a Garch model is that the error variance may be correlated over time because of phenomenon of volatility clustering. Thus it is appropriate to use the Garch model which incorporates an error process in which the variance is allowed to change systematically over time. Hence, the model can detect gradual departures from the random walk (weak form efficiency) through time.

Building AR (p) models

An important step in the model identification process is to find the order of the autoregressive process for the log returns. There are three basic steps to follow to fit AR(p) models to time series data. These steps involve plotting the data, possibly transforming the data, identifying the dependence orders of the model, parameter estimation, and diagnosis and model choice. The Box Jenkins methodology using auto-correlations is used to identify the order of the model.

Model selection for ADF Tests

The lag order, in addition to a sample size can affect the finite sample behavior of the ADF test. Proper correction for the lag effect in implementing the ADF test is desirable. Because appropriate values for the ADF test can be easily computed with desirable accuracy from response surface equation for any sample size and leg length, the analysis should be useful in practical applications (Cheung and Lai, 1995). The number of the augmenting lags (p) is determined by minimizing the Schwarts Bayesian information Criterion (SBI) or minimizing the Alkaike Information Criterion(AIC). In this study the SBI is used and the software automatically selects the appropriate lag length and hence the model.

4. Empirical results

This section discusses data source and data analysis. The data is on crude oil prices. The section also discusses the results of the random walk process and the results of the ADF test. Lastly, results from the Garch model with time-varying parameters approach are also discussed.

The data used in this study is monthly crude oil price from January 1980 to September 2010 with a total of 369 observations and is quoted in US dollars. The data is a monthly crude spot price of European Brent. This data is from the World Bank and the International Monetary Fund (IMF) and is available from the following website: http://www.mongabay.com/commodities/price-charts/price-of-uk-brent-oil.html.

The data is used to form two sub-segments of data namely, January 1980 to January 1994 and February 1994 to September 2010 segments. Two data segments are used for ADF tests for reasons which will become obvious after observing the results from Garch modelling. The data series from 1980 to 2010 when analysed as a whole with ADF test, the results are somewhat different, for instance, from 1980 to 1994, crude oil price does not follow a random walk model, i.e. it is mean reverting. The data series is also transformed into monthly log returns series by taking the first difference in the logarithm of the prices to give the log returns.

The ADF for the data in the period 1980 to 2010 (Full Data Set)

The ADF test is used to test stationarity for the data set from 1980 to 2010. Conclusions are made in line with Geman's (2007) paper. The ADF test statistic for untransformed crude oil price is -2.946064 with p value of 0.1493. At 10% significance level, the null hypothesis of non-stationary (unit root) is not rejected implying that crude oil prices are non stationary. Non stationarity implies the random walk (German, 2007). Similar results are obtained for log crude oil price data from 1980 to 2010, the ADF test statistic is -2.414055 (p value 0.3716). The p value is greater than 10%, the null hypothesis of unit root is not rejected. This result implies that the log crude oil price is non stationary and thus a random walk.

The ADF test for the data in the period January 1980 to January 1994 (First Segment)

The ADF test statistic on untransformed crude oil price data for the period 1980 to 1994 is -3.599062 with p value of 0.0829. Since p value is less than 10%, the null hypothesis of non-stationary (unit root) is rejected implying that crude oil price is stationary over the period of the first segment. This result is rather surprising and is not consistent with the results of the whole data set ranging from 1980 to 2010. This result suggests that crude oil price is mean reverting over the period 1980 to 1994. The ADF test statistic for log crude oil price data (period 1980 to 1994) is -2.963231. The p value $= 0.1459$ which is more than 10% , meaning that the null hypothesis of unit root is not rejected and hence implying that the log crude oil price are thus a random walk process. To summarize the results thus far, the data series from 1980 to 2010 when analysed as a whole, the conclusion is that the crude oil price follows a random walk model. However a shorter period, called the first segment, the conclusion is that that crude oil price is mean reverting over the period 1980-1994. However if the data is log transformed over the same period, 1980-1994, the conclusion is that the log crude oil price follows a random walk.

The ADF test for the data in the period February 1994 to September 2010 (Second Segment)

The value of the ADF test statistic is -1.703747 (p value $= 0.4278$), that is the hypothesis of a stationarity is not rejected and the monthly crude oil price follows a simple random walk for the period February 1994 to September 2010.The ADF test statistic of -2.816595 (p value of 0.1931) for log crude oil is less negative than the critical value at 10% significant level, the null hypothesis of unit root is not rejected. Thus the log crude oil price is non stationary hence implying a random walk for log crude oil price. The conclusion is crude oil and log crude oil price is a random walk over the period 1994-2010. These results of the random walk or mean reversion seem to depend on the period under consideration and whether the data is log transformed or not. The behavior of a random walk is more pronounced in log crude oil price.The conclusion is log crude oil price is a random walk over the period 1980-1994 but for the untransformed data the conclusion is that crude oil price is mean reverting for the same period (1980-1994).

Results from the Garch model with time varying parameters.

The results of using the Garch model with time varying parameter are presented in this section. Figures 1, 2 and 3 present the results of the changes towards/from the random walk. The figures show the paths of the estimated β_{it}'s, $i = 0, 1, 2$ coefficient (see equation (3.8)) with their respective 95 per cent confidence bands. For the period 1980 to 2010, the best model using Box Jenkins methodology is an AR (2) model:

$$r_t = \beta_{0t} + \beta_{1t} r_{t-1} + \beta_{2t} r_{t-2} + \mu_t$$

Garch model with time varying parameters (period from 1980 to 2010)

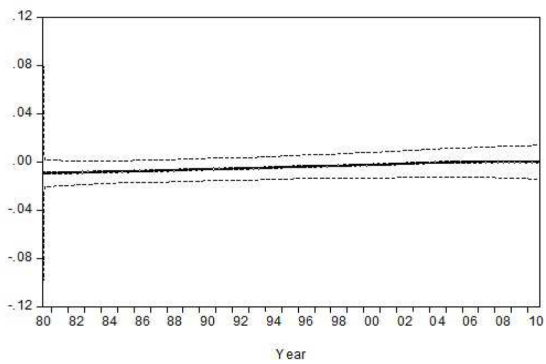

Figure 1. $\hat{\beta}_{0t}$, Drift parameter for crude oil price from 1980 to 2010. The estimates of $\hat{\beta}_{0t}$ are shown by a solid bold line and its confidence limits by dotted lines.

Consider Figure 1, which represents the results of the estimated drift parameter $\hat{\beta}_{0t}$ for the period 1980 to 2010. The estimate, $\hat{\beta}_{0t}$, has constant value of -0.0072 and is insignificantly different from zero considering its 95 percent confidence limit.

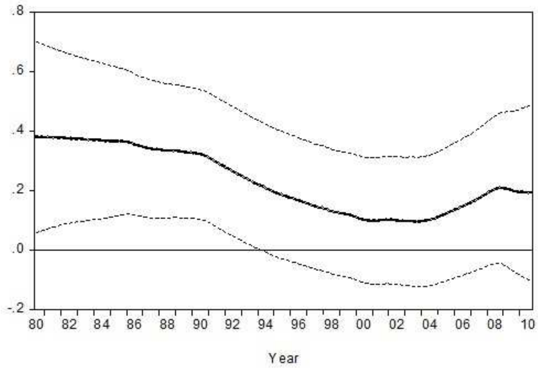

Figure 2. $\hat{\beta}_{1t}$ estimates for crude oil price from 1980 to 2010. The estimates of $\hat{\beta}_{1t}$ are shown by a solid bold line and its confidence limits by dotted lines.

Figure 2 shows the results of the parameter $\hat{\beta}_{1t}$ for the period 1980 to 2010. The estimate $\hat{\beta}_{1t}$ has an initial value of 0.39 and is significantly different from zero at 0.05 level. The magnitude of the estimated parameter gradually declines and first becomes insignificantlly different from zero in February 1994. The parameter remains insignificant for the rest of the period to end at a level of 0.19 in September 2010.

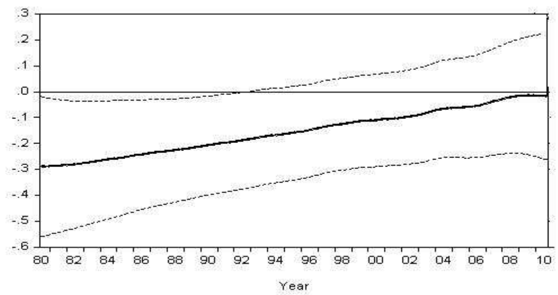

Figure 3. $\hat{\beta}_{2t}$ for crude oil price 1980 to 2010. The estimates of $\hat{\beta}_{2t}$ are shown by a solid bold line and its confidence limits by dotted lines.

Figure 3 shows the results of the parameter $\hat{\beta}_{2t}$ for the period 1980 to 2010. The estimate $\hat{\beta}_{2t}$ has an initial value of -0.29 and is significantly different from zero at 0.05 level. The magnitude of the estimated parameter gradually increase and first becomes insignificantly different from zero in March 1993. The parameter remains insignificant for the rest of the period to end at -0.02 in September 2010. Crude oil price follow the random walk from February 1994. Prior to the year 1994, the finding is that crude oil prices did not follow the random walk process i.e. crude oil price were mean reverting.

5. Conclusion

In this study, an attempt was made to determine whether crude oil price is mean reverting or a random walk process. Two approaches namely the Augmented Dickey-Fuller (ADF) test and the Garch model with time-varying properties are used. Before carrying out formal Augmented Dickey-Fuller (ADF) tests, the autocorrelation function (ACF) correlogram of crude oil price and log crude oil price are examined to investigate stationarity.

The untransformed data series from 1980 to 2010 shows evidence of a random walk process when using the ADF test, yet a shorter period (first segment) shows mean reversion for the period January 1980 to January 1994 according to the same test. The test also shows that crude oil price follows a random walk over the period February 1994 to 2010. Thus the results seem to depend on the period under consideration and this is rather puzzling. These results of the random walk or mean reversion also seem to depend on whether the data is log transformed or not. The behaviour of a random walk is more pronounced in log crude oil price. These results show that the ADF test approach has a limitation of depending on the period under consideration.

The Garch model with time-varying parameters approach shows the presence of mean reversion in log crude oil prices over the period January 1980 to January 1994. It shows a random walk as of February 1994. The conclusion of the ADF test seem to depend on the period under consideration. The cut off period of January 1994 for the first segment is suggested by the results of the Garch modelling approach. This approach does not depend on the period under consideration and can be deemed to be better than the ADF test in that sense. The Garch modelling approach confirms the ADF approach when the latter approach uses segmented data with log transformed data.

The results obtained in this paper are consistent with the results by Geman (2007) who concluded that the crude oil price follow a random walk for the period January 1999 to October 2004, using the ADF test. The result of the study also shows some similarity with Bessembinder et al. (1995), who confirm the existence of mean reversion over the period 1982 to 1991. The authors obtained inconclusive results over the period 2000 to 2005 and in this paper, a random walk prevails over that period.

This paper uses more current monthly data on crude oil prices up to September 2010. The Garch time-varying property approach used in the study produces results that are somewhat different to the ADF test. The results are similar only when the data is segmented after observing information from the Garch model. This study concludes the existence of mean reversion for crude oil prices over the period 1980 to 1994 and a random walk as of February 1994. The results also confirm a finding by Geman (2007) that the behaviour of a random walk is more pronounced when using log crude oil price. The ADF test using untransformed data shows that the crude oil price is mean reverting for the period January 1980 to February 1994 yet the log transformed data shows a random walk over the same period.

The paper recommends further study in using time-varying parameters approach to investigate mean reversion and random walk for asset prices of precious metals such as gold and platinum. Statisticians and econometricians should at least use the Garch approach before using the ADF tests to investigate whether prices are mean reverting or a random walk for a period under consideration, and avoid conflicting results which depend on size of the sample (period) under consideration.

References

[1] Bessembinder, H., Coughenour, J., Seguin, P. & Smoller, M. (1995) Mean reversion in equilibrium asset prices: Evidence from the futures term structure. Journal of Finance, 50: 361-375.

[2] Bernard, S.T., Khalaf, L., Kichian, M. & McMahon, S. (2008) Oil prices: Heavy Tails, Mean Reversion and the convenience Yield.
Available on line: http://ideas.repec.org/p/lvl/lagcr/0801.html

[3] Cheong, C.W. (2009) Modeling and forecasting crude markets using ARCH-type models. Energy policy, 2346-2355.

[4] Cheung, Y.W. and Lai, K.S. (1995) Log Order and critical values of the Augmented Dickey-Fuller Test. Journal of Business and Economic Statistics, 13(3): 277-280.

[5] Campbell, J.Y., Lo, A.W. & Mackinlay, A.C. (1997) The Econometrics of Financial Markets, Princeton: Princeton University Press.

[6] Emerson, R., Hall, S.G. & Zalaweska-Mitura, A. (1997) Evolving Market Efficiency with an application to some Bulgarian shares. Economics of planning, 30(1): 75-90.

[7] Geman, H. (2007) Mean Reversion versus Random Walk in Oil and Natural Gas Prices. United Kingdom: University of London.

[8] Harvey, H., Ruiz, E & Sentona, E. (1992) Unobserved component time series models with ARCH disturbances. Journal of Econometrics 52: 129-157.

[9] Jefferies, K. & Smith, G. (2005) The changing efficiency of African stock markets. South African Journal of Economics, 73: 54-67.

[10] Lo, A.W. & Mackinlay, A.C. (1999) A Non-Random Walk down Wall Street. Princeton: Princeton University Press.

[11] Moonis, S.A. & Shar, A. (2008) Teseting the time variation in beta in India. Journal of Emerging Markets Finance, 2(2): 163-180.

[12] Pindyck, R.S. (1999) The long run evolution of energy prices. The Energy Journal 20: 1-27.

[13] Zalewska-Mitura, A. & Hall, S.G.] (1999) Examining the first stages of market performance: a test for evolving market efficiency. Economics Letters, 64: 1-12.

[14] The mongabay website http://www.mongabay.com/commodities/price-charts/price-of-uk-brent-oil.html [Accessed 22July 2012]

1 306010 010091 >

İSTATİSTİK: JOURNAL OF THE TURKISH STATISTICAL ASSOCIATION İSTATİSTİK
Vol. 6, No. 1, January 2013, pp. 27–41
ISSN 1300-4077 | 13 | 1 | 27 | 41

SEMIPARAMETRIC INFERENCE AND BANDWIDTH CHOICE UNDER LONG MEMORY: EXPERIMENTAL EVIDENCE*

Uwe Hassler† and Maya Olivares
Goethe University Frankfurt

Abstract: The most widely used semiparametric estimators under fractional integration are variants of the local Whittle [LW] estimator. They are consistent for the long memory parameter d and follow a limiting normal distribution. Such properties require the bandwidth m to satisfy certain restrictions for the estimators to be "local" or semiparametric in large samples. Optimal rates for m are known and data-driven selection procedures have been proposed. A Monte Carlo study is conducted to compare the performance of the LW and the so-called exact LW estimators both in terms of experimental size when testing hypotheses about d and in terms of root mean squared error. In particular, the choice of the bandwidth is addressed. Further, competing approximations to limiting normality are compared.

Key words: Fractional integration, approximate normality, bandwidth selection.
History: Submitted: 11 June 2012; Revised: 14 September 2012; Accepted: 25 October 2012

1. Introduction

Persistence in the sense of slowly decaying autocorrelations is a stylized fact with many economic and financial time series, see Henry and Zaffaroni [12] for a survey. Such persistence is often called long memory. It can be captured by models that are fractionally integrated of order d, $I(d)$, with $0 < d < 1$, which extends the classical $I(0)/I(1)$ paradigm.

Two popular procedures to analyze long memory are the log-periodogram regression (GPH after Geweke and Porter-Hudak [4]), and the local Whittle estimator [LW], proposed by Künsch [18] and investigated by Robinson [22]. Both estimators are asymptotically normally distributed and consistent for $d \in (-0.5, 0.75)$, but the LW estimator is more efficient asymptotically. Still, LW is inconsistent for $d \geq 1$ and lacks asymptotic normality for $d > 3/4$, see Velasco [30]. Therefore, diverse methods have been proposed to improve its statistical properties. Well-known and easily implemented extensions include data differencing and periodogram tapering, see for example Hurvich and Ray [17], Hurvich and Chen [15] and Velasco [30]. However, Shimotsu and Phillips [28] argue that the first alternative requires prior knowledge of the degree of differencing, while the second one leads to an increase in the variance of the estimator. They propose a computationally more demanding variant of LW called Exact Local Whittle estimator [ELW]. It is consistent and follows the same limiting distribution as the LW estimator, however, it does so for a much larger parameter space. Alternatively, Abadir, Distaso and Giraitis [1] introduce a fully extended (or: nonstationarity-extended) version of the LW estimator. A thorough discussion of these LW-type estimators, and also of tapered LW estimators as well as wavelet-based competitors can be found in Faÿ, Moulines, Roueff and Taqqu [3].

* This paper is a revised and shortened version of a chapter from the PhD thesis of the second author. We thank an anonymous referee, Jörg Breitung, Dieter Nautz and Katsumi Shimotsu for helpful comments and suggestions. Financial support by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) through HA-3306/4-1 is gratefully acknowledged.

† Corresponding author: Goethe University Frankfurt, Statistics and Econometric Methods, Grüneburgplatz 1, D-60323 Frankfurt, Germany. Phone: +49 69 798 34762, Email: hassler@wiwi.uni-frankfurt.de.

In this paper, we focus on two variants of the LW procedure and compare the original LW and ELW estimators in an extensive Monte Carlo study. Several papers have studied the selection of the number of periodogram ordinates (bandwidth m) in relation to estimation bias and root mean squared error [RMSE], see for instance Robinson and Henry [24], Henry [10], and Nielsen and Frederiksen [20]. But it seems that size properties when testing for the true parameter value building on approximate normality have hardly been reported (an exception being the limited evidence in Hauser [9] and Hurvich and Chen [15, Table II] for just one bandwidth choice). For that reason we provide on the one hand experimental evidence not only on RMSE, but in particular on distortions of the nominal size (or coverage) as a function of the chosen bandwidth. Additionally, we compare deterministic with data-driven bandwidth selection rules. The latter ones have been proposed by Robinson [21], Delgado and Robinson [2] and Henry and Robinson [11]. On the other hand, we compare an alternative approximation of the test statistic ($\mathcal{R}^*$ from eq. (3.6)) with the usual asymptotic version ($\mathcal{R}$ from (3.4)), motivated by earlier findings in Hurvich and Chen [15]. We arrive at three relevant conclusions for empirical work that are summarized at the end of the paper.

The rest of the paper is organized as follows. Section 2 introduces the long memory model of fractional integration and the (E)LW estimator. In Section 3, we discuss approximations to the limiting normal distribution used for testing hypotheses about d, while simulation evidence is contained in Section 4. The final section summarizes our main findings.

2. Model and estimation

2.1. Fractional integration

The most widely used model to capture long memory is a fractionally integrated process $\{y_t, t \in \mathbb{Z}\}$, given by

$$(1-L)^d y_t = x_t\,, \quad -1 < d < 0.5\,, \tag{2.1}$$

where $(1-L)^d$ with the usual lag operator L is given by binomial expansion,

$$(1-L)^d = \sum_{j=0}^{\infty} \pi_{j,d} L^j\,, \quad \pi_{0,d} = 1,\ \pi_{j,d} = \frac{j-1-d}{j}\pi_{j-1,d}\,, \quad j \geq 1\,, \tag{2.2}$$

and $\{x_t\}$ is a purely stochastic, stationary process with short memory. More precisely, we assume that the spectral density of $\{x_t\}$ is bounded and bounded away from zero at frequency zero, such that the process is $I(0)$. In case of fractional integration one allows for non-integer values of $d \in \mathbb{R}$. Equation (2.1) defines a stationary process if and only if $d < 0.5$, see e.g. Granger and Joyeux [6] and Hosking [13].

In the time domain, the persistence of a fractionally integrated process is reflected by the behaviour of a hyperbolically decaying autocovariance sequence. For $0 < d < 0.5$, the autocovariances $\gamma_h = E(y_t y_{t+h})$ decay for a constant C_d depending on d so slowly,

$$\gamma_h \sim C_d\, h^{2d-1}, \quad h \to \infty\,, \tag{2.3}$$

that they are not summable, which characterizes long memory in the time domain:

$$\sum_{h=0}^{H} |\,\gamma_h\,| \to \infty, \quad H \to \infty\,.$$

Often, it is assumed that $\{x_t\}$ is an autoregressive moving-average process [ARMA], which we do not require here. For a fairly general sufficient condition on the short memory component $\{x_t\}$ that guarantees (2.3), see Hassler and Kokoszka [7, Coro. 2.1].

In the frequency domain, long memory translates into unboundedness of the spectral density at frequency zero. Particularly, it holds for $\{y_t\}$ with spectral density $f_y(\lambda)$ that

$$f_y(\lambda) \sim \lambda^{-2d} f_x(0), \quad \lambda \to 0\,, \tag{2.4}$$

where $f_x(\lambda)$ stands for the spectral density of the short memory component $\{x_t\}$. Hence, f_y is integrable for $d < 0.5$ notwithstanding the singularity at frequency zero.

Nonstationary fractionally integrated processes can be defined in terms of integer differences $(\Delta = 1 - L)$ for $0.5 \le d < 1.5$,

$$\Delta y_t = z_t \;\sim I(d-1)\,,$$

where $\{z_t,\, t \in \mathbb{Z}\}$ is $I(d-1)$ as defined in (2.1). Consequently,

$$y_t = y_0 + \sum_{j=1}^{t} z_j\,, \quad t = 1, \ldots, T\,, \tag{2.5}$$

is integrated of order d. Such processes have been labelled "type I" by Marinucci and Robinson [19], see also Robinson [23] for a discussion. Alternatively, many people work under the assumption of "type II" processes defined as

$$y_t = \mu + \sum_{j=0}^{t-1} \psi_{j,d}\, x_{t-j}\,, \quad t = 1, \ldots, T\,, \tag{2.6}$$

where $\psi_{j,d}$ are from the truncated expansion of Δ^{-d}, i.e. $\psi_{j,d} = \frac{j-1+d}{j}\psi_{j-1,d}$, and a constant μ is added to the process. The model from (2.6) can also be used for $d < 0.5$, although the process becomes stationary only asymptotically.

2.2. (Exact) Local Whittle [LW] estimation

Whittle [31, 32] suggested for stationary processes an approximation of the likelihood function in the frequency domain, which relies on the periodogram. Consider the discrete Fourier transform [DFT] $w_y(\cdot)$ of $\{y_t\}$ $(t = 1 \ldots T)$,

$$w_y(\lambda_j) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} y_t \exp\{i\lambda_j t\}\,, \quad i^2 = -1\,,$$

at the jth harmonic frequency $\lambda_j = \frac{2\pi j}{T}$. Then the periodogram simply is

$$I_y(\lambda_j) = |w_y(\lambda_j)|^2\,, \quad j = 1, \ldots, M = \left\lfloor \frac{T-1}{2} \right\rfloor\,, \tag{2.7}$$

where $\lfloor \cdot \rfloor$ stands for the floor operator. The log likelihood approximation in the frequency domain becomes

$$l_M(d) = -\sum_{j=1}^{M} \log f_y(\lambda_j) - \sum_{j=1}^{M} \frac{I_y(\lambda_j)}{f_y(\lambda_j)}\,.$$

Now, let us assume model (2.5). The LW estimator maximizes the log likelihood *locally* over a vicinity close to frequency 0, where the slope of f_y varies with d alone, see (2.4). To that end M is replaced by the bandwidth m. A crucial condition for the consistency of the estimator in the presence of short memory in $\{x_t\}$ is that the number of harmonic frequencies m used in the estimation must diverge more slowly than the sample size T:

$$\frac{1}{m} + \frac{m}{T} \to 0\,, \quad T \to \infty\,. \tag{2.8}$$

Replacing $f_y(\lambda_j) \sim G\lambda_j^{-2d}$ where $G = f_x(0)$, the negative local log likelihood becomes

$$-l_m(d) \approx \sum_{j=1}^{m}\left(\log G - 2d\log(\lambda_j) + \frac{I_y(\lambda_j)}{G\lambda_j^{-2d}}\right) =: Q_m(G,d) \,.$$

Concentrating G out yields $\widehat{G} = \frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2d} I_y(\lambda_j)$. Hence, the estimation of d requires minimizing

$$R(m,d) := \log\left\{\frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2d} I_y(\lambda_j)\right\} - \frac{2d}{m}\sum_{j=1}^{m}\log(\lambda_j)\,, \tag{2.9}$$

being the LW estimator thus defined as

$$\widehat{d}_{LW} = \arg\min R(m,d)\,.$$

Obviously, $\widehat{d}_{LW}$ crucially hinges on m as well as on the slope of $f_x(\lambda)$ at the origin, which is surpressed for notational convenience. Robinson [22] establishes limiting normality under (2.8) and further assumptions,

$$\sqrt{m}\,(\widehat{d}_{LW} - d) \Rightarrow \mathcal{N}\left(0, \frac{1}{4}\right)\,, \tag{2.10}$$

for $d \in (-0.5, 0.5)$ where "$\Rightarrow$" denotes convergence in distribution. Since the periodogram is shift invariant, the LW estimator is not affected by a mean μ different from zero and does not require an estimation of μ. Notice that the limiting variance is smaller than that of the famous semiparametric competitor, the log-periodogram regression by Geweke and Porter-Hudak [4]. Moreover, LW has been recommended since it is robust with respect to heteroskedasticity of a certain degree, see Robinson and Henry [24] and Shao and Wu [25]. Finally, Velasco [30] extended the results by Robinson [22] showing that the LW estimator is consistent for $d \in (-0.5, 1)$ and asymptotically normal for $d \in (-0.5, 0.75)$.

The issue of nonstationarity has been addressed more fully by Shimotsu and Phillips [28]. They propose to correct the DFT by adding a complementing term ensuring a valid approximation that holds for every value of d. This so-called exact LW procedure [ELW] implies replacing $\lambda_j^{2d} I_y(\lambda_j)$ in (2.9) by $I_{\Delta^d y}(\lambda_j)$, and it is valid if $\mu = 0$ in (2.6). For means different from zero, Shimotsu [27] suggests to demean $\{y_t\}$ with an appropriate estimator $\widehat{\mu}$, and to compute the exact LW estimator from the demeaned data. The objective function to be minimized becomes

$$R_E(m,d) := \log\left\{\frac{1}{m}\sum_{j=1}^{m} I_{\Delta^d(y-\widehat{\mu})}(\lambda_j)\right\} - \frac{2d}{m}\sum_{j=1}^{m}\log(\lambda_j)\,, \tag{2.11}$$

where $I_{\Delta^d(y-\widehat{\mu})}(\lambda_j)$ is the periodogram of $\{\Delta^d(y_t - \widehat{\mu})\}$. To determine the fractional differences, it is assumed that $\{y_t\}$ is given by a type II process as in (2.6). It turns out that the first sample observation y_1 is a reliable mean estimator in the case of large values of d, while the usual arithmetic mean $\overline{y}$ does a good job for small parameter values of d. Hence, Shimotsu [27] puts forward the following weighted estimator:

$$\begin{aligned} \widehat{\mu}(d) &= v(d)\,\overline{y} + (1 - v(d))\,y_1\,, \\ v(d) &= \begin{cases} 1\,, & d \leq 0.5 \\ \frac{1+\cos(4\pi d)}{2}\,, & 0.5 < d < 0.75 \\ 0\,, & d \geq 0.75 \end{cases}\,. \end{aligned}$$

To get a feasible procedure, he considers two steps. First, one determines an estimator of $\widehat{d}$ independent of μ in order to get an estimator of the constant: $\widehat{\mu} = \widehat{\mu}(\widehat{d})$. In a second step, the slope and Hessian of $R_E(m,d)$ are used to compute the feasible estimator (a MATLAB code is available from the homepage of K. Shimotsu):

$$\widehat{d}_{2ELW} = \widehat{d} - \frac{R'_E(m,\widehat{d})}{R''_E(m,\widehat{d})}.$$

Shimotsu [27, Theo. 3] shows that the two-step ELW estimator $\widehat{d}_{2ELW}$ is consistent and has the same limiting distribution as the LW and ELW estimators under $-0.5 < d < 2$.

3. Approximate inference

One goal of semiparametric inference is hypothesis testing about the long memory parameter d. Hypotheses of interest are $d \leq 0$ (short memory) vs. $d > 0$ (long memory), or $d \geq 0.5$ (nonstationarity) vs. $d < 0.5$ (stationarity). The tests are approximate in that they rely on limiting normality (as $m \to \infty$) of the appropriately standardized estimators.

3.1. Bandwidth selection

Robinson [22] proves limiting normality for the LW estimator under

$$m = T^{\alpha}, \quad 0 < \alpha < 0.8. \tag{3.1}$$

This rate is from Robinson [22, Assumption A4'] in the case $\{x_t\}$ is ARMA (for $\beta = 2$ in his notation); see Shimotsu and Phillips [28] for a slightly stronger assumption in the case of ELW estimation. In practice, the choice of the bandwidth m will be crucial for reliable inference. It balances a trade-off between variance and bias, see Henry and Robinson [11]. Thus, if m is chosen too large or too small, the outcome of the estimation may wrongly suggest a certain degree of persistence. To avoid such pitfalls, many empirical researchers typically opt for choosing a grid of bandwidth values and then plot the estimates against different values of m, see also Taqqu and Teverovsky [29] for graphical bandwidth selection. Ideally, one observes three regimes: With small values of m the estimates will display high variability, then the plot of the estimates should become approximately flat, while with further growing values of m the estimates may start to fall or to rise because of a bias due to a short memory component. In such an ideal situation one would choose m from the middle regime. In practice, however, such an ideal situation will rarely be encountered unless the sample size is very large.

As an alternative to graphical means, data-driven techniques for bandwidth choice have been proposed. They rely on a minimization of the asymptotic mean squared error and have been proposed by Delgado and Robinson [2], Henry and Robinson [11], Henry [10], Giraitis, Robinson and Samarov [5] and Hurvich and Deo [16]. While the latter two contributions concentrate on the log-periodogram regression estimator of d, Henry and Robinson [11] and Henry [10] derive an algorithm to obtain an optimal bandwidth for the LW estimator.

The approximative optimal spectral bandwidth derived heuristically in Henry and Robinson [11] must be iterated until convergence to an optimal bandwidth value is achieved. It is defined as

$$\begin{aligned} \widehat{d}^{(k)} &= \arg\min R(\widehat{m}^{(k)}; d)\,, \\ \widehat{m}^{(k+1)} &= \left(\frac{3T}{4\pi}\right)^{4/5} \left|\theta + \frac{\widehat{d}^{(k)}}{12}\right|^{-2/5}, \end{aligned} \tag{3.2}$$

with initial value $\widehat{m}^{(0)} = T^{0.8}$. Strictly speaking, the optimal rate of $T^{0.8}$ in (3.2) violates the condition in (3.1). In the case of the long memory representation defined in (2.4), if the first and second derivatives of $f_x(\lambda)$ exist, the parameter θ from (3.2) is defined as $\theta = f''_x(0)/2\,f_x(0)$, as shown by Delgado and Robinson [2]. The authors propose a simple feasible approximation for the unknown parameter θ, which is motivated by a Taylor expansion and consists in regressing the periodogram $I_y(\lambda_j)$ on the regressors $Z_{j\ell}(\widehat{d}^{(0)})$, $\ell = 0, 1, 2$:

$$I_y(\lambda_j) = \sum_{\ell=0}^{2} Z_{j\ell}(\widehat{d}^{(0)})\, \widetilde{\varphi}_\ell + \widetilde{\varepsilon}_j \quad j = 1, \ldots, \widehat{m}^{(0)},$$

where $Z_{j\ell}(d) = |1 - \exp\{i\lambda_j\}|^{-2d} \lambda_j^\ell/\ell!$. The estimates of $f_x(0)$ and $f_x''(0)$ are $\widetilde{\varphi}_0$ and $\widetilde{\varphi}_2$, respectively, so that the estimated parameter is given by

$$\widehat{\theta} = \frac{\widetilde{\varphi}_2}{2\widetilde{\varphi}_0}\,. \tag{3.3}$$

In principle, one could include the determination of $\widehat{\theta}$ as part of the kth iteration step, but Delgado and Robinson [2] advice against doing so, see also Henry [10].

3.2. Approximation of the asymptotic variance

Let now $\widehat{d}$ stand generically for the local Whittle estimator $\widehat{d}_{LW}$or the two-step mean-corrected version $\widehat{d}_{2ELW}$ of the ELW estimator by Shimotsu [27]. If we wish to test a null hypothesis about d_0, the asymptotic version of the test statistic becomes

$$\mathcal{R} = 2\sqrt{m}\,(\widehat{d} - d_0)\,, \tag{3.4}$$

which is compared with critical values from the standard normal distribution because of (2.10). In order to improve the size properties of the tests in finite samples, the bandwidth m in (3.4) can be replaced with an approximation m^* where

$$m^* = \sum_{j=1}^{m} \nu_j^2 \quad \text{with} \quad \nu_j = \log j - \frac{1}{m}\sum_{j=1}^{m} \log j\,, \tag{3.5}$$

because

$$\frac{m^*}{m} = 1 + O\left(\frac{\log^2 m}{m}\right) \to 1 \quad \text{as} \quad m \to \infty\,,$$

see Robinson [22, p.1645] or Hurvich and Beltrao [14, Lemma 1]. The rationale behind m^* stems from the Hessian of $R(m,d)$ from (2.9) evaluated at maximum likelihood,

$$\frac{\partial^2 R(m, \widehat{d})}{\partial d^2} = 4\,\frac{m^*}{m} + o_p(1)\,,$$

see Robinson [22, (4.10)] or Hurvich and Chen [15, p.163], and Shimotsu and Phillips [28, p.1916] for ELW. This motivates the usual maximum-likelihood approximation building on the Fisher information:

$$\sqrt{m}\,(\widehat{d} - d_0) \;\sim\; \mathcal{N}\left(0,\, \frac{m}{4\,m^*}\right).$$

The approximate test statistic $\mathcal{R}^*$ is therefore defined as

$$\mathcal{R}^* = 2\sqrt{m^*}\,(\widehat{d} - d_0)\,, \tag{3.6}$$

to be compared with standard normal percentiles.

Normalizing with m^* instead of m comes in naturally for the log-periodogram regression, too, due to studentizing the estimator with the regressor being $\log\left(4\sin^2(\lambda_j/2)\right) \sim 2\log\lambda_j$. In fact, this corresponds to the original proposal by Geweke and Porter-Hudak [4], which was found experimentally to outperform differing normalizations by Hassler, Marmol and Velasco [8, eqn. (7)]. For LW estimation, an asymptotically equivalent approximation of m building on $\log(2\sin(\lambda_j/2))$ instead of $\log j$ in (3.5) was advocated by Hurvich and Chen [15], while our asymptotically equivalent choice of m^* in (3.5) was used by Shimotsu [26].

In the next section we report experimental size properties of $\mathcal{R}$ and $\mathcal{R}^*$ under the null hypothesis that d_0 is the true value.

4. Experimental evidence

To investigate the finite-sample behaviour of the variants of the LW estimator, we now report the empirical size and RMSE from a simulation study for different bandwidth selection rules. These are deterministic as well as data-driven according to the iterative procedure described in (3.2).

The data generating process [DGP] is an ARFI$(1, d)$ model, where the short memory component $\{x_t\}$ is a stable AR(1) sequence,

$$x_t = a\, x_{t-1} + u_t\,, \tag{4.1}$$

with standard normal innovations $u_t \sim ii\mathcal{N}(0,1)$. We consider the following cases:

1. ARFI$(0, d)$, $d \in \{0, 0.45, 0.7\}$ and
2. ARFI$(1, 0)$ with $a = 0.5$.

ARFI$(0, d)$ denotes the case of fractionally integrated noise where $a = 0$. The scheme to generate fractional integration is throughout of type II, see (2.6), where we set $\mu = 0$ without loss of generalization. For each Monte Carlo DGP, 1000 replications with $T \in \{256, 512, 1024\}$ observations were performed. All computations were performed with MATLAB. To estimate d, we minimize over the interval $[-1; 3]$ using the routine "fminbnd".

From our experiments we present as empirical size $100\,\widehat{\alpha}$, where $\widehat{\alpha}$ is the relative frequency of rejection under the null at nominal level $\alpha_0 \in \{0.01, 0.05, 0.10\}$. Since $\widehat{\alpha}$ converges to α_0, the approximate 95% confidence interval is given by

$$\left[\widehat{\alpha} \pm 1.96\sqrt{\frac{\alpha_0(1-\alpha_0)}{1000}}\right] \quad \text{or} \quad \left[100\,\widehat{\alpha} \pm 1.96\sqrt{10\,\alpha_0(1-\alpha_0)}\right]. \tag{4.2}$$

The following table presents lengths of such intervals that allow to judge whether the percentages of rejections reported in the next two subsections are conformable with the corresponding nominal levels or not.

α_0	1%	5%	10%
CI from (4.2)	$[100\,\widehat{\alpha} \pm 0.617]$	$[100\,\widehat{\alpha} \pm 1.351]$	$[100\,\widehat{\alpha} \pm 1.859]$

In what follows, we highlight in bold face those experimental levels that are not significantly different from the nominal ones at the 95% level according to (4.2).

Table 1. White noise: ARFI(0,0)

		$T=256$			$T=512$			$T=1024$		
Test statistic	n.s.	$T^{0.55}$	$T^{0.65}$	$T^{0.75}$	$T^{0.55}$	$T^{0.65}$	$T^{0.75}$	$T^{0.55}$	$T^{0.65}$	$T^{0.75}$
	1%	6.7	4.5	3.4	3.9	3.2	2.3	3.8	2.4	**1.5**
$\mathcal{R}_{LW}$	5%	16.9	11.8	8.8	12.3	10.0	7.5	10.7	8.8	**6.0**
	10%	24.3	17.2	16.9	19.1	15.9	13.3	15.6	15.1	12.2
	1%	2.4	2.1	**1.5**	**1.6**	1.7	**1.1**	2.2	**1.2**	**1.1**
$\mathcal{R}^*_{LW}$	5%	7.8	6.7	**6.1**	**5.9**	**5.8**	**6.1**	6.6	**6.1**	**4.6**
	10%	14.4	12.1	**11.3**	**11.5**	**11.3**	**10.7**	**11.6**	**11.8**	**10.0**
RMSE		0.145	0.103	0.072	0.115	0.079	0.055	0.089	0.062	0.038
	1%	5.9	4.6	3.6	4.5	3.4	2.5	3.8	2.6	2.5
$\mathcal{R}_{2ELW}$	5%	15.6	11.9	9.1	12.6	10.7	7.7	10.6	8.6	9.3
	10%	24.5	18.7	18.5	19.5	16.0	14.5	16.7	15.9	14.4
	1%	1.8	2.3	1.9	1.9	1.9	**1.4**	2.3	**1.6**	1.7
$\mathcal{R}^*_{2ELW}$	5%	6.6	6.9	6.6	6.6	**6.2**	6.4	6.8	6.5	7.2
	10%	14.1	12.5	11.9	**11.5**	**11.7**	**11.8**	12.2	**11.8**	12.6
RMSE		0.148	0.108	0.076	0.118	0.085	0.058	0.092	0.067	0.042

Root mean squared errors and frequencies of rejections at nominal significance level n.s. when testing for the true value $d_0=0$ based on 1000 replications. The test statistics $\mathcal{R}$ and $\mathcal{R}^*$ are from (3.4) and (3.6), respectively. The bandwidths are $m=T^{0.55},T^{0.65},T^{0.75}$. In bold face experimental levels not significantly different from the nominal ones at 95% level according to (4.2).

4.1. Deterministic choice of m

As deterministic rules for bandwidth selection we include $m=T^{0.55},T^{0.65}$ and $T^{0.75}$. Tables 1-4 report the rejection frequencies of the test statistics (3.4) and (3.6) for the LW estimator and the two-step mean-corrected ELW estimator, respectively. All empirical sizes are computed for the bandwidth values m and for the sample sizes T mentioned above.

Table 2. Stationary fractionally integrated noise: ARFI(0,0.45)

		$T=256$			$T=512$			$T=1024$		
Test statistic	n.s.	$T^{0.55}$	$T^{0.65}$	$T^{0.75}$	$T^{0.55}$	$T^{0.65}$	$T^{0.75}$	$T^{0.55}$	$T^{0.65}$	$T^{0.75}$
	1%	5.3	4.1	3.1	4.5	2.6	**1.2**	3.1	1.7	2.4
$\mathcal{R}_{LW}$	5%	14.2	13.4	10.3	12.2	8.6	8.4	10.5	7.8	7.2
	10%	23.2	20.1	17.1	20.3	15.2	14.6	17.3	13.6	12.1
	1%	2.0	**1.4**	**1.1**	**1.6**	**1.1**	**0.7**	1.7	**1.0**	1.7
$\mathcal{R}^*_{LW}$	5%	**5.9**	7.8	6.4	6.5	**5.5**	**5.8**	**5.8**	**5.2**	**5.2**
	10%	12.0	13.7	12.6	**11.7**	**11.0**	11.9	**11.7**	**10.9**	**10.7**
RMSE		0.149	0.103	0.073	0.115	0.083	0.055	0.090	0.060	0.040
	1%	5.7	4.2	3.4	4.6	2.6	1.7	3.5	2.2	2.5
$\mathcal{R}_{2ELW}$	5%	14.5	12.8	10.5	12.6	8.7	8.7	11.5	7.8	8.4
	10%	22.6	19.3	19.7	20.4	16.5	14.5	18.5	14.2	14.7
	1%	2.8	2.1	1.9	1.7	**1.5**	1.8	1.8	**1.6**	1.9
$\mathcal{R}^*_{2ELW}$	5%	8.1	7.5	6.8	**6.3**	**5.8**	**5.9**	**5.9**	**5.8**	7.3
	10%	13.9	13.7	13.9	11.9	**11.2**	11.9	**11.6**	**11.2**	12.9
RMSE		0.148	0.110	0.078	0.117	0.082	0.056	0.092	0.062	0.045

Results when testing for the true value $d_0 = 0.45$; for further comments see Table 1.

Generally, it can be observed that the size distortion of the test tends to decrease as T and m increase. In the case of white noise (Table 1), $\mathcal{R}$ is noticeably oversized, and the performance of all estimators is very similar. The rejection probabilities seem indeed to be markedly sensitive to the bandwidth choice, though this problem is reduced for large sample sizes. The variance approximation used in $\mathcal{R}^*$ from (3.6) clearly reduces the size distortion in all cases and for both estimators. When $d = 0.45$, the size distortion is at least as large as in the white noise case. The performance of the LW estimator in terms of size and RMSE is a bit better than that of 2ELW in almost all cases, also when the sample size is small, but the overall performance of the two approaches is quite similar. Again, the variance approximation used in (3.6) reduces the size distortion of all tests remarkably. In the case d takes on the nonstationary value 0.7, the two-stage ELW estimator outperforms the LW estimator, above all in terms of size distortion, see Table 3.

Table 3. Nonstationary fractionally integrated noise: ARFI(0,0.7)

		$T = 256$			$T = 512$			$T = 1024$		
Test statistic	n.s.	$T^{0.55}$	$T^{0.65}$	$T^{0.75}$	$T^{0.55}$	$T^{0.65}$	$T^{0.75}$	$T^{0.55}$	$T^{0.65}$	$T^{0.75}$
	1%	5.8	3.7	3.4	4.0	3.4	4.2	3.6	2.9	4.4
$\mathcal{R}_{LW}$	5%	15.1	11.7	10.1	12.3	11.8	11.6	13.1	10.9	12.0
	10%	22.5	19.5	16.0	19.1	18.2	17.8	18.4	17.0	19.8
	1%	**1.6**	**1.3**	**1.5**	**1.5**	1.9	2.6	1.9	1.8	3.6
$\mathcal{R}_{LW^*}$	5%	7.2	6.6	7.0	**5.8**	7.3	8.5	7.4	7.3	10.6
	10%	13.1	12.4	12.5	**11.8**	13.4	15.0	13.8	13.6	17.1
RMSE		0.148	0.103	0.075	0.114	0.083	0.058	0.092	0.065	0.047
	1%	4.4	3.1	2.1	3.8	2.4	3.1	2.7	2.3	2.0
$\mathcal{R}_{2ELW}$	5%	12.5	9.9	9.6	11.8	10.8	9.2	9.7	7.3	8.0
	10%	21.2	17.2	14.6	18.9	17.5	15.3	17.2	13.8	14.4
	1%	1.7	**1.6**	**1.4**	**1.3**	**1.5**	**1.6**	**1.1**	**1.0**	**1.5**
$\mathcal{R}^*_{2ELW}$	5%	**5.6**	**5.0**	**5.2**	**5.8**	7.1	7.5	**5.6**	**5.0**	**5.7**
	10%	**11.3**	**11.1**	**11.7**	**10.9**	**11.7**	12.5	**11.4**	**9.6**	11.9
RMSE		0.144	0.102	0.072	0.112	0.081	0.056	0.088	0.058	0.042

Results when testing for the true value $d_0 = 0.7$; for further comments see Table 1.

When short-run dynamics are added to the model, the selection of an adequate bandwidth becomes even more decisive. Table 4 shows the simulation results for an ARFI(1,0) process with a moderate autoregressive coefficient $a = 0.5$. The influence of the short-term component dominates the behaviour of all estimators, and the contamination of the periodogram leads to a large bias in the estimation of d. The size distortion caused by the autoregressive term increases excessively as m gets large. If we allowed for a moving-average component instead of the AR(1) dynamics, similarly detrimental effects are obtained. The theoretical charm of semiparametric estimators consists in their (asymptotic) robustness against the presence of short memory. The simulations show, however, that the finite sample performance may be far from the asymptotic promise. We conclude that even for samples as large as $T = 1000$ the bandwidth must be chosen very conservatively ($m < T^{0.65}$) to get somewhere close to the nominal size. Once more, the variance approximation used in (3.6) clearly outperforms the test from (3.4).

Table 4. AR(1)

		$T=256$			$T=512$			$T=1024$		
Test statistic	n.s.	$T^{0.55}$	$T^{0.65}$	$T^{0.75}$	$T^{0.55}$	$T^{0.65}$	$T^{0.75}$	$T^{0.55}$	$T^{0.65}$	$T^{0.75}$
	1%	8.1	23.1	91.8	4.3	16.6	94.6	3.6	11.2	96.9
$\mathcal{R}_{LW}$	5%	18.5	41.8	97.3	13.1	31.2	98.4	13.0	24.3	99.2
	10%	25.7	52.3	97.9	20.1	43.0	99.4	19.5	34.7	99.5
	1%	3.1	13.4	88.2	**1.6**	9.8	92.0	1.9	7.2	95.7
$\mathcal{R}^*_{LW}$	5%	9.6	30.6	95.6	**6.1**	24.2	97.9	7.2	19.4	99.2
	10%	15.9	42.2	97.5	12.5	35.2	99.2	14.6	30.0	99.5
RMSE		0.158	0.172	0.276	0.120	0.120	0.220	0.091	0.087	0.174
	1%	7.1	28.6	95.6	4.8	20.5	98.5	3.4	10.9	98.4
$\mathcal{R}_{2ELW}$	5%	17.8	45.1	98.6	12.6	38.5	99.6	10.7	25.9	99.6
	10%	26.2	55.0	99.4	19.5	47.7	99.8	19.3	35.3	99.7
	1%	2.6	16.2	94.0	**1.6**	13.0	96.4	1.7	7.6	97.1
$\mathcal{R}^*_{2ELW}$	5%	8.4	35.5	97.6	6.7	31.0	99.6	7.1	20.7	99.6
	10%	15.5	46.1	99.3	12.3	42.6	99.8	13.0	31.2	99.7
RMSE		0.159	0.183	0.316	0.116	0.130	0.244	0.095	0.087	0.192

The model is ARFI$(1,0)$ with $a=0.5$; for further comments see Table 1.

4.2. Data-driven choice of m

In this subsection we study the iterative procedure defined in (3.2), concentrating on the LW estimator in order to save space. In the white noise case we observe in Table 5 similar empirical sizes as under deterministic bandwidth selection in Table 1. For the ARFI(0,d) cases (Tables 5 and 6) the size distortion is larger than under the deterministic rules reported in Tables 2 and 3. Finally, in the ARFI(1,0) case (Table 6) the data-driven selection is superior only to large bandwidths chosen deterministically ($T^{0.65}$ or $T^{0.75}$). This reinforces the above warning to select m rather conservatively in practice to circumvent bias and size distortion due to short-memory. All in all, data-driven bandwidth determination has no benevolent effect on the size distortion. What is more, there were many cases where no convergence according to (3.2) was achieved (see Tables 5 and 6).

Table 5. LW and data-driven bandwidth selection

		White noise			ARFI(0,0.45)		
Test statistic	n.s.	$T = 256$	$T = 512$	$T = 1024$	$T = 256$	$T = 512$	$T = 1024$
	1%	4.9	3.5	2.7	8.2	6.5	4.1
$\mathcal{R}_{LW}$	5%	12.3	8.5	7.1	17.6	15.7	12.3
	10%	19.9	15.2	12.5	25.0	23.2	18.2
	1%	2.5	2.2	2.1	3.3	2.3	2.0
$\mathcal{R}^*_{LW}$	5%	7.5	**5.5**	**5.4**	9.7	9.8	8.3
	10%	13.7	11.3	**9.8**	15.9	15.9	14.3
Non convergence		70	50	26	29	8	8

Based on 1000 replications of the ARFI$(0,0)$ and ARFI(0,0.45) models. The number of non convergence cases is also reported. For further comments see Table 1.

Table 6. LW and data-driven bandwidth selection

		ARFI(0,0.7)			ARFI(1,0)		
Test statistic	n.s.	$T = 256$	$T = 512$	$T = 1024$	$T = 256$	$T = 512$	$T = 1024$
	1%	14.2	9.8	9.7	14.2	8.3	5.7
$\mathcal{R}_{LW}$	5%	23.6	19.0	18.4	24.0	16.4	15.0
	10%	31.2	26.6	26.1	31.8	23.4	21.6
	1%	8.1	4.1	3.8	7.1	4.0	3.0
$\mathcal{R}^*_{LW}$	5%	15.4	10.4	10.6	15.1	10.6	10.5
	10%	21.5	18.9	18.5	21.6	18.3	16.1
Non convergence		48	17	8	15	7	8

Based on 1000 replications of ARFI(0,0.7), and ARFI(1,0) with $a = 0.5$; for further comments see Table 5.

5. Concluding remarks

The LW estimator is asymptotically more efficient than other popular semiparametric long memory estimators. Moreover, the procedure has been proven to be robust to heteroskedasticity of a certain degree, see for instance Robinson and Henry [24]. At the same time it has at least three limitations. First, consistency and limiting normality only hold for a restricted parameter range excluding relevant cases of nonstationarity (see Robinson [22] and Velasco [30]). To overcome this shortcoming, Shimotsu and Phillips [28] proposed a computationally more involved variant called the exact LW estimator. It has the same limiting properties as LW but covers also the region of nonstationarity. In most practical situations, a mean-corrected version of the exact LW has to be worked with, see Shimotsu [27]. Second, the normal distribution of the normalized estimator holds of course only asymptotically. In finite samples it may be hard to control the probability of a type I error. It is a priori not clear how the estimator should be normalized to get a test statistic with satisfactory size properties in finite samples. Third, the semiparametric nature of LW hinges on an appropriate choice of a bandwidth m that balances the trade-off between bias and variance. The optimal rate of divergence for m was determined such as to minimize the asymptotic mean squared error [mse]. The resulting rate, however, violates the rate given in (3.1) ensuring limiting normality. Hence, the question arises how the normalized estimator will behave as a test statistic when the bandwidth is chosen according to a data-driven criterion that is mse optimal.

All three issues just raised with the (exact) LW estimator are addressed in our paper by means of a Monte Carlo study. Our contributions can be summarized as follows. First, the 2ELW by Shimotsu [27] is superior to the LW estimator when the process is indeed nonstationary. In the region of stationarity, however, there are many situations where LW dominates the exact variant in terms of size distortion and mse. Not knowing in practice, whether the true d is less than 0.5 or not, both ELW and LW can equally be advised. Second, with respect to the normal approximation we compared the usual asymptotic version of the test statistic $\mathcal{R}$ from (3.4) with the finite sample modification $\mathcal{R}^*$ from (3.6). Our experimental results on size distortions are very clear-cut: the finite sample approximation $\mathcal{R}^*$ constitutes a uniform improvement in that it always outperforms the classical variant. This holds for LW just as well as for 2ELW. Finally, we ran a horserace between deterministic bandwidth selection and data-driven selection rules relying on the mse optimal bandwidth rate. It turns out, generally speaking, that a data-driven bandwidth choice affects subsequent inference even in large samples. Resulting size distortions are large compared to the case of careful, moderate choices of deterministic values of m. Our experiments show that even a small bandwidth choice of $m = T^{0.55}$ may result in a too liberal size performance under short memory for samples with length between $T = 250$ and $T = 1000$.

References

[1] Abadir, K. M., Distaso, W. and Giraitis, L. (2007). Nonstationarity-extended local whittle estimation. *Journal of Econometrics*, **141**, 1353–1384.

[2] Delgado, M. and Robinson, P. M. (1996). Optimal spectral bandwidth for long memory. *Statistica Sinica*, **6**, 97–112.

[3] Faÿ, G., Moulines, E., Roueff, F. and Taqqu, M. S. (2009). Estimators of long memory: Fourier versus wavelets. *Journal of Econometrics*, **151**, 159–177.

[4] Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, **4**, 221–238.

[5] Giraitis, L., Robinson, P. M. and Samarov, A. (2000). Adaptive semiparametric estimation of the memory parameter. *Journal of Multivariate Analysis*, **72**, 183–207.

[6] Granger, C. W. J. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, **1**, 15–29.

[7] Hassler, U. and Kokoszka, P. (2010). Impulse responses of fractionally integrated processes with long memory. *Econometric Theory*, **26**, 1855–1861.

[8] —, Marmol, F. and Velasco, C. (2006). Residual log-periodogram inference for long-run relationships. *Journal of Econometrics*, **130**, 165–207.

[9] Hauser, M. A. (1997). Semiparametric and nonparametric testing for long memory: A monte carlo study. *Empirical Economics*, **22**, 247–271.

[10] Henry, M. (2001). Robust automatic bandwidth for long memory. *Journal of Time Series Analysis*, **22**, 293–316.

[11] — and Robinson, P. M. (1996). Bandwidth choice in gaussian semiparametric estimation of long range dependence. In P. M. Robinson and M. Rosenblatt (eds.), *Athens Conference on Applied Probability and Time Series, Vol. II*, New York: Springer, pp. 220–232.

[12] — and Zaffaroni, P. (2003). The long-range dependence paradigm for macroeconomics and finance. In P. Doukhan, G. Oppenheim and M. S. Taqqu (eds.), *Theory and Applications of Long-Range dependence*, New York: Birkhäuser Boston, pp. 417–438.

[13] Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, **68**, 165–176.

[14] Hurvich, C. M. and Beltrao, K. I. (1994). Automatic semiparametric estimation of the memory parameter of a long memory time series. *Journal of Time Series Analysis*, **15**, 285–302.

[15] — and Chen, W. W. (2000). An efficient taper for overdifferenced series. *Journal of Time Series Analysis*, **21**, 155–180.

[16] — and Deo, R. S. (1999). Plug-in selection of the number of frequencies in regression estimates of the memory parameter of a long-memory time series. *Journal of Time Series Analysis*, **20**, 331–341.

[17] — and Ray, B. K. (1995). Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated processes. *Journal of Time Series Analysis*, **16**, 17–41.

[18] Künsch, H. R. (1987). Statistical aspects of self-similar processes. *In: Proceedings of the First World Congress of the Bernoulli Society*, **1**, 67–74.

[19] Marinucci, D. and Robinson, P. M. (1999). Alternative forms of fractional Brownian motion. *Journal of Statistical Planning and Inference*, **80**, 111–122.

[20] Nielsen, M. Ø. and Frederiksen, P. H. (2005). Finite sample comparison of parametric, semiparametric, and wavelet estimators of fractional integration. *Econometric Reviews*, **24**, 405–443.

[21] Robinson, P. M. (1994). Rates of convergence and optimal bandwidth in spectral analysis of processes with long range dependence. *Probability Theory and Related Fields*, **99**, 443–473.

[22] — (1995). Gaussian semiparametric estimation of long range dependence. *Annals of Statistics*, **23**, 1630–1661.

[23] — (2005). The distance between rival nonstationary fractional processes. *Journal of Econometrics*, **128**, 283–399.

[24] — and Henry, M. (1999). Long and short memory conditional heteroskedasticity in estimating the memory parameter of levels. *Econometric Theory*, **15**, 299–336.

[25] Shao, X. and Wu, W. B. (2007). Local whittle estimation of fractional integration for nonlinear processes. *Econometric Theory*, **23**, 899–951.

[26] Shimotsu, K. (2006). Simple (but effective) tests of long memory versus structural breaks. *Working Paper, Queen's University.*

[27] — (2010). Exact local Whittle estimation of fractional integration with unknown mean and trend. *Econometric Theory*, **26**, 501–540.

[28] — and Phillips, P. C. B. (2005). Exact local Whittle estimation of fractional integration. *The Annals of Statistics*, **33**, 1890–1933.

[29] Taqqu, M. S. and Teverovsky, V. (1996). Semi-parametric graphical estimation techniques for long-memory data. In P. M. Robinson and M. Rosenblatt (eds.), *Athens Conference on Applied Probability and Time Series, Vol. II*, New York: Springer, pp. 420–432.

[30] Velasco, C. (1999). Gaussian semiparametric estimation of non-stationary time series. *Journal of Time Series Analysis*, **20**, 87–127.

[31] WHITTLE, P. (1951). *Hypothesis testing in time series analysis.* Almqvist and Wiksell.

[32] — (1953). Estimation and information in stationary time series. *Arkiv för Matematik*, **2**, 423–434.

1 306010 270419 >

İSTATİSTİK: JOURNAL OF THE TURKISH STATISTICAL ASSOCIATION
Vol. 6, No. 1, January 2013, pp. 42–55
ISSN 1300-4077 | 13 | 1 | 42 | 55

İSTATİSTİK

PERFORMANCE EVALUATION OF A FINITE BUFFER SYSTEM WITH VARYING RATES OF IMPATIENCE

Amit Choudhury and Pallabi Medhi*
Dept. of Statistics
Gauhati University, Guwahati-781014, Assam, India.

Abstract: In the analysis of reneging behavior, one approach is to assume random patience time. Each customer is assumed to follow identical distribution of patience time. However, there are many real life queuing systems where this assumption is not satisfied. Customers waiting in the system are often aware of their position in it and hence the reneging rate varies with the position of the customer in the system. This paper is an attempt to model such a reneging phenomena along with balking. Explicit closed form expressions of a number of performance measures have been presented. A numerical example with design aspects has also been presented to demonstrate the results derived.

Key words: Balking, Finite buffer queue, Impatience, Queuing, Reneging.
History: Submitted: 21 May 2012; Revised: 5 October 2012; Accepted: 14 November 2012

1. Introduction

In this competitive world, the key to survival of service oriented organizations rest in positive customer experience at the point of service delivery. It is argued that unless the customer is satisfied with the quality of service offered, customer loyalty may suffer. It is in this context that waiting for service attains importance, as queues are unavoidable. That being so, customers get impatient on having to experience a queue. This impatience finds reflection in two ways viz: balking and reneging. Even though these concepts have been dealt with in literature, closed form expressions are not always available. This paper is an attempt to fill this gap in literature.

By balking, we mean the phenomenon of customers arriving for service into a non-empty queue and leaving without joining the queue. There is no balking from an empty queue. Haight (1957) has provided a rationale, which might influence a person to balk. It relates to the perception of the importance of being served which induces an opinion somewhere in between urgency, so that a queue of certain length will not be joined, to indifference where a non-zero queue is also joined.

A customer will be said to have reneged if after joining the system it gets impatient and leaves the system without receiving service. On joining the system, it has a patience time such that in case service is unavailable within this patience time, the customer reneges.

Reneging can be of two types-viz. reneging till beginning of service (henceforth referred to as R_BOS) and reneging till end of service (henceforth referred to as R_EOS). A customer can renege only as long as it is in the queue and we call this as reneging of type R_BOS. It cannot renege once it begins receiving service. A common example is the barbershop. A customer can renege while he is waiting in queue. However once service gets started i.e. hair cut begins, the customer cannot leave till hair cutting is over. On the other hand, if customers can renege not only while waiting in queue but also while receiving service, we call such behavior as reneging of type R_EOS. "Ward and Bambos (2003) detailed a case of web system. In such a system, request for specific pages

* Corresponding author: pmedhi16@gmail.com (P.Medhi)

arrive at random points in time and are served in the order received. Impatient users may cancel their request at any time and that may occur either during service or before service has begun. Overall such dynamics arise in any queuing situation with time sensitive jobs, where reneging may occur both in queues and at the server".

In the analysis of reneging phenomena, one approach is to assume that each customer has a Markovian patience time, the distribution of which is position independent. However, it is our common day observation that there are systems where the customer is aware of its position in the system. For example customers queuing at the O.P.D. (out patient department) clinic of a hospital would know of their position in the queue. This invariably causes waiting customers to have higher rates of reneging in case their position in the queue is towards the end. It is not unreasonable then to expect that such customers who are positioned at a distance from the service facility have reneging rates, which are higher than reneging rates of customers who are near the service facility. In other words, we assume that customers are "position aware" and in this paper we model the reneging phenomenon in such a manner that the Markovian reneging rate is a function of the position of the customer in the system. Customers at higher states will be assumed to have higher reneging rates.

The subsequent sections of this paper are structured as follows. Section 2 contains a brief review of the literature. Section 3 and section 4 contains the derivation of steady state probabilities and performance measures respectively. We perform sensitivity analysis in section 5. A numerical example is discussed in section 6. Concluding statements are presented in section 7. The appendix contains some derivation.

2. Literature Survey

Barrer (1957a) carried out one of the early work on reneging where he considered deterministic reneging with single server Markovian arrival and service rates. Customers were selected randomly for service. In his subsequent work, Barrer (1957b) also considered deterministic reneging (of both R_BOS and R_EOS type) in a multi server scenario with FCFS discipline. The general method of solution was extended to two related queuing problems. Another early work was by Haight (1959). Ancher and Gafarian (1963) carried out an early work on Markovian reneging with Markovian arrival and service pattern. Kok and Tijms (1985) considered a single server queuing system where a customer becomes a lost customer when its service has not begun within a fixed time. Haghighi et al. (1986) considered a Markovian multiserver queuing model with balking as well as reneging. Each customer had a balking probability which was independent of the state of the system. Reneging discipline considered by them was R_BOS. Liu et al. (1987) considered an infinite server Markovian queuing system with reneging of type R_BOS. Customers had a choice of individual service or batch service, batch service being preferred by the customer. Brandt and Brandt (1999) considered a S-server system with two FCFS queues, where the arrival rates at the queues and the service may depend on number of customers 'n' being in service or in the first queue. Customers in the first queue were assumed impatient customers with deterministic reneging. Boots and Tijms (1999) considered an $M/M/C$ queue in which a customer leaves the system when its service has not begun within a fixed interval after its arrival. They have given the probabilistic proof of 'loss probability', which was expressed in a simple formula involving the waiting time probabilities in the standard $M/M/C$ queue. Ke and Wang (1999) considered the machine repair problem in which failed machines balk with probability $(1-b)$ and renege according to a negative exponential distribution. Another work using the concepts of balking and reneging in machine interference queue has been carried out by Al-Seedy and Al-Ibraheem (2001). Choi et al. (2001) introduced a simple approach for the analysis of the $M/M/C$ queue with a single class of customers and constant patience time by finding simple Markov process. Applying this approach, they analyzed the M/M/1 queue with two classes of customer in which class 1 customer have impatience of constant duration and class 2 customers have no impatience and lower priority than class 1 customers. Performance measures of both $M/M/C$ and $M/M/1$ queues were discussed. Zhang et al. (2005) considered an $M/M/1/N$ framework with Markovian reneging where they derived the steady state probabilities and formulated a cost model. Some performance measures were also discussed. Choudhury (2008) provided a detailed and lucid derivation of the distribution of virtual waiting time in a single server Markovian queuing system under R-BOS.

El- Paoumy (2008a) derived the analytical solution of $M^x/M/2/N$ queue for batch arrival system with Markovian reneging. Another paper on Markovian reneging was by Altman and Yechiali (2008). They derived the probability generating function of number of customers present in the system and some performance measures were discussed. Xiong and Altiok (2009) have provided approximations for some performance measures of a multi server queue with Poisson arrivals, general service time distribution and deterministic reneging.

Yechiali (2007) considered a multi server Markovian queue suffering occasional disaster breakdown. During such breakdown, all customers in the system are cleared. New arrivals during the breakdown period have an exponentially distributed patience time, such that if the service is not reactivated during this patience time, the customer reneges.

Other attempts at modeling reneging phenomenon include those by Baccelli et al. (1984), Martin and Artalejo (1995), Shawky (1997), Choi et al. (2004), Singh et al. (2007), El- Sherbiny (2008) and El-Paoumy and Ismail (2009) etc.

An early work on balking was carried out by Haight (1957). Jouini et al. (2008) modeled a call center as an M/M/s+M queue with endogenized customer reactions to announcements. They assumed that customers react by balking upon hearing the delay announcement and may subsequently renege if their realized waiting time exceeds the delay that has originally announced to them. They calculated the waiting time distribution i.e. announcement coverage and subsequent performance in terms of balking and reneging. Al-Seedy et al. (2009) presented an analysis for the M/M/c queue with balking and reneging. They assumed that arriving customers balked with a fixed probability and reneged according to a negative exponential distribution. The generating function technique was used to obtain the transient solution of system those results in a simple differential equation. Yue et al. (2009) considered an M/M/2 queuing system with balking and two heterogeneous servers, server 1 and server 2. They assumed that customers arrived according to a Poisson process and form a single waiting line where two parallel servers provided heterogeneous exponential service on a first-come first-served basis. It is also assumed that server 1 is perfectly reliable and server 2 is subject to breakdowns. They obtained the stationary condition where the system reaches a steady state and derived the steady state probabilities in a matrix form by using matrix-geometric solution method. They produced explicit expressions of some performance measures such as the mean system size, the average balking rate and the probabilities that server 2 is in various states. Choudhury and Medhi (2011) analyzed a multiserver Markovian queuing system under the assumption that customers may balk as well as renege. They assumed that each arriving customer has probability $(1-p)$ of balking from a system with no idle servers and for reneging they assumed that each customer joining the system have a random patience time. Explicit closed form expressions were presented. A numerical example with design aspects was also discussed to demonstrate results derived.

Some other papers which have considered both balking and reneging are the work by Shawky and El-Paoumy (2000), El- Paoumy (2008a and 2008b), El- Sherbiny (2008), Shawky and El-Paoumy (2008), Pazgal et al. (2008).

3. The Model and System State Probabilities

The model we deal with in this paper is the traditional M/M/1/k model with the restriction that customers may balk from a non-empty queue as well as may renege after they join the queue.

We shall assume that the inter arrival and service rates are λ and μ respectively. Analysis of M/M/1/k queuing model assumes significance from the fact that in the classical M/M/1 model, "it is assumed that the system can accommodate any number of units. In practice, this may seldom be the case. We have thus to consider the situation such that the system has limited waiting space and can hold a maximum number of k units (including the one being served)" {Medhi (1994)}. Our formulation differs from that of Medhi (1994) in that his model did not consider balking and reneging. As stated earlier we impose the additional restrictions of state dependent balking and position dependent reneging.

For balking it will be assumed that if the customer on arrival observes the system to be in state 'i', the probability that he will balk is 'i/k', $i = 1, 2, \ldots, k$. With this set up, the finite buffer restriction can also be seen as the state from which customer balks with probability 1(=k/k). There is no balking from an empty queue.

For reneging, we shall assume that customers joining the system are of Markovian reneging type. We shall assume that on joining the system, the customer is aware of its position in the same. Consequently, the reneging rate is taken as a function of the customer's position in the system. In particular, a customer who is at state 'n' will be assumed to have random patience time following exp (ν_n). Under R_BOS, we shall assume that

$$v_n = \begin{cases} 0 & \text{, for } n = 0, 1. \\ v + c^{n-1} & \text{, for } n = 2, 3, \ldots, k \end{cases}$$

and under R_EOS,

$$v_n = \begin{cases} 0 & \text{for } n = 0. \\ v + c^{n-1} & \text{for } n = 1, 2, \ldots, k \end{cases}$$

where c is a constant (c≥0 and c≠1).

Our aim behind this formulation is to ensure that customer's at higher positions have monotonic reneging rates. This requires that as a customer progresses in the system from position n to (n-1), the reneging distribution shift from $\exp(v_n)$ to $\exp(v_{n-1})$. In view of the memory less property, this shifting of reneging distribution is mathematically tractable as we shall demonstrate in the subsequent sections.

Our work stands out on a number of counts. First, one can observe from section 2 that existing reneging literature does not analyze the case where the reneging behavior is position dependent. All such Markovian reneging rules assume that reneging rate is constant irrespective of the position of the customer. To the best of our knowledge, formulation of position dependent reneging rule has not been attempted in literature. However, as mentioned earlier, there are many systems where customers are position aware and hence have variable reneging rates. This formulation is an important focus of this paper. Second, even though one can observe reneging and balking in our day-to-day life, very little work has analyzed these features together. This has been attempted here. Third, an important focus of this paper is the derivation of explicit closed form expressions of performance measures which can be used 'off the shelf' by practitioners. Reneging and balking literature seldom provide explicit closed form expressions; much less when reneging and balking are both involved as in the case of this paper.

The steady state probabilities are derived by the Markov process method. We first analyze the case where customers renege only from the queue. Under R_BOS, letp_n denote the probability that there are 'n' customers in the system. The steady state probabilities under R_BOS are given below,

$$\lambda p_0 = \mu p_1 \tag{3.1}$$

$$\begin{aligned} \lambda\{1-(n-1)/k\}p_{n-1} + \{\mu + nv + c(c^n-1)/(c-1)\}p_{n+1} = \lambda\{1-n/k\}p_n + \\ \{\mu + (n-1)v + c(c^n-1)/(c-1)\}p_{n-1}\ ;\ n = 1, 2, \ldots, k-1 \end{aligned} \tag{3.2}$$

$$\lambda\{1-(k-1)/k\}\,p_{k-1}=\left\{\mu+(k-1)v+c(c^{k-1}-1)/(c-1)\right\}p_k \tag{3.3}$$

Solving recursively, we get (under R_BOS)

$$p_n=\left[\left\{\lambda^n\prod_{r=1}^{k}(1-\overline{r-1}/k\right\}/\left\{\prod_{r=1}^{n}\left(\mu+\overline{r-1}v+c\overline{c^{r-1}-1}/\overline{c-1}\right)\right\}\right]p_0\ ;\ n=1,2,\ldots,k \tag{3.4}$$

where p_0 is obtained from the normalizing condition $\sum_{n=0}^{k}p_n=1$ and is given as

$$p_0=\left[1+\sum_{n=1}^{k}\left\{\lambda^n\prod_{r=1}^{k}(1-\overline{r-1}/k\right\}/\left\{\prod_{r=1}^{n}\left(\mu+\overline{r-1}v+c\overline{c^{r-1}-1}/\overline{c-1}\right)\right\}\right]^{-1} \tag{3.5}$$

The steady state probabilities satisfy the recurrence relation, under R_BOS

$$p_n=\left[\{\lambda(1-\overline{n-1}/k)\}/\{\mu+\overline{n-1}\,\nu+c(c^{n-1}-1)/(c-1)\}\right]p_{n-1};n=1,2,\ldots,k$$

Under R_EOS where customers may renege from queue as well as while being served, let q_n denote the probability that there are n customers in the system. Applying the Markov theory, we obtain the following set of steady state equations.

$$\lambda q_0=(\mu+v)q_1 \tag{3.6}$$

$$\begin{aligned}\lambda\{1-(n-1)/k\}\,q_{n-1}+\{\mu+(n+1)v+c(c^n-1)/(c-1)\}\,q_{n+1}&=\lambda\{1-n/k\}\,q_n\\ +\left\{\mu+nv+c(c^{n-1}-1)/(c-1)\right\}q_n\ &;\ n=1,2,\ldots,k-1\end{aligned} \tag{3.7}$$

$$\lambda\{1-(k-1)/k\}\,q_{k-1}=\left\{\mu+kv+c(c^{k-1}-1)/(c-1)\right\}q_k \tag{3.8}$$

Solving recursively, we get (under R_EOS)

$$q_n=\left[\left\{\lambda^n\prod_{r=1}^{n}(1-\overline{r-1}/k\right\}/\left\{\prod_{r=1}^{n}\left(\mu+rv+c\overline{c^{r-1}-1}/\overline{c-1}\right)\right\}\right]q_0\ ;\ n=1,2,\ldots,k \tag{3.9}$$

where q_0 is obtained from the normalizing condition $\sum_{n=0}^{k}q_n=1$ and is given as

$$q_0=\left[1+\sum_{n=1}^{k}\left\{\lambda^n\prod_{r=1}^{k}(1-\overline{r-1}/k\right\}/\left\{\prod_{r=1}^{n}\left(\mu+rv+c\overline{c^{r-1}-1}/\overline{c-1}\right)\right\}\right]^{-1} \tag{3.10}$$

The steady state probabilities satisfy the recurrence relation, under R_EOS

$$q_n=\left[\left\{\lambda(1-\overline{n-1}/k)\right\}/\left\{\mu+nv+c(c^{n-1}-1)/(c-1)\right\}\right]q_{n-1}\ ;\ n=1,2,\ldots,k$$

4. Performance Measures

The main objective of any queuing study is to assess some well-defined parameters through which the nature of the quality of service can be studied. These parameters are known as performance measures. Performance measures are important as issues or problems caused by queuing situations are often related to customer's dissatisfaction with service or may be the root cause of economic losses in a business. Analysis of the relevant performance measures of queuing models allows the cause of queuing issues to be identified and the impact of proposed changes to be assessed.

An important measure is the mean number of customers in the system, which is traditionally denoted by 'L'. We have presented the derivation of this important performance measure separately for the two reneging disciplines in the appendix. These are denoted by $L_{\text{R\_BOS}}$ and $L_{\text{R\_EOS}}$.

Let $P(s)$ be the p.g.f of the steady state probability under R_BOS rule. Then we note that

$$\begin{aligned} L_{\text{R\_BOS}} &= \sum_{n=0}^{k} np_n \\ &= P'(1) \\ &= \frac{d}{ds} P(s)\,|_{s=1} \end{aligned}$$

(See the appendix for more derivations)

From (A.8) and (B.2), the mean system sizes under the two reneging rules are

$$L_{\text{R\_BOS}} = k[\lambda - (\mu - \nu)(1 - p_0) - p_0 + \{c - (p_0/p_0(c\lambda, \mu, \nu, k))/(c-1)\}]/(\lambda + k\nu) \tag{4.1}$$

$$L_{\text{R\_EOS}} = k[\lambda - \mu(1 - q_0) - q_0 + \{c - (q_0/q_0(c\lambda, \mu, \nu, k))/(c-1)\}]/(\lambda + k\nu) \tag{4.2}$$

The mean queue size formulas for the two cases can now be obtained and are given by

$$\begin{aligned} L_{q(\text{R\_BOS})} &= \sum_{n=2}^{k} (n-1)p_n \\ &= L_{R\_BOS} - (1 - p_0) \\ &= k[\lambda - (\mu - \nu)(1 - p_0) - p_0 + \{c - (p_0/p_0(c\lambda, \mu, \nu, k))/(c-1)\}]/(\lambda + k\nu) - (1 - p_0) \end{aligned}$$

where p_0 and $p_0(c\lambda, \mu, \nu, k)$ are defined in (3.5) and (A.4) respectively.
Similarly,

$$\begin{aligned} L_{q(R\_EOS)} &= \sum_{n=2}^{k} (n-1)q_n \\ &= L_{R\_EOS} - (1 - q_0) \\ &= k[\lambda - \mu(1 - q_0) - q_0 + \{c - (q_0/q_0(c\lambda, \mu, \nu, k))/(c-1)\}]/(\lambda + k\nu) - (1 - q_0) \end{aligned}$$

where q_0 and $q_0(c\lambda, \mu, \nu, k)$ are defined in (3.10) and (B.3) respectively.

Customers arrive into the system at the rate λ. However all the customers who arrive do not join the system because of balking and due to finite buffer restriction. The effective arrival rate into the system is thus different from the overall arrival rate and is given by

$$\begin{aligned} {\lambda^e}_{(\text{R\_BOS})} &= \lambda \sum_{n=0}^{k} (1 - n/k)p_n \\ &= \lambda[1 - L_{\text{R\_BOS}}/k] \\ &= \lambda[k\nu + (\mu - \nu)(1 - p_0) + p_0 - \{c - (p_0/p_0(c\lambda, \mu, \nu, k))/(c-1)\}]/(\lambda + k\nu) \end{aligned} \tag{4.3}$$

Similarly in case of R_EOS

$$\lambda^{e}{}_{R\_EOS} = \lambda[k\nu + \mu(1-q_0) + q_0 - \{c - (q_0/q_0(c\lambda,\mu,\nu,k))/(c-1)\}]/(\lambda + k\nu) \tag{4.4}$$

We have assumed that each customer has a random patience time following exp(v). Clearly then, the reneging rate of the system would depend on the state of the system as well as the reneging rule. The average reneging rate (avg rr) under two reneging rules are given by

$$\begin{aligned} Avgrr_{(R\_BOS)} &= \sum_{n=2}^{k} \{(n-1)\nu + c(c^{n-1}-1)/(c-1)\}q_n \\ &= \nu\{P'(1) - p_1\} - \nu(1-p_0-p_1) + \{1 \big/ (c-1)\} \sum_{n=2}^{k} c^n p_n - \{c/(c-1)\} \sum_{n=2}^{k} p_n \\ &= \nu L_{R\_BOS} - \nu(1-p_0) - c/(c-1) + p_0/\{p_0(c\lambda,\mu,\nu,k)(c-1)\} + p_0 \end{aligned} \tag{4.5}$$

$$\begin{aligned} Avgrr_{(R\_EOS)} &= \sum_{n=1}^{k} \{n\nu + c(c^{n-1}-1)/(c-1)\}q_n \\ &= \nu Q'(1) + \{1 \big/ (c-1)\} \sum_{n=1}^{k} c^n p_n - \{c/(c-1)\} \sum_{n=1}^{k} q_n \\ &= \nu L_{R\_EOS} + c/(c-1) - q_0/\{q_0(c\lambda,\mu,\nu,k)(c-1)\} + q_0 \end{aligned} \tag{4.6}$$

In a real life situation, customers who balk or renege represent the business lost. Customers are lost to the system in three ways, due to balking, due to finite buffer restriction and due to reneging. Management would like to know the proportion of total customers lost in order to have an idea of total business lost.

Hence the mean rate at which customers are lost (under R_BOS) is

$$\lambda - \lambda^{e}{}_{(R\_BOS)} + avgrr_{(R\_BOS)} = \lambda - \mu(1-p_0) \tag{4.7}$$

and the mean rate at which customers are lost (under R_EOS) is

$$\lambda - \lambda^{e}{}_{(R\_EOS)} + avgrr_{(R\_EOS)} = \lambda - \mu(1-q_0) \tag{4.8}$$

These rates helps in the determination of proportion of customers lost which is of interest to the system manager as also an important measure of business lost. This proportion (under R_BOS) is given by

$$\lambda - \lambda^{e}{}_{(R\_BOS)} + avgrr_{(R\_BOS)} \big/ \lambda = 1 - (\mu/\lambda)(1-p_0)$$

and the proportion (under R_EOS) is given by

$$\lambda - \lambda^{e}{}_{(R\_EOS)} + avgrr_{(R\_EOS)} \big/ \lambda = 1 - (\mu/\lambda)(1-q_0)$$

The proportion of customer completing receipt of service can now be easily determined from the above proportion.

The customers who leave the system from the queue do not receive service. Consequently, only those customers who reach the service station constitute the actual load of the server. From the server's point of view, this provides a measure of the amount of work he has to do. Let us call the rate at which customers reach the service station as λ^s .Then under R_BOS

$$\begin{aligned} &= \lambda^{e}{}_{(R\_BOS)} \left\{ 1 - \sum_{n=2}^{k} (n-1)\nu p_n / \lambda^{e}{}_{(R\_BOS)} \right\} \\ &= \lambda^{e}{}_{(R\_BOS)} - avgrr_{(R\_BOS)} \\ &= \mu(1-p_0) \end{aligned}$$

In case of R_EOS, one needs to recall that customers may renege even while being served and only those customers who renege from the queue will not constitute any work for the server. Thus,

$$\begin{aligned}
&\lambda^s_{(\text{R_EOS})} = \lambda^e_{(\text{R_EOS})}(\text{1-proportion of customers lost due to reneging out of those joining the system})\\
&= \lambda^e{}_{(\text{R_EOS})}\left\{1 - \sum_{n=2}^{k}(n-1)\nu q_n / \lambda^e{}_{(\text{R_EOS})}\right\}\\
&= \lambda^e{}_{(R_EOS)} - avgrr_{(R_EOS)}\\
&= \mu(1-q_0)
\end{aligned}$$

5. Sensitivity Analysis

We have assumed that there are essentially four parameters viz: λ, μ, ν, k relating to the stochastic nature of arrival, service, reneging patterns and system size. Various reasons may influence these parameters so that on different occasions these may undergo change. From managerial point of view, an idle server is a waste. So also for low server utilization. It is therefore interesting to examine and understand how server utilization varies in response to change in system parameters. We place below the effect of change in these system parameters on server utilization. For this purpose, we shall follow the following notational convention in the rest of this section.

$p_n\ (\lambda, \mu, \nu, k)$ and $q_n\ (\lambda, \mu, \nu, k)$ will denote the probability that there are 'n' customers in a system with parameters in (λ, μ, ν, k) stead state under R_BOS and R_EOS respectively.

i. Let $\lambda_1 > \lambda_0$, then

$$\begin{aligned}
&\tfrac{p_0(\lambda_1,\mu,\nu,k)}{p_0(\lambda_0,\mu,\nu,k)} < 1\\
&\Rightarrow \tfrac{(\lambda_0-\lambda_1)}{\mu} + \tfrac{(1-1/k)({\lambda_0}^2-{\lambda_1}^2)}{\mu(\mu+\nu+c)} + \ldots + \tfrac{(1-1/k)\ldots\{1-(k-1)/k\}({\lambda_0}^k-{\lambda_1}^k)}{\mu(\mu+\nu+c)\ldots\{\mu+(k-1)\nu+c(c^{k-1}-1)/(c-1)\}} < 0
\end{aligned}$$

which is true. Hence $p_0 \downarrow$ as $\lambda \uparrow$.

ii. Let $\mu_1 > \mu_0$, then

$$\begin{aligned}
&\tfrac{p_0(\lambda,\mu_1,\nu,k)}{p_0(\lambda,\mu_0,\nu,k)} > 1\\
&\Rightarrow \lambda\left(\tfrac{1}{\mu_0} - \tfrac{1}{\mu_1}\right) + \lambda^2(1-1/k)..\left\{\tfrac{1}{\mu_0(\mu_0+\nu+c)} - \tfrac{1}{\mu_1(\mu_1+\nu+c)}\right\} + \ldots +\\
&+\lambda^k(1-1/k)(1-2/k)\ldots\{1-(k-1)/k\}\left\{\begin{array}{c}\frac{1}{\mu_0(\mu_0+\nu+c)\ldots\{\mu_0+(k-1)\nu+c(c^{k-1}-1)/(c-1)\}} - \\ \frac{1}{\mu_1(\mu_1+\nu+c)\ldots\{\mu_1+(k-1)\nu+c(c^{k-1}-1)/(c-1)\}}\end{array}\right\} > 0
\end{aligned}$$

which is true. Hence $p_0 \uparrow$ as $\lambda \uparrow$.

iii. Let $v_1 > v_0$, then

$$\begin{aligned}
&\tfrac{p_0(\lambda,\mu,\nu_1,k)}{p_0(\lambda,\mu,\nu_0,k)} > 1\\
&\Rightarrow \lambda^2(1-1/k)\left\{\tfrac{1}{\mu(\mu+\nu_0+c)} - \tfrac{1}{\mu(\mu+\nu_1+c)}\right\} + \ldots\\
&+\lambda^k(1-1/k)(1-2/k)\ldots.\{1-(k-1)/k\}\left\{\begin{array}{c}\frac{1}{\mu(\mu+\nu_0+c)\ldots\{\mu+(k-1)\nu_0+c(c^{k-1}-1)/(c-1)\}} - \\ \frac{1}{\mu(\mu+\nu_1+c)\ldots\{\mu+(k-1)\nu_1+c(c^{k-1}-1)/(c-1)\}}\end{array}\right\} > 0
\end{aligned}$$

which is true. Hence $p_0 \uparrow$ as $v \uparrow$.

iv. Let $k_1 > k_0$, then

$$\begin{aligned}
&\tfrac{p_0(\lambda,\mu,\nu,k_1)}{p_0(\lambda,\mu,\nu,k_0)} < 1\\
&\Rightarrow \sum_{n=1}^{k_0} \frac{\lambda^n \prod_{r=1}^{n}\{1-(r-1)/k_0\}}{\prod_{r=1}^{n}\left\{\mu+(r-1)\nu+c(c^{r-1}-1)/(c-1)\right\}} - \sum_{n=1}^{k_1} \frac{\lambda^n \prod_{r=1}^{n}\{1-(r-1)/k_1\}}{\prod_{r=1}^{n}\left\{\mu+(r-1)\nu+c(c^{r-1}-1)/(c-1)\right\}} < 0
\end{aligned}$$

which is true. Hence $p_0 \downarrow$ as $k \uparrow$.

The following can similarly be shown
v. $q_0 \downarrow$ as $\lambda \uparrow$
vi. $q_0 \uparrow$ as $\mu \uparrow$
vii. $q_0 \uparrow$ as $v \uparrow$
viii. $q_0 \downarrow$ as $k \uparrow$
Under R_BOS, these results state that an increase in arrival rate would result in lowering of the fraction of time the server is idle. An increase in service rate would mean the server is able to work efficiently so that it can process same amount of work quickly. This translates to higher server idle time. An increase in reneging rate would mean the server has fewer work to do and hence higher fraction of idle time. An increase in system size translates to the lowering of the fraction of time the server is idle. Similar conclusions can be drawn under R_EOS.

6. Numerical Example

To illustrate the use of our results, we apply them to a queuing scenario. We quote below an example from Taha (2003, page 610).

'The time for barber Joe to give a haircut is exponential with mean of 12 minutes. Because of his popularity, customers usually arrive (according to a Poisson distribution) at a rate much higher than Joe can handle 6 customers per hour. Joe really will feel comfortable if the arrival rate is effectively reduced to about 4 customers per hour. To accomplish this goal, he came up with the idea of providing limited seating in the waiting area so that newly arriving customers would go elsewhere when they discover that all the seats are taken. How many seats should Joe provide to accomplish his goal?'

This is a design problem where the system manager (Joe, the barber) desires a system design in respect of size of the waiting area (number of chairs for waiting customers).

Here $\lambda = 6/hr$ and $\mu = 5/hr$. As required by Joe, we examine the effect of limited seating arrangement in the waiting area with different choices of k. Though not explicitly stated, it is necessary to assume reneging and balking. Customers these days are very hard pressed for time. Prompt customer service being the expectation, it is all the more reasonable to assume that customers are all of reneging type. We shall assume that reneging distribution is state dependent following $\exp(v_n)$ where v_n is as described in section 3. Specifically, we shall assume $v = 0.1/hr$ and considered the scenario with $c = 1.05$. Given the fact that service in a barbershop is being analyzed, clearly the reneging rule would be R_BOS. We further assume that the probability of balking by an arriving customer is 'i/k', $i = 1, 2, \ldots, k$ where i is the state of the customer observes the system to be in on its arrival.

Various performance measures of interest computed under different scenarios are given in Table in below. These measures were arrived at using a FORTRAN 77 program coded by the authors. Different choices of k were considered. Results relevant with regard to Joe's desire to limit arrival rate of customers into his service station to something around $4/hr$ are presented in the table.

Since a larger waiting area would also entail additional expenditure/ investment, Joe needs to examine how the performance measures differ across different choices of k. In case the reneging behavior of customer follows exp(0.1) distribution and when $c = 1.05$, it appears from the above table that an ideal choice of k could be 20 (seating space in waiting area =19) with $\lambda^s = 3.99628$.

Two interesting observations can be made from the above table. To a layman, Joe's aim of reducing λ from 6 to 4 effectively boils down to turning away one third of his customers. Our analysis confirms the same. In the scenario examined above, the percentage of customers lost due to reneging together with finite buffer at the level of ideal choice of k hovers very close to one third at 33.39%. Second, at the level of λ^s nearest to 4, the fraction of time Joe would be idle (p_0) is almost constant at 20% in the above scenarios. This results stand to reason.

TABLE 1. Performance measures assuming $\lambda = 6/hr$, $\mu = 5/hr$, $v = 0.1/hr$ and $c = 1.05$

Performance Measure	Size of Waiting Area		
	19 ($k=20$)	20 ($k=21$)	21 ($k=22$)
λ^s (i.e. arrival rate of customers reaching service station)	3.99628	4.00249	4.00816
Effective arrival rate (λ^e)	5.40723	5.43109	5.45309
Fraction of time server is idle	0.20074	0.19950	0.19837
Average length of queue	1.17665	1.19069	1.20369
Average length of system	1.97590	1.99119	2.00532
Mean reneging rate	1.41095	1.42860	1.44494
Rate of loss due to balking and finite buffer	0.59277	0.56891	0.54691
Mean rate of customers lost	2.00371	1.99751	1.99184
Proportion of customers lost due to reneging, balking and finite buffer	0.33395	0.33292	0.33197

7. Conclusion

The analysis of a single server Markovian queuing system with state dependent balking and position dependent reneging has been presented. Even though balking and reneging have been discussed by others, explicit expressions are not always available. Besides, to the best of our knowledge, modeling of position dependent reneging has not been attempted in literature. This paper makes a contribution here. Closed form expressions of number of performance measures have been derived. To study the change in the system corresponding to change in system parameters, sensitivity analysis has also been presented. A numerical example has been discussed to demonstrate results derived. The numerical example is of indicative nature meant to illustrate the benefits of our theoretical results in a design context. The limitations of this work stem from the Markovian assumptions. Extension of our results for general distribution is a pointer to future research.

Appendix

A. Derivation of $P'(1)$ under R_BOS

Let $P(s)$ denote the probability generating function, defined by $P(s) = \sum_{n=0}^{\infty} p_n s^n$. From (3.2) we have

$$\lambda\{1-(n-1)/k\}p_{n-1} + \{\mu + nv + c(c^n-1)/(c-1)\}p_{n+1} = \lambda\{1-n/k\}p_n + \left\{\mu + (n-1)v + c(c^{n-1}-1)/(c-1)\right\}p_n \;;\; n = 1, 2, \ldots, k-1$$

Multiplying both sides of the equation by s^n and summing over n

$$\lambda s\sum_{n=1}^{k-1}\{1-(n-1)/k\}p_{n-1}s^{n-1} - \lambda\sum_{n=1}^{k-1}(1-n/k)p_n s^n = \sum_{n=1}^{k-1}\{\mu + (n-1)v + (c^n-c)/(c-1)\}p_n s^n - \frac{1}{s}\sum_{n=1}^{k-1}\{\mu + nv + (c^{n+1}-c)/(c-1)\}p_{n+1}s^{n+1} \tag{A.1}$$

$\Rightarrow \lambda s[p_0 s^0 + (1-1/k)p_1 s^1 + \ldots + \{1-(k-1)/k\}p_{k-2}s^{k-2}] - \lambda[(1-1/k)p_1 s^1 + (1-2/k)p_2 s^2 + \ldots + \{1-(k-1)/k\}p_{k-1}s^{k-1} = \mu(p_1 s^1 + \ldots + p_{k-1}s^{k-1}) + \nu\{p_2 s^2 + 2p_3 s^3 + \ldots + (k-1)p_k s^k + \{1/(c-1)\}\{cp_1 s + c^2 p_2 s^2 + \ldots + c^{k-1}p_{k-1}s^{k-1} - c(p_1 s^1 + \ldots + p_{k-1}s^{k-1}\} - (1/s)[\mu(p_2 s^2 + \ldots + p_k s^k) + \nu\{p_2 s^2 + 2p_3 s^3 + \ldots + (k-1)p_k s^k + \{1/(c-1)\}\{c^2 p_2 s^2 + \ldots + c^k p_k s^k - c(p_2 s^2 + \ldots + p_k s^k\}$

$\Rightarrow \lambda s[(p_0 s^0 + \ldots + p_{k-2}s^{k-2}) - (1/k)\{p_1 s + 2p_2 s^2 + \ldots + (k-2)p_{k-2}s^{k-2}\}] - \lambda[(p_1 s^1 + \ldots + p_{k-1}s^{k-1}) - (1/k)\{p_1 s + 2p_2 s^2 + \ldots + (k-1)p_{k-1}s^{k-1}\}] = \mu\{P(s) - p_0 - p_k s^k\} + \nu s\{2p_2 s + \ldots + (k-1)p_{k-1}s^{k-2}\} -$

$\nu s(p_2 s+...+p_{k-1}s^{k-2})+\{1/(c-1)\}\{P(cs)-p_0-c^k p_k s^k-cP(s)+cp_0+cp_k s^k\}-(1/s)[\mu\{P(s)-p_0-p_1 s\}+\nu s\{2p_2 s+...+kp_k s^{k-1}\}-\nu s(p_2 s+...+p_k s^{k-1})+\{1/(c-1)\}\{P(cs)-p_0-cP(s)+cp_0\}]$

$\Rightarrow \lambda s[\{P(s)-p_{k-1}s^{k-1}-p_k s^k\}-(s/k)\{P'(s)-(k-1)p_{k-1}s^{k-2}-kp_k s^{k-1}\}]-\lambda[\{P(s)-p_k s^k-p_0\}-(s/k)\{P'(s)-kp_k s^{k-1}\}]=\mu\{P(s)-p_0-p_k s^k\}+\nu s\{P'(s)-p_1-kp_k s^{k-1}\}-\nu\{P(s)-p_0-p_1 s-p_k s^k\}+\{1/(c-1)\}\{P(cs)+p_0(c-1)-cP(s)-cp_k s^k(c^{k-1}-1)\}-(\mu/s)\{P(s)-p_0-p_1 s\}-\nu\{P'(s)-p_1\}+(\nu/s)\{P(s)-p_0-p_1 s\}-\{1/(c-1)s\}\{P(cs)+p_0(c-1)-cP(s)$

$\Rightarrow \{(\lambda s/k)+\nu\}P'(s)(1-s)=(\nu/s)P(s)(1-s)-(\mu/s)P(s)(1-s)+(\mu/s)p_0(1-s)-(\nu/s)p_0(1-s)-\{P(cs)(1-s)/(c-1)s\}-(p_0/s)(1-s)+\lambda P(s)(1-s)+\{\mu+(k-1)+c(c^{k-1}-1)/(c-1)\}p_k s^k-\{\mu-\nu+\nu k+c(c^{k-1}-1)/(c-1)\}p_k s^k$

$\Rightarrow \{(\lambda s+k\nu)/k\}P'(s)=[(\nu/s)-(\mu/s)+\lambda+c/\{s(c-1)\}]P(s)+\{(\mu/s)-(\nu/s)-1/s\}p_0-P(cs)/\{(c-1)s\}$

Now

$\lim_{s\to 1^-}P'(s)=\lim_{s\to 1^-}\{k/(\lambda+k\nu)\}[\{(\nu/s)-(\mu/s)+\lambda+c/(c-1)s\}P(s)+\{(\mu/s)-(\nu/s)-1/s)p_0-\{P(c)/(c-1)s\}$

$$\Rightarrow P'(1)=\{k/(\lambda+k\nu)\}[\lambda-(\mu-\nu)(1-p_0)-p_0+\{c-P(c)\}/(c-1)] \tag{A.2}$$

Here $P(c)=\sum_{n=0}^{k}p_n(\lambda,\mu,\nu,k)c^n$ where the symbol $p_n(\lambda,\mu,\nu,k)$ is as described in section 5. We use p_n and $p_n(\lambda,\mu,\nu,k)$ interchangeably. However should any of the parameters λ,μ,v,k change, it is explicitly stated. To obtain a closed form expression for $P(c)$, let us for the time being, consider another queuing system with parameter and assumptions similar to the queuing system we are presently considering except that the arrival rate is '$c\lambda$'. For this new system, the steady state equations are same as (3.1), (3.2) and (3.3) with 'λ' replaced by '$c\lambda$'. Denoting the steady state probabilities of this new system by $p_n(c\lambda,\mu,\nu,k)$, we can obtain

$$p_n(c\lambda,\mu,\nu,k)=\left[(c\lambda)^n\prod_{r=1}^{n}\{1-(r-1)/k\}\Big/\left\{\prod_{r=1}^{n}\left(\mu+\overline{r-1}v+c\overline{c^{r-1}-1}/\overline{c-1}\right)\right\}\right]p_0(c\lambda,\mu,\nu,k)\quad ;\ n=1,2,\dots,k \tag{A.3}$$

where

$$p_0(c\lambda,\mu,\nu,k)=\left[1+\sum_{n=1}^{k}(c\lambda)^n\prod_{r=1}^{n}\{1-(r-1)/k\}\Big/\left\{\prod_{r=1}^{n}\left(\mu+\overline{r-1}v+c\overline{c^{r-1}-1}/\overline{c-1}\right)\right\}\right]^{-1} \tag{A.4}$$

Let $P(S;c\lambda,\mu,\nu,k)$ denotes the probability generating function of this new queuing system so that

$$P(S;c\lambda,\mu,\nu,k)=\sum_{n=0}^{k}p_n(c\lambda,\mu,\nu,k)s^n$$

Now

$$\begin{aligned}P(c)&=\sum_{n=0}^{k}p_n(\lambda,\mu,\nu,k)c^n\\&=p_0+\sum_{n=1}^{k}\left[(c\lambda)^n\prod_{r=1}^{n}\{1-(r-1)/k\}\Big/\{\prod_{r=1}^{n}(\mu+\overline{r-1}\,\nu+c\overline{c^{r-1}-1}/\overline{c-1})\}\right]p_0\\ \Rightarrow[\{P(c)-p_0\}/p_0]&=\sum_{n=1}^{k}\left[(c\lambda)^n\prod_{r=1}^{n}\{1-(r-1)/k\}\Big/\{\prod_{r=1}^{n}(\mu+\overline{r-1}\,\nu+c\overline{c^{r-1}-1}/\overline{c-1})\}\right]p_0\end{aligned} \tag{A.5}$$

Now putting $S=1$ in $P(S;c\lambda,\mu,\nu,k)$ we get

$$\begin{aligned}
&P(1;c\lambda,\mu,\nu,k)=p_0(c\lambda,\mu,\nu,k)+\sum_{n=1}^{k}p_n(c\lambda,\mu,\nu,k)\\
&\Rightarrow 1=p_0(c\lambda,\mu,\nu,k)+\sum_{n=1}^{k}\left[(c\lambda)^n\prod_{r=1}^{n}\{1-(r-1)/k\}\Big/\{\prod_{r=1}^{n}(\mu+\overline{r-1}\nu+c\overline{c^{r-1}-1}/\overline{c-1})\}\right]p_0(c\lambda,\mu,\nu,k)\\
&\Rightarrow 1=p_0(c\lambda,\mu,\nu,k)+\{(P(c)-p_0)/p_0\}\,p_0(c\lambda,\mu,\nu,k)\\
&\Rightarrow P(c)=p_0/p_0(c\lambda,\mu,\nu,k)
\end{aligned} \tag{A.6}$$

{using (A.3) and (A.5)}
Similarly under R_EOS,

$$Q(c)=q_0/q_0(c\lambda,\mu,\nu,k) \tag{A.7}$$

Using equation (A.6) in (A.2) we obtain

$$P'(1)=\{k/(\lambda+k\nu)\}[\lambda-(\mu-\nu)(1-p_0)-p_0+\{c-p_0/p_0(c\lambda,\mu,\nu,k)\}/(c-1)] \tag{A.8}$$

where $p_0(c\lambda,\mu,\nu,k)$ is given in (A.4).

B. Derivation of $Q'(1)$ under R_EOS

From equation (3.7), we have

$$\begin{aligned}
\lambda\{1-(n-1)/k\}q_{n-1}+\{\mu+(n+1)\nu+c(c^n-1)/(c-1)\}\,q_{n+1}&=\lambda(1-n/k)q_n\\
&+\left\{\mu+n\nu+c(c^{n-1}-1)/(c-1)\right\}q_n;\, n=1,2,\ldots,k-1
\end{aligned}$$

Multiplying both sides of this equation by s^n and summing over n we get

$$\begin{aligned}
\lambda s\sum_{n=1}^{k-1}\{1-(n-1)/k\}q_{n-1}s^{n-1}-\lambda\sum_{n=1}^{k-1}\{1-n/k\}q_ns^n&=\sum_{n=1}^{k-1}\{\mu+n\nu+(c^n-c)/(c-1)\}q_ns^n\\
&-\tfrac{1}{s}\sum_{n=1}^{k-1}\{\mu+(n+1)\nu+(c^{n+1}-c)/(c-1)\}q_{n+1}s^{n+1}
\end{aligned} \tag{B.1}$$

Proceeding in a manner similar to previous section, we obtain

$$Q'(1)=\{k/(\lambda+k\nu)\}[\lambda-\mu(1-q_0)-q_0+\{c-q_0/q_0(c\lambda,\mu,\nu,k)\}/(c-1)] \tag{B.2}$$

where

$$q_0(c\lambda,\mu,\nu,k)=\left[1+\sum_{n=1}^{k}(c\lambda)^n\prod_{r=1}^{n}\{1-(r-1)/k\}\Big/\{\prod_{r=1}^{n}(\mu+r\nu+c\overline{c^{r-1}-1}/\overline{c-1})\}\right]^{-1} \tag{B.3}$$

References

[1] Altman, E. and Yechiali, U. (2008). Infinite- Server Queues with System's Additional Tasks and Impatient Customers. Probability in the Engineering and Informational Sciences, Vol. 22, No. 4, pp: 477-493.

[2] Al-Seedy, R. O. and Al-Ibraheem, F. M. (2001). An Interarrival Hyperexponential Machine Interference with Balking, Reneging, State-dependent, Spares and an Additional Server for Longer Queues. International Journal of Mathematics and Mathematical Sciences, Vol. 27, No. 12, pp: 737-747.

[3] Al-Seedy, R. O., El-Sherbiny, A. A., El-Shehawy, S. A. and Ammar, S. I. (2009). Transient solution of the M/M/c queue with balking and reneging. Computers and Mathematics with Applications, Vol. 57, No. 8, pp: 1280-1285.

[4] Ancker, C. J. Jr. and Gafarian, A. V. (1963). Queuing Problems with Balking and Reneging I, Operations Research, Vol. 11, No. 1, pp: 88-100.

[5] Baccelli, F., Boyer, P. and Hebuterne, G. (1984). Single Server Queues with Impatients Customers. Advances in Applied Probability, Vol.16, No. 4, pp: 887-905.

[6] Barrer, D.Y. (1957a). Queuing with impatient customers and ordered Service. Operations Research, vol. 5, No. 5, pp: 650-656.

[7] Barrer, D.Y. (1957b). Queuing with impatient customers and indifferent clerks. Operations Research, vol. 5, No. 5, pp: 644-649.

[8] Boots, N.K. and Tijms, H. (1999). An M/M/c Queues with Impatient Customers. TOP, Vol. 7, No. 2, pp: 213-220.

[9] Brandt, A. and Brandt, M. (1999). On a Two-Queue Priority System with Impatience and its Application to a Call Centre. Methodology and Computing in Applied Probability, Vol. 1, No. 2, pp: 191-210.

[10] Choi, B. D., Kim, B. and Chung, J. (2001). M/M/1 Queue with Impatient Customers of Higher Priority. Queuing Systems, Vol. 38, No. 1, pp: 49-66.

[11] Choi. B. D., Kim, B. and Zhu, D. (2004). MAP/M/C Queue with Constant Impatients Times. Mathematics of Operations Research, Vol. 29, No. 2, pp: 309-325.

[12] Choudhury, A. (2008). Impatience in Single Server Queuing Model. American Journal of Mathematical and Management Sciences, Vol. 28, No. 1, pp: 177-211.

[13] Choudhury, A and Medhi, P. (2011). Balking and Reneging in multiserver Markovian Queuing System. International Journal of Mathematics in Operational Research, Vol. 3, No. 4, pp: 377-394.

[14] El-Paoumy, M. S. (2008a). On Poisson Arrival Queue: $M^x/M/2/N$ with Balking, Reneging and Heterogeneous Servers. Applied Mathematical Sciences, Vol. 2, No. 24, pp: 1169-1175.

[15] El-Paoumy, M. S. (2008b). On a Truncated Erlangian Queuing system with State-Dependent Service rate, Balking and Reneging. Applied Mathematical Sciences, Vol. 2, No. 24, pp: 1161-1167.

[16] El-Paoumy, M.S. and Ismail, M. M. (2009). On a truncated Erlang Queuing System with Bulk Arrivals, Balking and Reneging. Applied Mathematical Sciences, Vol. 3, No. 23, pp: 1103-1113.

[17] El-Sherbiny, A. A. (2008). The Non-Truncated Bulk Arrival Queue $M^x/M/2/N$ with Reneging, Balking, State-Dependent and an Additional Server for Longer Queues. Applied Mathematical Sciences, Vol. 2, No. 15, pp: 747-752.

[18] Haghighi, A. Montazer, Medhi, J. and Mohanty, S. G. (1986). On a multi server Markovian queuing system with Balking and Reneging. Computer and Operational Research, Vol. 13, No. 4, pp: 421-425.

[19] Haight, F.A. (1957). Queuing with Balking. Biometrika, Vol. 44, No. 3/4, pp: 360-369.

[20] Haight, F.A. (1959). Queuing with Reneging. Metrika, Vol. 2, No. 1, pp: 186-197.

[21] Jouini, O., Aksin, Z. and Dallery, Y. (2008). Call Centers with Delay Information: Models and Insights. Downloaded from the site http://www.uclouvain.be/cps/ucl/doc/core/documents/OR_jouini.pdf(accessed on 03.08.2010).

[22] Ke, J. C. and Wang, K-H. (1999). Cost analysis of the M/M/R machine repair problem with balking, reneging and server breakdowns. The Journal of the Operational Research Society, Vol. 50, No. 3, pp: 275-282.

[23] Kok, A.G.D and Tijms, H. (1985). A Queuing System with Impatient customers. Journal of Applied Probability, Vol. 22, No. 3, pp: 688-696.

[24] Liu, L, Kashyap, B. R. K. and Templeton, J. G. C. (1987). The service system $M/M^R/\infty$ with impatient customers. Queuing Systems, Vol. 2, No. 4, pp: 363-372.

[25] Martin, M. and Artalejo, J. R. (1995). Analysis of an M/G/1 Queue with two types of Impatient units. Advances in Applied Probability, Vol. 27, No. 3, pp: 840-861.

[26] Medhi, J (1994). Stochastic Processes, Second Edition, Wiley Eastern Limited, India.

[27] Pazgal, A. and Radas, S. (2008). Comparison of customer balking and reneging behavior to queuing theory predictions: An experimental study. Computers and Operations Research, Vol. 35, No. 8, pp: 2537-2548.

[28] Shawky, A. I. (1997). The single server machine interference model with balking, reneging and an additional server for longer queues. Microelectronics Reliability, Vol. 37, No.2, pp: 355-357.

[29] Shawky, A. I. and El- Paoumy, M. S. (2000). The Interarrival Hyperexponential Queues: $H_k/M/C/N$ with Balking and Reneging. Stochastics An International Journal of Probability and Stochastic Processes, Vol. 69, No. 1 and 2, pp: 67-76.

[30] Shawky, A. I. and El- Paoumy, M. S. (2008). The Truncated Hyper- Poisson Queues: $H_k/M^{a,b}/c/N$ with Balking, Reneging and General Bulk- Service Rule. Yugoslav Journal of Operations Research, Vol. 8, No. 1, pp: 23-36.

[31] Singh, C.J., Kumar, B. and Jain, M. (2007). Single server interdependent queuing model with controllable arrival rates and reneging. Pakistan J. Statistics, Vol. 23, No. 2, pp: 171-178.

[32] Taha, H. A., (2003). Operations Research, Fifth Edition, Prentice Hall of India Private Limited.

[33] Ward, A. R. and Bambos, N. (2003). On Stability of Queuing Networks with Job Deadlines. Journal of Applied Probaboility, Vol. 40, No. 2, pp: 293-304.

[34] Xiong, W. and Altiok, T. (2009). An approximation for multi-server queues with deterministic reneging times. Annals of Operations Research, Vol. 172, No. 1, pp: 143-151.

[35] Yechiali, U. (2007). Queues with system disasters and impatient customers when system is down. Queueing Systems: Theory and Applications. Vol. 56, No. 3-4, pp: 195-202.

[36] Yue, D., Yue, W., Yu, J. and Tian, R. (2009). A heterogeneous two-server queuing system with balking and server breakdowns. Paper presented in the Eight International Symposium on Operations Research and its Applications (ISORA'09), Zhangjiajie, Chaina, Sept 20-22. Downloaded from the site http://www.aporc.org/LNOR/10/ISORA2009F31.pdf. (accessed on 04.08.2010).

[37] Zhang, Y., Yue, D. and Yue, W. (2005). Analysis of an M/M/1/N queue with balking, reneging and server vacations. Paper presented in the Vth. International Symposium on OR and its Application, 2005 and downloaded from http://www.aporc.org/LNOR/5/ISORA2005F04.pdf. (accessed on 04.08.2010).