Least squares estimation in a simple random coefficient autoregressive model.*

Søren Johansen † and Theis Lange‡
University of Copenhagen and CREATES

January 25, 2013

Abstract

The question we discuss is whether a simple random coefficient autoregressive model with infinite variance can create the long swings, or persistence, which are observed in many macro economic variables. The model is defined by $y_t = s_t \rho y_{t-1} + \varepsilon_t$, $t = 1, \ldots, n$, where s_t is an i.i.d. binary variable with $p = P(s_t = 1)$, independent of ε_t i.i.d. with mean zero and finite variance. We say that the process y_t is persistent if the autoregressive coefficient $\hat{\rho}_n$ of y_t on y_{t-1}, is close to one. We take $p < 1 < p\rho^2$ which implies $1 < \rho$ and that y_t is stationary with infinite variance. Under this assumption we prove the curious result that $\hat{\rho}_n \overset{P}{\to} \rho^{-1}$. The proof applies the notion of a tail index of sums of positive random variables with infinite variance to find the order of magnitude of $\sum_{t=1}^n y_{t-1}^2$ and $\sum_{t=1}^n y_t y_{t-1}$ and hence the limit of $\hat{\rho}_n$.

Keywords: Time series, explosive processes, bubble models, stable limits

JEL Classification: C32.

1 Introduction

A simple case of a random coefficient autoregressive (RCA) model, see Nicholls and Quinn (1982), is given by

$$y_t = s_t \rho y_{t-1} + \varepsilon_t, \; t = 1, \ldots, n,$$

*We are grateful to Roman Frydman and Michael Goldberg for suggesting the problem and for financial support from the Center for Imperfect Knowledge Economics, University of Copenhagen (funded by INET, Institute for New Economic Thinking) and the Center for Research in Econometric Analysis of Time Series (CREATES, funded by the Danish National Research Foundation). The anonymous referees and the editors Allan Timmermann and Herman van Dijk made many useful suggestions which helped relate the results to existing literature on RCA model.

†Address: Department of Economics, University of Copenhagen, Øster Farimagsgade 5, DK-1353 Copenhagen K. Denmark. Email: soren.johansen@econ.ku.dk

‡Address: Institute of Public Health, University of Copenhagen, Øster Farimagsgade 5, DK-1014 Copenhagen K. Email: t.lange@biostat.ku.dk

where ε_t are i.i.d. $(0, \sigma^2)$ and independent binary variables s_t for which

$$P(s = 1) = p; P(s = 0) = 1 - p = q.$$

We assume $\rho > 1$, and for notational reasons that $s_0 = 0$ and $y_0 = \varepsilon_0$. The process y_t is explosive in the periods where $s_t = 1$ and creates an excursion which stops once $s_t = 0$, an event that has probability $q = 1 - p$.

We define persistence of y_t as a value close to one of the limit of the autoregressive estimator

$$\hat{\rho}_n = \frac{\sum_{t=1}^n y_t y_{t-1}}{\sum_{t=1}^n y_{t-1}^2}. \tag{1}$$

Frydman and Goldberg (2007) raised the question if such a RCA model with infinite variance could create the long swings, or persistence, which is typical of macro variables and often associated with unit root processes.

Our main result is therefore to analyse the above question and we prove that if the process is stationary, $p < 1$, and has infinite variance, $1 < p\rho^2$, then $\hat{\rho}_n$ converges to $\rho^{-1} < p^{1/2} < 1$. General inference on the parameters in this model is not attempted in the present paper, even though we comment below on some of the results from the literature for the case of a finite variance, $p\rho^2 < 1$.

The general RCA model has been studied by many authors, see Andél (1976), Quinn and Nicholls (1981), Quinn (1982), the monograph by Nicholls and Quinn (1982), Lux and Sornette (2002), and Aue, Horvath, Steinebach (2006). The simple model studied here was suggested by Blanchard and Watson (1982) as a model for bubbles in the economy,but the discussion of Diba and Grosmann (1982) indicates that only positive bubbles can occur in the economy, and since rational bubbles must have expected initial value of zero the impossibility of negative bubbles also rules out positive bubbles: thus we shall not use such an interpretation of the model.

2 Review of the properties of $\hat{\rho}_n$ for the case of stationary y_t with finite variance

Many results for stationary processes with finite variance, that is $p < 1$ and $p\rho^2 < 1$, can be found as special cases of the general theory of the RCA models. Conditions for the existence of a stationary solution and finite moments were given in Nicholls and Quinn (1982), see also Aue, Horvath and Steinebach (2006) who give minimal conditions. For the present model, $p < 1$ implies the existence of a stationary solution, which has finite mean if $p\rho < 1$, and finite variance if $p\rho^2 < 1$. For a finite variance process the autoregressive coefficient, $\hat{\rho}_n$, was shown to converge in probability to $p\rho$ in Nicholls and Quinn (1982), who also proved asymptotic normality and found the asymptotic variance, if $E(y_t^4) < \infty$. Quinn and Nicholls (1981) suggested conducting inference using a quasi Gaussian likelihood based upon the conditional mean $E(y_t|y_{t-1}) = \rho p y_{t-1}$ and variance $Var(y_t|y_{t-1}) = pq\rho^2 y_{t-1}^2 + \sigma^2$, and proved consistency and asymptotic normality under the assumption $p\rho^2 < 1$. This was later analysed under weaker conditions by Aue, Horvath and Steinebach (2006). There seems to be no results for the likelihood based upon the assumption that ε_t are i.i.d. Gaussian. Finally Lux and

Sornette (2002) show, using results of Kesten (1973) and Goldie (1991), that if $p < 1 < \rho$, then the limit distribution of y_t has heavy tails, which in this case means $P(y_t \geq x) \approx cx^{-\alpha}$ with the tail index $\alpha = -\ln p / \ln \rho$.

3 Convergence of $\hat{\rho}_n$ for the case of stationary y_t with infinite variance, $p < 1 < p\rho^2$

For the stationary and infinite variance case, $p < 1 < p\rho^2$, we have not found any results in the literature, and the analysis of $\hat{\rho}_n$ in this section seems to be new.

In Subsection 3.1 we analyze the process and its product moments. We show how y_t can be represented as a sum of independent components belonging to successive periods, and that the main term of the product moments is a sum of positive random variables with infinite variance. For such processes we apply the concept of tail index and the theory in Feller (1971), to find the order of magnitude of the product moments. Once the product moments are analysed, the proof in Subsection 3.2 of the main result in Theorem 1, see (1) and (9), $\hat{\rho}_n \xrightarrow{P} \rho^{-1}$ is easily established. We end the section in Subsection 3.3 by a simulation illustrating the convergence result.

3.1 Representation of the process and product moments

If $s_t = 0$ we say that a collapse has occurred, and we let N_n denote the number of collapses before time n, and define the time points of the collapses as

$$0 = T_0^* < T_1^* < T_2^* < \cdots < T_{N_n}^* \leq n < T_{N_n+1}^*.$$

Here $N_n = \sum_{t=1}^n (1 - s_t)$ is the number of collapses before or at n, and we let $T_i = T_i^* - T_{i-1}^*$ be the length of the periods, $i = 1, 2, \ldots$. Note that if $s_{t_0} = s_{t_0+1} = 0$, then the corresponding period has length one. The last period before n is of length $n - T_{N_n}^*$. The variables T_i are independent and have the same geometric distribution

$$P(T_i = k) = p^{k-1}q, \quad k = 1, 2, \ldots$$

Using the distribution of ε_1 we now construct a double array of i.i.d. $(0, \sigma^2)$ random variables $\varepsilon_{it}, i = 1, 2, \ldots$ and $t = 0, 1, \ldots$, and construct the process y_t as follows. In the first period we use $\varepsilon_{1t}, t = 0, 1, \ldots$ and find for $t = 1, \ldots, T_1 - 1$ that, starting at $y_0 = \varepsilon_{10}$, we get, because $s_1 = \cdots = s_{T_1-1} = 1$, that

$$y_t = \sum_{v=0}^{t} \rho^{t-v} \varepsilon_{1v} = \rho^t \sum_{v=0}^{\infty} \rho^{-v} \varepsilon_{1v} - \rho^{-1} \sum_{v=0}^{\infty} \rho^{-v} \varepsilon_{1,v+t+1} = \rho^t Z_1 - \rho^{-1} Z_{1t}, \tag{2}$$

where Z_{1t} has the same distribution as Z_1 with $E(Z_1) = 0$ and $Var(Z_1) = \sigma^2/(1 - \rho^{-2})$. The last observation of the first period, y_{T_1}, has $s_{T_1} = 0$, and we define

$$y_{T_1} = \varepsilon_{20},$$

which acts as initial value for the second period, where $\varepsilon_{2t}, t = 0, 1, \ldots$ are used to construct the process.

Similar expressions can be found for the i'th period, $t = T^*_{i-1} + 1, \ldots, T^*_i - 1$

$$y_t = \sum_{v=0}^{t-T^*_{i-1}} \rho^{t-T^*_{i-1}-v}\varepsilon_{iv} = \rho^{t-T^*_{i-1}}\sum_{v=0}^{\infty}\rho^{-v}\varepsilon_{iv} - \rho^{-1}\sum_{v=0}^{\infty}\rho^{-v}\varepsilon_{i,v+t-T^*_{i-1}+1} = \rho^{t-T^*_{i-1}}Z_i - \rho^{-1}Z_{it}, \quad (3)$$

and define $y_{T^*_i} = \varepsilon_{i+1,0}$. Note that by using a double array $\{\varepsilon_{it}\}$ we have made sure that T_i and Z_i are independent and i.i.d. We next apply this representation to find an approximation to the autoregressive estimator. Note also that the process y_t is recurrent in the sense that it starts afresh each time $s_t = 0$, so the periods are independent.

Lemma 1 *The product moments have the representation*

$$\sum_{t=1}^{n} y_{t-1}^2 = \frac{1}{\rho^2 - 1}[\sum_{i=1}^{N_n}\rho^{2T_i}Z_i^2 + \sum_{i=1}^{N_n+1}A_i],$$
$$\sum_{t=1}^{n} y_{t-1}y_t = \frac{\rho^{-1}}{\rho^2 - 1}[\sum_{i=1}^{N_n}\rho^{2T_i}Z_i^2 + \sum_{i=1}^{N_n+1}B_i],$$

where the remainder terms satisfy

$$E(|A_i|^{\xi} + |B_i|^{\xi}|T_i) \le c\rho^{\xi T_i}, \; i = 1, \ldots, N_n + 1, \quad (4)$$

for $0 < \xi \le 1$. *It follows that the estimator based on* $y_t, t = 0, 1, \ldots, n$, *has the representation*

$$\hat{\rho}_n = \frac{\sum_{t=1}^{n} y_t y_{t-1}}{\sum_{t=1}^{n} y_{t-1}^2} = \rho^{-1}\frac{\sum_{i=1}^{N_n}\rho^{2T_i}Z_i^2 + \sum_{i=1}^{N_n+1}B_i}{\sum_{i=1}^{N_n}\rho^{2T_i}Z_i^2 + \sum_{i=1}^{N_n+1}A_i}. \quad (5)$$

Proof. We find from (2) that

$$\sum_{t=1}^{T_1} y_{t-1}^2 = \sum_{t=1}^{T_1}(\rho^{t-1}Z_1 - \rho^{-1}Z_{1t-1})^2 = \frac{\rho^{2T_1}}{\rho^2 - 1}Z_1^2 + A_1$$
$$A_1 = \frac{-1}{\rho^2 - 1}Z_1^2 + \rho^{-2}\sum_{t=1}^{T_1}Z_{1t-1}^2 - 2Z_1\sum_{t=1}^{T_1}\rho^{t-2}Z_{1t-1}.$$

We need the well-known inequality valid for $a \ge 0$ and $b \ge 0$

$$(a+b)^{\xi} = b^{\xi} + \xi\int_0^a (b+x)^{\xi-1}dx \le b^{\xi} + \xi\int_0^a x^{\xi-1}dx = b^{\xi} + a^{\xi}, \; 0 < \xi \le 1. \quad (6)$$

This implies that

$$E(|A_1|^{\xi}|T_1) \le a_1 E(Z_1^{2\xi}) + a_2 T_1 E(Z_1^{2\xi}) + a_3\rho^{\xi T_1}E(Z_1^{2\xi}) \le c\rho^{\xi T_1},$$

which shows (4) for A_1. Here the last term is evaluated as follows

$$E(|\sum_{t=1}^{T_1}\rho^{t-2}Z_{1t-1}Z_1|^\xi|T_1) \leq \sum_{t=1}^{T_1}\rho^{\xi(t-2)}E(|Z_{1t-1}Z_1|^\xi) \leq \sum_{t=1}^{T_1}\rho^{\xi(t-2)}E(Z_{1t-1}^{2\xi})^{1/2}E(Z_1^{2\xi})^{1/2}$$
$$\leq E(Z_1^{2\xi})\sum_{t=1}^{T_1}\rho^{\xi(t-2)} \leq c\rho^{\xi T_1}E(Z_1^{2\xi}).$$

The same proof can be used for $A_i, i = 2, \ldots, N_n$, and it is seen that the bound c does not depend on i. For $i = N_n + 1$, we have $A_{N_n+1} = \sum_{t=T^*_{N_n}+1}^{n} y_{t-1}^2 \leq \sum_{t=T^*_{N_n}+1}^{T^*_{N_n+1}} y_{t-1}^2$ and the same proof works.

Next we find, noting that $y_{T_1} = \varepsilon_{20}$, that

$$\sum_{t=1}^{T_1} y_{t-1}y_t = \sum_{t=1}^{T_1-1}(\rho^{t-1}Z_1 - \rho^{-1}Z_{1t-1})(\rho^t Z_1 - \rho^{-1}Z_{1t}) + (\rho^{T_1-1}Z_1 - \rho^{-1}Z_{1T_1-1})\varepsilon_{20}$$
$$= \frac{\rho^{2(T_1-1)}}{\rho^2-1}\rho Z_1^2 + B_1$$
$$B_1 = \frac{-\rho Z_1^2}{\rho^2-1} + \rho^{-2}\sum_{t=1}^{T_1-1}[Z_{1t-1}Z_{1t} - \rho^{t+1}Z_{1t-1}Z_1 - \rho^t Z_1 Z_{1t}] + \rho^{-1}(\rho^{T_1}Z_1 - Z_{1T_1-1})\varepsilon_{20}.$$

The same proof shows that $E(|B_1|^\xi|T_1)$ satisfies (4). The terms $B_i, i = 2, \ldots, N_n + 1$ can be handled similarly. ■

In the following we apply the assumption that $p\rho^2 > 1$ so that the variance of y_t is infinite. We want to find the limit of the regression estimator given in (5) and for that we need the order of magnitude of the main term $\sum_{i=1}^{N_n}\rho^{2T_i}Z_i^2$ and bounds on the remainder terms.

We find the order of magnitude of these terms using the theory of sums of positive random variables with infinite variance, see for instance Feller (1971, Chapter IX, Section 8). It turns out that it is not possible to normalize the main term to convergence, because of the discrete nature of the geometric distribution, but instead we bound T_i using an auxiliary exponentially distributed variable U_i, for which $\sum_{i=1}^{N_n}\rho^{2U_i}Z_i^2$ can be normalized to convergence.

Let U_i be i.i.d. exponentially distributed variables with parameter $\lambda = -\log p$, and note that the waiting times can be represented as one plus the integer part of U_i :

$$T_i = [U_i] + 1,$$

since as required

$$P(T_i = k) = P(k-1 \leq U_i < k) = e^{-\lambda(k-1)} - e^{-\lambda k} = p^{k-1}q.$$

We have the evaluations

$$U_i \leq T_i \leq U_i + 1,$$

and hence the bounds for any finite m

$$\sum_{i=1}^{m}\rho^{2U_i}Z_i^2 \leq \sum_{i=1}^{m}\rho^{2T_i}Z_i^2 \leq \rho^2\sum_{i=1}^{m}\rho^{2U_i}Z_i^2. \tag{7}$$

This shows that it is enough to find the order of magnitude of $\sum_{i=1}^m \rho^{2U_i} Z_i^2$, and for this we need the so-called tail index of a positive random variable. We find from

$$P(\rho^{2U} > x) = P(U \geq \frac{\log x}{2\log\rho}) = e^{-\frac{\lambda \log x}{2\log\rho}} = x^{\frac{\log p}{2\log\rho}} = x^{-\alpha/2},\ \alpha = -\frac{\log p}{\log\rho} > 0,$$

that

$$P(\rho^{2U} Z^2 > x) = E[P(\rho^{2U} Z^2 > x|Z)] = E(xZ^{-2})^{-\alpha/2} = x^{-\alpha/2} E(Z^\alpha).$$

Thus the tails of the distributions of ρ^{2U} and $\rho^{2U} Z^2$ decrease as $x^{-\alpha/2}$, and we say that the tail index of ρ^{2U} and $\rho^{2U} Z^2$ is $\alpha/2$. Note that $p < 1 < p\rho^2$ implies $0 < \alpha < 2$.

3.2 Convergence of $\hat{\rho}_n$ in the infinite variance case

With these tools we can now prove the main result.

Theorem 1 *For $p < 1 < p\rho^2$, and $E(\varepsilon_t^2) < \infty$, it holds that*

$$m^{-2/\alpha} \sum_{i=1}^m \rho^{2U_i} Z_i^2 \xrightarrow{d} V_{\alpha/2}, m \to \infty, \tag{8}$$

where $Z_i = \sum_{v=0}^\infty \rho^{-v} \varepsilon_{iv}$ is given by (3) and $V_{\alpha/2}$ is a stable distribution of index $\alpha/2$. This implies that

$$\hat{\rho}_n = \frac{\sum_{t=1}^n y_t y_{t-1}}{\sum_{t=1}^n y_{t-1}^2} \xrightarrow{P} \rho^{-1}. \tag{9}$$

Proof. *Proof of* (8): By construction $\rho^{2U_i} Z_i^2, i = 1, \ldots, m$ are i.i.d. and Z_i is independent of U_i, so the tail index of $\rho^{2U_i} Z_i^2$ is $\alpha/2$. The result (8) now follows from Feller (1971, Theorem 2 p. 305, and (8.14)).

Proof of (9): We use the representation (5) for $\hat{\rho}_n$ but first replace the stochastic N_n by the nonstochastic m and show that

$$R_m = \frac{m^{-2/\alpha} \sum_{i=1}^m \rho^{2T_i} Z_i^2 + m^{-2/\alpha} \sum_{i=1}^{m+1} B_i}{m^{-2/\alpha} \sum_{i=1}^m \rho^{2T_i} Z_i^2 + m^{-2/\alpha} \sum_{i=1}^{m+1} A_i} \xrightarrow{P} 1,\ m \to \infty.$$

From the bounds (7) and (8) it follows that it is enough to show that

$$m^{-2/\alpha} \sum_{i=1}^{m+1} (|A_i| + |B_i|) \xrightarrow{P} 0,\ m \to \infty.$$

The expectation of this need not be finite when $\rho > 1$, but because $0 < \alpha < 2$ we can choose ξ so that $\alpha/2 < \xi < \min(1, \alpha)$. Then, because $\lim_{\xi \to \alpha/2} p\rho^\xi = p\rho^{-\frac{\log p}{2\log\rho}} = p^{1/2} < 1$, we can choose ξ so that $p\rho^\xi < 1$ and therefore $E(\rho^{\xi T_1}) = \sum_{k=0}^\infty (\rho^\xi p)^k p^{-1} q < \infty$. From (4) and (6) we find

$$E(m^{-2/\alpha} \sum_{i=1}^{m+1} (|A_i| + |B_i|))^\xi \leq m^{-2\xi/\alpha} (m+1) E(|A_1|^\xi + |B_1|^\xi) \leq c m^{1-2\xi/\alpha} E(\rho^{\xi T_1}) \to 0,$$

because $\alpha/2 < \xi$ and $E(\rho^{\xi T_i}) < \infty$ when $\xi < \alpha$.

Next we want to prove that we can replace m by $N_n = \sum_{t=1}^{n}(1 - s_t)$. By the law of large numbers we have $n^{-1}N_n \xrightarrow{P} q$ so that for given $\varepsilon > 0, \delta > 0$ we can choose an n_0 so that for $n \geq n_0$ we have with probability greater than $1 - \delta$

- $[n(q - \varepsilon)] \leq N_n \leq [n(q + \varepsilon)]$
- $(\frac{[n(q+\varepsilon)]}{[n(q-\varepsilon)]})^{2/\alpha} \leq 1 + \varepsilon$
- $[n(q + \varepsilon)]^{-2/\alpha} \sum_{i=1}^{[n(q+\varepsilon)]+1}(|A_i| + |B_i|) \leq \varepsilon/(1 + \varepsilon)$.

Then it follows that with probability greater than $1 - \delta$

$$N_n^{-2/\alpha} \sum_{i=1}^{N_n+1}(|A_i| + |B_i|) \leq (\frac{[n(q + \varepsilon)]}{[n(q - \varepsilon)]})^{2/\alpha}[n(q + \varepsilon)]^{-2/\alpha} \sum_{i=1}^{[n(q+\varepsilon)]+1}(|A_i| + |B_i|) \leq \varepsilon,$$

so that $N_n^{-2/\alpha} \sum_{i=1}^{N_n+1}(|A_i| + |B_i|) \xrightarrow{P} 0$. Together with (5) this proves (9) and completes the proof of Theorem 1. ■

3.3 A numerical illustration of the convergence of $\hat{\rho}_n$ to ρ^{-1}

Figure 1 shows simulated values of the median (and 2.5% and 97.5% quantiles) of 1000 simulations of $\hat{\rho}_n$ for $n = 10.000$, $\rho = 1.032$ and p in the range 0.9 to 1. It is seen that for $p < \rho^{-2} = 0.939$ (the finite variance case) the limit is almost proportional to p with slope ρ and for $p > \rho^{-2} = 0.939$ (the infinite variance case) the limit is almost constantly equal to $\rho^{-1} = 0.969$, which illustrates the result that $\hat{\rho}_n \xrightarrow{P} \min(p\rho, \rho^{-1})$. In addition the simulations indicates that the rate of convergence in the infinite variance case, $p > 0.939$, is faster than $n^{1/2}$.

Figure 1 here

4 Conclusion

We have considered a simple RCA model with the purpose of finding out if the autoregressive coefficient $\hat{\rho}_n$ can be close to one when the process is stationary, but explosive in periods and the variance is infinite. We have shown that $\hat{\rho}_n \xrightarrow{P} \rho^{-1} < p^{1/2} < 1$, and hence only equal to one if $\rho = p = 1$, which is the case of a random walk, and in this sense the model cannot explain the long swings seen in macro variables.

5 References

1. Andél, J., 1976. Autoregressive series with random parameters. Mathematisches Operationsforschung und Statistik 7, 735–41.

2. Aue, A., Horváth, L., Steinebach, J., 2006. Estimation in random coefficient autoregressive models. Journal of Time Series Analysis 27, 61–76.

3. Blanchard, O., Watson, M., 1982. Bubbles, Rational Expectations, and Financial Markets. In Paul Wachter (ed.) Crises in the Economic and Financial Structure. Lexington, MA: Lexington Books, pp. 295–315.

4. Diba, B.T., Grossman, H.I., 1987. On the inception of rational bubbles. Quarterly Journal of Economics 102, 697-700.

5. Feller, W., 1971. An Introduction to Probability Theory and Its Applications. Vol. II. Second edition. Wiley, New York.

6. Frydman, R., Goldberg, M.D., 2007. Imperfect Knowledge Economics. Princeton University Press, Princeton.

7. Goldie, C.M., 1991. Implicit Renewal Theory and Tails of Solutions of Random Equations. Annals of Applied Probability 1, 126-66.

8. Kesten, H., 1973. Random Difference Equations and Renewal Theory for Products of Random Matrices. Acta Mathematica 131, 207-48.

9. Lux. T., Sornette, D., 2002. On Rational Bubbles and Fat Tails. Journal of Money, Credit, and Banking 34, 589-610.

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

11. Quinn, B. G., 1982. A note on the existence of strictly stationary solutions to bilinear equations. Journal of Time Series Analysis 3, 249–252.

12. Quinn, B. G., Nicholls, D.F., 1981. The estimation of random coefficient autoregressive models II. Journal of Time Series Analysis 2, 185–203.

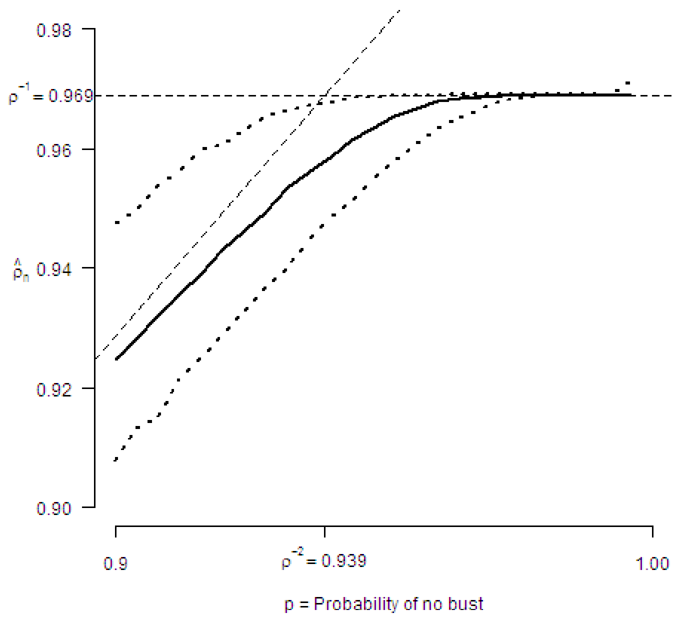

Figure 1: The figure shows the result of 1.000 simulations for $n = 10.000$ of $\hat{\rho}_n$ for $\rho = 1.032$ and $0.9 < p < 1$. We have plotted the median and the 2.5% and 97.5% quantiles of the simulations.