Christopeit, Norbert; Massmann, Michael

Working Paper

A Note on an Estimation Problem in Models with Adaptive Learning

Tinbergen Institute Discussion Paper, No. 13-151/III

Provided in Cooperation with:
Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: Christopeit, Norbert; Massmann, Michael (2013) : A Note on an Estimation Problem in Models with Adaptive Learning, Tinbergen Institute Discussion Paper, No. 13-151/III, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:
http://hdl.handle.net/10419/87485

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
A Note on an Estimation Problem in Models with Adaptive Learning

Norbert Christopeit¹
Michael Massmann²

¹ University of Bonn, Germany;
² Faculty of Economics and Business Administration, VU University Amsterdam, and Tinbergen Institute, The Netherlands.
Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at [http://www.tinbergen.nl](http://www.tinbergen.nl)

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: [http://www.dsf.nl/](http://www.dsf.nl/)

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579
A note on an estimation problem in models with adaptive learning

Norbert Christopeit  Michael Massmann*
University of Bonn Vrije Universiteit Amsterdam
and Tinbergen Institute

September 24, 2013

Abstract

This paper provides an example of a linear regression model with predetermined stochastic regressors for which the sufficient condition for strong consistency of the ordinary least squares estimator by Lai & Wei (1982, *Annals of Statistics*) is not met. Nevertheless, the estimator is strongly consistent, as shown in a companion paper, cf. Christopeit & Massmann (2013b). This is intriguing because the Lai & Wei condition is the best currently available and is referred to as “in some sense the weakest possible”. Moreover, the example discussed in this paper arises naturally in a class of macroeconomic models with adaptive learning, the estimation of which has recently gained popularity amongst researchers and policy makers.

keywords: least-squares regression, stochastic regressors, strong consistency, minimal sufficient condition, adaptive learning.

JEL codes: C22, C51, D83

*corresponding author: Michael Massmann, Department of Econometrics, Vrije Universiteit, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands, phone: +31 (0)20 5986014, email: m.massmann@vu.nl
Motivation and result

Consider the general linear regression model

\[ y_t = \theta' z_t + \varepsilon_t, \quad t = 1, 2, \ldots, \] (1)

where the \( \varepsilon_t \) are unobservable errors, \( \theta = (\theta_1, \ldots, \theta_K)' \) is a vector of unknown parameters and \( y_t \) is the observed response to the inputs \( z_t = (z_{t1}, \ldots, z_{tK})' \). For serially uncorrelated \( \varepsilon_t \), the parameter vector \( \theta \) is usually estimated by the ordinary least squares (OLS) estimator

\[ \hat{\theta}_T = (Z'_T Z_T)^{-1} Z'_T y(T) \]

based on the observations \( y(T) = (y_1, \ldots, y_T)' \) and \( Z_T = (z_{tk})_{1 \leq t \leq T, 1 \leq k \leq K} \). For deterministic regressors \( z_{tk} \), there exists a well established theory for strong consistency (i.e. convergence with probability one) of the OLS estimator, cf. Anderson & Taylor (1976), Drygas (1976), Lai & Robbins (1977) and Lai, Robbins & Wei (1978, 1979), providing both sufficient and necessary conditions. For stochastic regressors, the situation is more delicate. The best result obtained so far is by Lai & Wei (1982a, 1982b) and draws on two assumptions:

(A1) \( (\varepsilon_T) \) is a martingale difference sequence with respect to some basic filtration \( (F_T) \).

(A2) \( (z_T) \) is a predetermined sequence with respect to \( (F_T) \), i.e. \( z_T \) is \( F_{T-1} \)-measurable for each \( T \).

Denote the sample second moment matrix of the regressors by

\[ M_T = Z'_T Z_T, \]

and let \( \lambda_{\text{max}}(T) \) and \( \lambda_{\text{min}}(T) \) be the maximal and the minimal eigenvalue of \( M_T \), respectively.

Lai & Wei (1982a) then prove the following.

**Theorem (Lai & Wei, 1982a)** If the disturbances satisfy the condition \( \sup_T E(\varepsilon_T^2 | F_{T-1}) < \infty \) a.s., a sufficient condition for strong consistency of the OLS estimators of \( \theta \) is the following:

\[ \lambda_{\text{min}}(T) \to \infty \quad \text{and} \quad \left[ \log \lambda_{\text{max}}(T) \right]^{1+\rho} = o\left( \lambda_{\text{min}}(T) \right) \quad \text{a.s.} \] (2)

for some \( \rho > 0 \). If, in addition, \( \sup_T E(\varepsilon_T^{2+\eta} | F_{T-1}) < \infty \) for some \( \eta > 0 \), then it suffices to require (2) for \( \rho = 0 \).

Lai & Wei (1982a) refer to the condition in (2) as “in some sense the weakest possible” (p. 155), since even a marginal violation of it can lead to inconsistency. For \( \rho = 0 \), they give an example to that effect. The purpose of this Note is to provide a different example, see Result...
below, which may be considered as a converse of that in Lai & Wei (1982) in the sense that, despite a marginal violation of (2), the strong consistency of $\hat{\theta}_T$ continues to hold. This example arises in a typical macroeconomic model in which agents are not assumed to be fully rational but to form expectations by means of an adaptive learning algorithm. Estimation of models in this model class have recently gained popularity amongst researchers and policy makers; see for instance the New Keynesian Phillips curve models estimated by Milani (2007) and Chevillon, Massmann & Mavroeidis (2010), the European Central Bank’s New Multi-Country Model by Dieppe, González Pandiella, Hall & Willman (2011), and the inflation model by Malmedier & Nagel (2012).

Specifically, consider the simple linear regression model

$$y_t = \delta + \beta a_{t-1} + \varepsilon_t,$$

where $\varepsilon_t$ is an i.i.d. Gaussian disturbance with mean 0 and variance $\sigma^2$, where the regressor $a_t$ is determined by the recursion

$$a_t = a_{t-1} + \gamma_t (y_t - a_{t-1}),$$

$a_0 = 0$, and where

$$\gamma_t = \frac{\gamma}{t}$$

is a weighting, or gain, sequence. It is assumed throughout that $\gamma > 0$ and $\beta < 1$. This model is derived from the more general specification $y_t = \delta w_t + \beta y_{t-1}^e + \varepsilon_t$ where $y_{t-1}^e$ denotes agents’ expectations of $y_t$ given the information set $\mathcal{F}_{t-1} = \sigma (y_s, s \leq t-1; w_s, s \leq t)$, see for instance the classical cobweb model in Bray & Savin (1986) or the Lucas (1973) aggregate supply model. Importantly, in the present setting, agents would no longer possess sufficient knowledge to form rational expectations $y_{t-1}^e = E(y_t|\mathcal{F}_{t-1})$ but, instead, are assumed to know only the structure of the rational expectations equilibrium $y_t = \alpha w_t + \varepsilon_t$ and to estimate the unknown parameter $\alpha$ recursively by means of a stochastic approximation algorithm such that their forecast is $y_{\hat{t}|t-1} = a_{t-1} w_t$. Assuming for analytical tractability that $w_t$ is constant and, then, setting $w = 1$ without loss of generality, the model in (3)-(4) obtains. The weighting sequence in (5) makes of the updating mechanism in (4) an instance of a so-called decreasing-gain recursion, since $\gamma_t \to 0$. The asymptotic properties of this system are analysed in detail in two companion papers, viz. Christopeit & Massmann (2013a, 2013b). The properties turn out to depend crucially on $c = \gamma (1 - \beta)$. In particular, Christopeit & Massmann (2013b) prove that the OLS estimators of $\delta$ and $\beta$ are strongly consistent for $c > 1/2$. Nevertheless, the following result shows that Lai & Wei’s condition (2) is violated; the proof is presented in the following section.
Result For the model in (3)-(5), the two eigenvalues of the sample second moment matrix $M_T$ are given by

$$
\lambda_{\min} = \frac{A_T + O(1)}{\sigma^2 + 1 + o(1)} (1 + o(1))
$$

$$
\lambda_{\max} = \frac{T}{T} \left[ T \left[ \sigma^2 + 1 + o(1) \right] (1 + o(1)) \right]
$$

where

$$A_T = \sum_{t=1}^{T} (a_t - \bar{a}_T)^2.
$$

When $c > 1/2$, although $\lambda_{\min}(T) \to \infty$ it turns out that

$$
\plim_{T \to \infty} \frac{\log \lambda_{\max}(T)}{\lambda_{\min}(T)} = \sigma^2 \frac{\alpha^2 + 1}{2c - 1}.
$$

Consequently, Lai & Wei’s condition (2) is violated for $c > 1/2$.

It will also be seen below that condition (2) does hold when $c < 1/2$; accordingly, the estimators in this case are strongly consistent, cf. Christopeit & Massmann (2013a). Investigation of the boundary case $c = 1/2$ is left to future research.

Proof and discussion

Calculating the eigenvalues

Model (3) is a special bivariate case of (1) with $\theta = (\delta, \beta)'$ and $z_t = (1, x_t)'$, where $x_t = a_{t-1}$.

The sample second moment matrix is, then,

$$M_T = T \begin{pmatrix} 1 & p_T \\ p_T & r_T \end{pmatrix}
$$

where

$$p_T = \bar{x}_T = \bar{x}_T - \frac{1}{T} \bar{a}_T,
$$

$$r_T = \bar{x}_T^2 = \bar{x}_T^2 - \frac{1}{T} \bar{a}_T^2.
$$

We adopt the usual notation $\bar{x}_T = \frac{1}{T} \sum_{t=1}^{T} x_t$ and $\bar{x}_T^2 = \frac{1}{T} \sum_{t=1}^{T} x_t^2$. Recall also that $a_0 = 0$. For ease of notation, we suppress the subscript $T$ of $p_T$ and $r_T$ for a moment. Simple calculation shows that the eigenvalues of $M_T$ are given by

$$\lambda_{\pm} = \frac{T}{2} \left[ r + 1 \pm \sqrt{(r + 1)^2 - 4 (r - p^2)} \right]
$$

$$= \frac{T}{2} \left[ r + 1 \pm \sqrt{(r - 1)^2 + 4p^2} \right].
$$
Therefore $\lambda_{\text{max}} \geq T|p|$. To obtain the exact rate, consider the expansion

$$
\lambda_{\text{max}} = \frac{T}{2} \left[ r + 1 + \sqrt{(r + 1)^2 - 4(r - p^2)} \right]
= \frac{T(r + 1)}{2} \left[ 1 + \sqrt{1 - 4\frac{r - p^2}{(r + 1)^2}} \right]
= \frac{T(r + 1)}{2} [1 + (1 - 2d - 2R(d))]
= T(r + 1) [1 - d - R(d)],
$$

(8)

where we have put

$$
d = d_T = \frac{r - p^2}{(r + 1)^2}
$$

(9)

and where $R(d)$ is the residual in the Taylor expansion of the square root function about one:

$$
\sqrt{1 - 4d} = 1 - 2d - \frac{2d^2}{(1 - 4\theta d)^{3/2}} = 1 - 2d - 2R(d),
$$

with $0 < \theta < 1$. Similarly, for $\lambda_{\text{min}},$

$$
\lambda_{\text{min}} = \frac{T}{2} \left[ r + 1 - \sqrt{(r + 1)^2 - 4(r - p^2)} \right]
= \frac{T(r + 1)}{2} \left[ 1 - \sqrt{1 - 4\frac{r - p^2}{(r + 1)^2}} \right]
= \frac{T(r + 1)}{2} [1 - (1 - 2d - 2R(d))]
= T(r + 1) [d + R(d)],
$$

(10)

for the same $d$ and $R(d)$ as above. Note that

$$
R(d) = O(d^2) \text{ as } d \to 0.
$$

(11)

**Asymptotics of eigenvalues**

It is shown in Christopeit & Massmann (2013a) that $a_T \to \alpha$ with probability one. Therefore, $p_T$ and $r_T$ in (6)-(7) satisfy, with probability one,

$$
p_T = \bar{\pi}_T + O(T^{-1})
$$

$$
r_T = \bar{a}_T^2 + O(T^{-1})
$$

as well as

$$
r_T - p_T^2 = \bar{a}_T^2 - \bar{\pi}_T^2 + O(T^{-1}).
$$

(12)

Due to the convergence of $a_T$ it follows that $p_T = \alpha + o(1)$ and $r_T = \alpha^2 + o(1)$ such that $d_T$ and $R(d_T)$ in (9) and (11) satisfy, almost surely,

$$
d_T = o(1),
$$

$$
R(d_T) = o(1)
$$
and

\[ TR(d_T) = Td_T \cdot O(d_T) = Td_T \cdot o(1). \]

To ascertain the convergence rate of \( d_T \), introduce

\[ A_T = \sum_{t=1}^{T} (a_t - \pi_T)^2 = T \left( \alpha_T^2 - \pi_T^2 \right). \]

This is the quantity whose behaviour turns out crucial for the consistency of the OLS estimator of \( \delta \) and \( \beta \), see Christopeit & Massmann (2013a). By virtue of (12),

\[ r_T - p_T^2 = \frac{1}{T} A_T + O(T^{-1}), \]

such that \( Td_T \) may be written as

\[ Td_T = \frac{A_T + O(1)}{(r_T + 1)^2} = \frac{A_T + O(1)}{(\alpha^2 + 1 + o(1))} T = \frac{A_T}{(\alpha^2 + 1)^2} \left[ 1 + O(A_T^{-1}) + o(1) \right]. \]

Therefore, by (10),

\[
\begin{align*}
\lambda_{\min} &= (r_T + 1) T \left[ d_T + R(d_T) \right] \\
&= (r_T + 1) T d_T \left[ 1 + o(1) \right] \\
&= \left[ \alpha^2 + 1 + o(1) \right] \frac{A_T}{(\alpha^2 + 1)^2} \left[ 1 + O(A_T^{-1}) + o(1) \right] \\
&= \frac{A_T}{\alpha^2 + 1} \left[ 1 + O(A_T^{-1}) + o(1) \right]
\end{align*}
\]

and by (8),

\[
\begin{align*}
\lambda_{\max} &= T (r_T + 1) \left[ 1 - d_T - R(d_T) \right] \\
&= T \left[ \alpha^2 + 1 + o(1) \right] (1 + o(1)).
\end{align*}
\]

As a consequence,

\[
\frac{\log \lambda_{\max}(T)}{\lambda_{\min}(T)} = (\alpha^2 + 1) \frac{\log T + O(1)}{A_T [1 + O(A_T^{-1}) + o(1)]}. \tag{13}
\]

It is shown in Christopeit & Massmann (2013a, equation (B.14)) that, for \( c > \frac{1}{2} \),

\[
\frac{A_T}{\log T} \xrightarrow{p} \frac{\gamma^2 \sigma^2}{2c - 1}. \tag{14}
\]

Since \( A_T \) is monotone increasing, this implies that \( \lim_{T \to \infty} A_T = \infty \text{ a.s.} \). It is not known whether (14) also holds with probability one. As a consequence of (14),

\[
\lim_{T \to \infty} \frac{\log \lambda_{\max}(T)}{\lambda_{\min}(T)} = (\alpha^2 + 1) \frac{2c - 1}{\gamma^2 \sigma^2}. \]

In other words, Lai & Wei’s condition (2) is violated for \( c > \frac{1}{2} \), as claimed in Result. Nevertheless, Christopeit & Massmann (2013b) prove that strong consistency of the OLS estimator for \( \delta \) and \( \beta \) does hold for \( c > \frac{1}{2} \).
Remark 1 For \( c < 1/2 \), it is shown in Christopeit & Massmann (2013a, equation (B.22)) that the exact divergence rate of \( A_T \) is

\[
A_T = \frac{\gamma^2 v^2}{1 - 2c} T^{1 - 2c} (1 + o(1)) \quad \text{a.s.,}
\]

where \( v \) is a random variable which is nonzero with probability one. Hence, from (13),

\[
\lim_{T \to \infty} \frac{\log \lambda_{\max}(T)}{\lambda_{\min}(T)} = 0 \quad \text{a.s.}
\]

such that Lai & Wei’s condition (2) is satisfied for \( c < 1/2 \).

Remark 2 The divergence rate of \( A_T \) for the boundary case \( c = 1/2 \) is not known. Hence the question of whether or not Lai & Wei’s condition (2) is satisfied is left to future research.

Remark 3 For the bivariate case of simple regression models like (3), a somewhat weaker condition for strong consistency of the slope estimator is given in Lai & Wei (1982a), namely,

\[
\frac{A_T}{\log T} \to \infty \quad \text{a.s.}
\]  

(15)

In view of (14), (15) is not satisfied for \( c > 1/2 \). For \( c < 1/2 \), it is since (15) is implied by (2).

References


