Dynamic censored regression and the Open Market Desk reaction function

Robert de Jong* Ana María Herrera†

June 30, 2005

Abstract

The censored regression model and the Tobit model are standard tools in econometrics. This paper provides a formal asymptotic theory for dynamic time series censored regression when lags of the dependent variable have been included among the regressors. We derive fading memory properties of the model under the assumption that the regression error is strong mixing. We show the formal asymptotic correctness of conditional maximum likelihood estimation of the dynamic Tobit model, and the correctness of Powell’s least absolute deviations procedure for the estimation of the dynamic censored regression model. The paper is concluded with an application of the dynamic censored regression methodology to temporary purchases of the Open Market Desk.

1 Introduction

The censored regression model and the Tobit model are standard tools in econometrics. In a time series framework, censored variables arise when the dynamic optimization behavior of

*Department of Economics, Ohio State University, 429 Arps Hall, Columbus, OH 43210, email dejong@econ.ohio-state.edu
†Department of Economics, Michigan State University, 215 Marshall Hall, East Lansing, MI 48824, email herrer20@mail.msu.edu. We thank Selva Demiralp and Oscar Jordà for making their data available, and Bruce Hansen for the use of his code for kernel density estimation.
a firm or individual leads to a corner response for a significant proportion of time. Censored regression models apply to the level of open market operations or foreign exchange intervention carried out by a central bank, to the clearing price in commodity markets where the government imposes price floors, to the quantity of imports and exports of goods subject to quotas, and to numerous other series.

The asymptotic theory for the Tobit model in cross-section situations has long been understood; see for example the treatment in Amemiya (1973). However, there appears to be no asymptotic theory for the time series dynamic censored regression model where lagged values of the dependent variable are included among the regressors, and this paper seeks to provide such a treatment. The primary analytical issue is to show that under some conditions, the dynamic censored regression model as defined below is stationary. This issue is reminiscent of showing the stationarity of ARMA models under conditions about the roots of the AR lag polynomial. The dynamic censored regression model under consideration is

\[ y_t = \max(0, \sum_{i=1}^{p} \rho_i y_{t-i} + \gamma' x_t + \varepsilon_t), \]  

where \( x_t \) denotes the regressor and \( \varepsilon_t \) is a regression error, and we assume that \( \gamma \in \mathbb{R}^{q} \), and we define \( \sigma^2 = E\varepsilon_t^2 \). One feature of the treatment of the censored regression model in this paper is that \( \varepsilon_t \) is itself allowed to be strong mixing, and therefore potentially correlated, process. While stationarity results for general nonlinear models have been derived in e.g. Meyn and Tweedie (1994), there appear to be no results for the case where innovations are not i.i.d. (i.e. weakly dependent or heterogeneously distributed). The reason for this appears to be that the derivation of results such as those of Meyn and Tweedie (1994) depends on a Markov chain argument, and this line of reasoning appears to break down when the i.i.d. assumption is dropped. This means that in the current setting, Markov chain techniques cannot be used for the derivation of stationarity properties, which complicates our analysis substantially, but also puts our analysis on a similar level of generality as can be achieved for the linear model.

A second feature is that no assumption is made on the lag polynomial other than that

\[ \rho_{\max}(z) = 1 - \sum_{i=1}^{p} \max(0, \rho_i) z^i \]  

has its roots outside the unit circle. Therefore, in terms of the conditions on \( \rho_{\max}(z) \) and the dependence allowed for \( \varepsilon_t \), the aim of this paper is to attempt to analyze the dynamic Tobit model on a level of generality that is comparable to the level of generality under which results for the linear model AR(\( p \)) model can be derived. Note that intuitively, negative values for \( \rho_j \) can never be problematic when considering the stationarity properties of \( y_t \), since they “pull \( y_t \) back to zero”. This intuition is formalized by the fact that only \( \max(0, \rho_j) \) shows up in our stationarity requirement.

The literature on the dynamic Tobit model appears to mainly consist of (i) results and applications in panel data settings, and (ii) applications of the dynamic Tobit model in a
time series setting without providing a formal asymptotic theory. For a treatment of the
dynamic Tobit model in a panel setting, the reader is referred to Arellano and Honoré (1998,
section 8.2). The asymptotic justification for panel data Tobit models is always through
a large-$N$ type argument, which distinguishes this work from the treatment of this paper.
Wei (1999) considers dynamic Tobit models in a Bayesian framework. Finally, de Jong and
Woutersen (2003) consider the dynamic time series binary choice model and derive the weak
dependence properties of this model. Both this paper and de Jong and Woutersen (2003)
alow the error distribution to be weakly dependent. The proof in de Jong and Woutersen
(2003) establishes a contraction mapping type result for the dynamic binary choice model;
however, the proof in this paper is completely different, since other analytical issues arise in
the censored regression context.

An alternative formulation for the dynamic censored regression model could be

\[ y_t = y_t^* I(y_t^* > 0) \]

where \( B \) denotes the backward operator. This model will not be considered in this paper, and
its fading memory properties appear straightforward to derive. The formulation considered
in this paper appears the appropriate one if the 0 values in the dynamic Tobit are not caused
by a measurement issue, but have a genuine interpretation. In the case of a model for the
difference between the price of an agricultural commodity and its government-instituted price
floor, we may expect economic agents to react to the actually observed price in the previous
period rather than the latent market clearing price, and the model considered in this paper
appears more appropriate. However, if our aim is to predict tomorrow’s temperature from
todays’s temperature as measured by a lemonade-filled thermometer that freezes at zero
degrees Celsius, we should expect that the alternative formulation of the dynamic censored
regression model of Equation (2) is more appropriate.

Papers that estimate censored regression models in a time series framework cover di-
verse topics. In the financial literature, prices subject to price limits imposed in stock mar-
kets, commodity future exchanges, and foreign exchange futures markets have been treated
as censored variables. Kodres (1988, 1993) uses a censored regression model to test the
unbiasedness hypothesis in the foreign exchange futures markets. Wei (2002) proposes a
censored-GARCH model to study the return process of assets with price limits, and applies
the proposed Bayesian estimation technique to Treasury bill futures.

Censored data are also common in commodity markets where the government has histor-
ically intervened to support prices or to impose quotas. An example is provided by Chavas
and Kim (2001) who use a dynamic Tobit model to analyze the determinants of U.S. butter
prices with particular attention to the effects of market liberalization via reductions in floor
prices. Zangari and Tsurumi (1996), and Wei (1999) use a Bayesian approach to analyze the
demand for Japanese exports of passenger cars to the U.S., which were subject to quotas negotiated between the U.S. and Japan after the oil crisis of the 1970’s.

Applications in time series macroeconomics comprise determinants of borrowing behavior, open market operations and foreign exchange intervention. For instance, Peristiani (1994) estimates a censored regression model to study the decision of individual banks to borrow from the Federal Reserve’s discount window. Dynamic Tobit models have been used by Demiralp and Jordà (2002) to study the determinants of the daily transactions conducted by the Open Market Desk, and Kim and Sheen (2002) and Frenkel, Pierdzioch and Stadtmann (2003) to estimate the intervention reaction function for the Reserve Bank of Australia and the Bank of Japan, respectively.

The structure of this paper is as follows. Section 2 present our weak dependence results for \((y_t, x_t)\) in the censored regression model. In Section 3, we show the asymptotic validity of the dynamic Tobit procedure. Powell’s LAD estimation procedure for the censored regression model, which does not assume normality of errors, is considered in Section 4. Section 5 studies the determinants of temporary purchases of the Open Market Desk. The Appendix contains all proofs of our results.

2 Main results

We will prove that \(y_t\) as defined by the dynamic censored regression model satisfies a weak dependence concept called \(L_r\)-near epoch dependence. Near epoch dependence of random variables \(y_t\) on a base process of random variables \(\eta_t\) is defined as follows:

**Definition 1** Random variables \(y_t\) are called \(L_r\)-near epoch dependent on \(\eta_t\) if

\[
\sup_{t \in \mathbb{Z}} E |y_t - E(y_t|\eta_{t-M}, \eta_{t-M+1}, \ldots, \eta_{t+M})|^r = \nu(M)^r \to 0 \quad \text{as} \quad M \to \infty. \tag{3}
\]

The base process \(\eta_t\) needs to satisfy a condition such as strong or uniform mixing or independence in order for the near epoch dependence concept to be useful. For the definitions of strong (\(\alpha\)-) and uniform (\(\phi\)-) mixing see e.g. Gallant and White (1988, p. 23) or Pötscher and Prucha (1997, p. 46). The near epoch dependence condition then functions as a device that allows approximation of \(y_t\) by a function of finitely many mixing or independent random variables \(\eta_t\).

For studying the weak dependence properties of the dynamic censored regression model, assume that \(y_t\) is generated as

\[
y_t = \max(0, \sum_{i=1}^{p} \rho_i y_{t-i} + \eta_t). \tag{4}
\]
Later, we will set $\eta_t = \gamma x_t + \varepsilon_t$ in order to obtain weak dependence results for the general dynamic censored regression model that contains regressors.

When postulating the above model, we need to resolve the question as to whether there exists a strictly stationary solution to it and whether that solution is unique in some sense. See for example Bougerol and Picard (1992) for such an analysis in a linear multivariate setting. In the linear model $y_t = \rho y_{t-1} + \eta_t$, these issues correspond to showing that $\sum_{j=0}^{\infty} \rho^j \eta_{t-j}$ is a strictly stationary solution to the model that is unique in the sense that no other function of $(\eta_t, \eta_{t-1}, \ldots)$ will form a strictly stationary solution to the model.

An alternative way of proceeding to justify inference could be by considering arbitrary initial values $(y_1, \ldots, y_p)$ for the process instead of starting values drawn from the stationary distribution, but such an approach will be substantially more complicated.

We will first proceed by deriving a moment bound for $y_t$. The following theorem provides such a result:

**Theorem 1** If $\eta_t$ is strictly stationary, $\rho_{\max}(B)$ has all its roots outside the unit circle, and $\| \max(0, \eta_t) \|_r < \infty$ for some $r \geq 1$, then $\sup_{t \in \mathbb{Z}} \| y_t \|_r < \infty$.

The idea of the strict stationarity proof of this paper is to show that by writing the dynamic censored regression model as a function of the lagged $y_t$ that are sufficiently remote in the past, we obtain an arbitrarily accurate approximation of $y_t$. Let $B$ denote the backward operator, and define the lag polynomial $\rho_{\max}(B) = 1 - \sum_{i=1}^{p} \max(0, \rho_i) B^i$. The central result of this paper, the formal result showing the existence of a unique backward looking strictly stationary solution that satisfies a weak dependence property for the dynamic censored regression model is now the following:

**Theorem 2** If $\eta_t$ is a sequence of strictly stationary strong mixing random variables with $\alpha$-mixing numbers $\alpha(M)$, $\rho_{\max}(B)$ has all its roots outside the unit circle, $\| \max(0, \eta_t) \|_2 < \infty$, and

$$P[\eta_{t-1} \leq y_1, \ldots, \eta_{t-p} \leq y_p | \eta_{t-p-1}, \eta_{t-p-2}, \ldots] \geq F(y_1, \ldots, y_p) > 0$$

for some function $F(\ldots, \ldots)$ and all $(y_1, \ldots, y_p) \in \mathbb{R}^p$, then (i) there exists a solution $y_t$ to the model of Equation (4) such that $(y_t, \eta_t)$ is strictly stationary; (ii) if $z_t = f(\eta_t, \eta_{t-1}, \ldots)$ is a solution to the model, then $y_t = z_t$ a.s.; and (iii) $y_t$ is $L_2$-near epoch dependent on $\eta_t$. If in addition, $\alpha(M) \leq c_1 \exp(-c_2 M)$ for positive constants $c_1$ and $c_2$ and $\| \max(0, \eta_t) \|_2 < \infty$ for some $\delta > 0$, then the near epoch dependence sequence $\nu(M)$ satisfies $\nu(M) \leq c_4 \exp(-c_2 M^{1/3})$ for positive constants $c_4$ and $c_2$.

The condition of Equation (5) is needed in the proof to ensure that the probability of reaching 0 given the last $p$ values of $\eta_t$ is always positive. By assuming that the joint conditional
distribution of \((\eta_{t-1}, \ldots, \eta_{t-p})\) is nonzero for values of the arguments that are negative and large in absolute value, it rules out a situation in which it can be predicted with certainty that in the next period, we will have a non-zero \(y_t\) value.

One interesting aspect of the condition on \(\rho_{max}(B)\) is that negative \(\rho_i\) are not affecting the strict stationarity of the model. The intuition is that because \(y_t \geq 0\) a.s., negative \(\rho_i\) can only “pull \(y_t\) back to zero” and because the model has the trivial lower bound of 0 for \(y_t\), unlike the linear model, this model does not have the potential for \(y_t\) to tend to minus infinity.

3 The dynamic Tobit model

Define \(\beta = (\rho', \gamma', \sigma')'\), where \(\rho = (\rho_1, \ldots, \rho_p)\), and define \(b = (r', c', s)'\) where \(r\) is a \((p \times 1)\) vector and \(c\) is a \((q \times 1)\) vector. The scaled tobit loglikelihood function conditional on \(y_1, \ldots, y_p\) under the assumption of normality of the errors equals

\[
L_T(b) = L_T(c, r, s) = (T - p)^{-1} \sum_{t=p+1}^T l_t(b),
\]

where

\[
l_t(b) = I(y_t > 0) \log(s^{-1} \phi((y_t - \sum_{i=1}^p r_i y_{t-i} - c' x_t)/s))
\]

\[
+ I(y_t = 0) \log(\Phi((- \sum_{i=1}^p r_i y_{t-i} - c' x_t)/s)).
\]

In order for the loglikelihood function to be maximized at the true parameter \(\beta\), it appears hard to achieve more generality than to assume that \(\varepsilon_t\) is distributed normally given \(y_{t-1}, \ldots, y_{t-p}, x_t\). This assumption is close to assuming that \(\varepsilon_t\) given \(x_t\) and all lagged \(y_t\) is normally distributed, which would then imply that \(\varepsilon_t\) is i.i.d. and normally distributed. Therefore in the analysis of the dynamic Tobit model below, we will not attempt to consider a situation that is more general than the case of i.i.d. normal errors. Alternatively to the result below, we could also find conditions under which \(\hat{\beta}_T\) converges to a pseudo-true value \(\beta^*\). Such a result can be established under general mixing assumptions on \((x_t', \varepsilon_t)\), by the use of Theorem 2.

Let \(\hat{\beta}_T\) denote a maximizer of \(L_T(b)\) over \(b \in B\). Define \(w_t = (y_{t-1}, \ldots, y_{t-p}, x_t', 1)\). The “1” at the end of the definition of \(w_t\) allows us to write “\(b' w_t\)”. For showing consistency, we need the following two assumptions.
Assumption 1 \((x'_t, \varepsilon_t)'\) is a sequence of strictly stationary strong mixing random variables with \(\alpha\)-mixing numbers \(\alpha(M)\), where \(x_t \in \mathbb{R}^q\) and
\[
y_t = \max(0, \sum_{i=1}^{p} \rho_i y_{t-i} + \gamma' x_t + \varepsilon_t).
\] (8)

Assumption 2
1. \((x'_t, \varepsilon_t)'\) is a sequence of strictly stationary strong mixing random variables with \(\alpha\)-mixing numbers \(\alpha(M)\), and \(E|x_t|^2 < \infty\).
2. Conditional on \((x_1, \ldots, x_T)\), \(\varepsilon_t\) is independently normally distributed with mean zero and variance \(\sigma^2 > 0\).
3. \(\beta \in B\), where \(B\) is a compact subset of \(\mathbb{R}^{p+q+1}\).
4. \(Ew_tw_t'(\sum_{i=1}^{p} \rho_i y_{t-i} + \gamma' x_t > \delta)\) is positive definite for some positive \(\delta\).

Theorem 3 Under Assumption 1 and 2, \(\hat{\beta}_T \xrightarrow{p} \beta\).

For asymptotic normality, we need the following additional assumption.

Assumption 3
1. \(\beta\) is in the interior of \(B\).
2. \(I = E(\partial/\partial b)l_t(\beta)(\partial/\partial b')l_t(\beta) = -E(\partial/\partial b)(\partial/\partial b')l_t(\beta)\) is invertible.

Theorem 4 Under Assumptions 1, 2, and 3, \(T^{1/2}(\hat{\beta}_T - \beta) \xrightarrow{d} N(0, I^{-1})\).

4 Powell’s LAD for dynamic censored regression

For this section, define \(\beta = (\rho', \gamma')'\), where \(\rho = (\rho_1, \ldots, \rho_p)\), and define \(b = (r', c')'\) where \(r\) is a \((p \times 1)\) vector and \(c\) is a \((q \times 1)\) vector and \(w_t = (y_{t-1}, \ldots, y_{t-p}, x_t')\). This redefines the \(b\) and \(\beta\) vectors such as to not include \(s\) and \(\sigma\) respectively; this is because Powell’s LAD
estimator does not provide a first-round estimate for $\sigma^2$. Powell’s LAD estimator $\tilde{\beta}_T$ of the dynamic censored regression model is defined as a minimizer of

$$S_T(b) = S_T(c, r, s) = (T - p)^{-1} \sum_{t=p+1}^{T} s(y_{t-1}, \ldots, y_{t-p}, x_t, \varepsilon_t, b)$$

$$= (T - p)^{-1} \sum_{t=p+1}^{T} \left| y_t - \max(0, \sum_{i=1}^{p} r_i y_{t-i} + c' x_t) \right|$$

(9)

over a compact set subset $B$ of $\mathbb{R}^{p+q}$. We can prove consistency of Powell’s LAD estimator of the dynamic time series censored regression model under the following assumption.

**Assumption 4**

1. $\beta \in B$, where $B$ is a compact subset of $\mathbb{R}^{p+q}$.

2. The conditional distribution $F(\varepsilon_t|w_t)$ satisfies $F(0|w_t) = 1/2$, and $f(\varepsilon|w_t) = \partial F(\varepsilon|w_t) / \partial \varepsilon$ is continuous in $\varepsilon$ on a neighborhood of 0 and satisfies $c_2 \geq f(0|w_t) \geq c_1 > 0$ for constants $c_1, c_2 > 0$.

3. $E|x_t|^3 < \infty$, and $Ew_t w_t' I(\sum_{i=1}^{p} \rho_i y_{t-i} + \gamma' x_t > \delta)$ is nonsingular for some positive $\delta$.

**Theorem 5** Under Assumptions 1 and 4, $\tilde{\beta}_T \xrightarrow{p} \beta$.

For asymptotic normality, we need the following additional assumption. Below, let

$$\psi(w_t, \varepsilon_t, b) = I(b' w_t > 0)(1/2 - I(\varepsilon_t + (\beta - b)' w_t > 0)) w_t.$$  

(10)

$\psi(\ldots, \ldots)$ can be viewed as a “heuristic derivative” of $s(\ldots)$ with respect to $b$.

**Assumption 5**

1. $\beta$ is in the interior of $B$.

2. Defining $G(z, b, r) = E I(|w_t| z | w_t |^r$, we have for $z$ near 0, for $r = 0, 1, 2$,

$$\sup_{|b-\beta| < \zeta_0} |G(z, b, r)| \leq K_1 z.$$  

(11)
3. The matrix
\[
\Omega = \lim_{T \to \infty} E\left( T^{-1/2} \sum_{t=1}^{T} \psi(w_t, \varepsilon_t, \beta) \right) \left( T^{-1/2} \sum_{t=1}^{T} \psi(w_t, \varepsilon_t, \beta) \right)'
\] (12)
is well-defined, and \( N = Ef(0|w_t)I(w'_t\beta > 0)w_tw'_t \) is invertible.

4. For some \( r \geq 2, E|x_t|^{2r} < \infty, E|\varepsilon_t|^{2r} < \infty, \) and the strong mixing numbers \( \alpha(\cdot) \) for \( (x'_t, \varepsilon'_t) \) satisfy \( \alpha(M) \leq c_1 \exp(-c_2m) \) for constants \( c_1, c_2 > 0, \) and \( r > 2(p+q). \)

5. The conditional density \( f(\varepsilon|w_t) \) satisfies, for a nonrandom Lipschitz constant \( L_0, \)
\[
|f(\varepsilon|w_t) - f(\tilde{\varepsilon}|w_t)| \leq L_0 |\varepsilon - \tilde{\varepsilon}|.
\] (13)

**Theorem 6** Under Assumptions 1, 4 and 5, \( T^{1/2}(\tilde{\beta}_T - \beta) \xrightarrow{d} N(0, N^{-1}\Omega N^{-1}). \)

Assumption 5.1 is identical to Powell’s Assumption P.2, and Assumption 5.2 is the same as Powell’s Assumption R.2. Theorem 6 imposes moment conditions of order 4 or higher. The conditions imposed by Theorem 6 are moment restrictions that involve the dimensionality \( p + q \) of the parameter space. These conditions originate from the stochastic equicontinuity proof of Hansen (1996), which is used in the proof. One would expect that some progress in establishing stochastic equicontinuity results for dependent variables could aid in relaxing condition 4 imposed in Theorem 6.

5 Empirical Application

In this section we discuss an application of the dynamic censored regression model. Although there is a significant number of papers that model and estimate the Federal Open Market Committee’s (FOMC’s) reaction function, we are only aware of a recent study where lags of the dependent variable (i.e. open market operations) are included among the regressors. Even though the asymptotic properties of the dynamic Tobit model where not known, Demiralp and Jordà (2002) use a dynamic Tobit model to analyze whether the February 4, 1994, Fed decision to publicly announce changes in the federal funds rate target affected the manner in which the Open Market Desk conducts operations. In this section we re-evaluate their findings.
5.1 Data

The data used in the analysis can be found in Demiralp and Jordà (2002). The data are daily and span the period between April 25, 1984 and August 14, 2000. We classify open market operations in four groups: (a) temporary purchases comprise overnight reversible repurchase agreements (RP) and term RP; (b) permanent purchases include T-bill purchases and coupon purchases; (c) temporary sales are overnight and term matched sale-purchases; and (d) permanent sales comprise T-bill sales and coupon sales. We restrict our analysis to the change in the maintenance-period-average level of reserves brought about by temporary purchases of the Open Market Desk. Because the computation of reserves is based on a 14-day maintenance period that starts on Thursday and finishes on the "Settlement Wednesday" two weeks later, the maintenance-period average is the object of attention of the Open Desk. Thus, all operations are adjusted according to the number of days spanned by the transaction, and standardized by the aggregate level of reserves held by depository institutions in the maintenance period previous to the execution of the transaction. Daily values for temporary purchases are plotted in Figure 1. Note that, although not clearly apparent in the figure, the Open Market Desk engaged in temporary purchases only 37% of the time.

Demiralp and Jordà (2002) separate deviations of the federal funds rate from the target into three components:

\[
NEED_t = f_t - \left[ f_{m(t)-1}^* + w_t E_{m(t)-1} (\Delta f_{m(t)}) \right]
\]

\[
EXPECT_t = E_{m(t)-1} (\Delta f_{m(t)})
\]

\[
SURPRISE_t = \Delta f_t^* - E_{m(t)-1} (\Delta f_{m(t)})
\]

where the maintenance period to which observation in day \( t \) belongs is denoted by \( m(t) \), \( f_t \) denotes the federal funds rate in day \( t \); \( f_{m(t)-1}^* \) denotes the value of the target in the maintenance period previous the one to which observation \( t \) belongs; \( E_{m(t)-1}(\Delta f_{m(t)}) \) denotes the expectation of a target change in day \( t \), conditional on the information available at the beginning of the 14-day maintenance period; and \( w_t \) denotes the probability of a target change on date \( t \). Both the expected change in the target, \( E_{m(t)-1}(\Delta f_{m(t)}) \), and the weights \( w_t \) are calculated by Demiralp and Jordà (2002) using the ACH model of Hamilton and Jordà (2002). This decomposition is intended to reflect three different motives for open market purchases: (1) to add liquidity in order to accommodate shocks to reserve's demand; (2) to accommodate expectations of future changes in the target; and (3) to adjust to a new target level. Thus, \( NEED_t \) represents a proxy for the projected reserve need, and changes in the federal funds rate are separated into an expected component, \( EXPECT_t \), and a surprise component, \( SURPRISE_t \). The latter takes a non-zero value for the 115 days in the sample when there was a change in the target, and zero otherwise.
5.2 Model and estimation procedure

The following dynamic censored regression model is used to describe temporary purchases by the Open Market Desk:

\[
TB_t = \max(0, \gamma + \sum_{m=1}^{4} \gamma_m D_{tm} + \sum_{j=1}^{3} \rho_j TB_{t-j} + \sum_{j=1}^{3} \gamma_j^{TS} TS_{t-j} + \sum_{j=1}^{3} \gamma_j^{PB} PB_{t-j} \\
+ \sum_{j=1}^{3} \gamma_j^{PS} PS_{t-j} + \sum_{m=1}^{10} \gamma_m^{N} NEED_{t-m} \times DAY_{tm} + \sum_{m=1}^{10} \gamma_m^{E} EXPECT_{t-m} \times DAY_{tm} \\
+ \sum_{j=0}^{3} \gamma_j^{S} SURPRISE_{t-j} + \epsilon_t) \tag{17}
\]

where \(TB_t\) denotes temporary purchases, \(TS_t\) denotes temporary sales, \(PB_t\) denotes permanent purchases, \(PS_t\) denotes permanent sales, \(DAY_{tm}\) denotes a vector of maintenance-day dummies, \(D_{tm}\) is such that \(D_{t1} = DAY_{t1}, D_{t2} = DAY_{t2}, D_{t3} = DAY_{t7}\) and \(D_{t4} = DAY_{t10}\), and \(\epsilon_t\) is a stochastic disturbance.

This model is a simplified version of the one estimated by Demiralp and Jordà (2002) in that it does not include dummies for all days in the maintenance period. Instead, to control for differences in the reserve levels that the Federal Reserve might want to leave in the system at the end of the day, we include only dummies for the first Thursday, the first Friday, the second Friday, and Settlement Wednesday. However, we do allow the response of temporary purchases to reserve needs and expected changes in the fed funds rate to vary across all days of the maintenance period. Furthermore, we estimate the dynamic censored model using a method known to be robust to both non-normality and heteroskedasticity: Powell’s CLAD estimator.

The CLAD estimates, \(\hat{b}\), are obtained by using the iterative linear programming algorithm proposed by Buchinsky (1994). This procedure amounts to first solving the linear programming (LP) representation of the optimization problem

\[
\min_b \left\{ \frac{1}{T-p} \sum_{t=p+1}^{T} \frac{1}{2} \text{sgn} (y_t - b'w_t)(y_t - b'w_t) \right\} \tag{18}
\]

to obtain the estimates \(\hat{b}^{(1)}\). Then, solve the LP problem for \(\hat{b}^{(2)}\) using the observations for which \(\hat{b}^{(1)} w_t > 0\). This procedure is repeated until the set of observation used in two consecutive iterations is the same. Standard errors for \(\hat{b}\) are obtained according to Equation
We compute \( \tilde{\Omega} \) as the long-run variance of \( \tilde{\psi}(w_t, \tilde{b}) = I(\tilde{b}'w_t > 0)[\frac{1}{2} - I(y_t < \tilde{b}'w_t)]w_t \), following the suggestions of Andrews (1991) to select the bandwidth for the Bartlett kernel. To compute \( \tilde{N} \), we estimate \( f(0|w_t) \) using a higher-order Gaussian kernel with the order and bandwidth selected according to Hansen (2003, 2004). The reported standard errors for the Tobit estimates are the quasi-maximum likelihood standard errors.

5.3 Estimation Results

Maximum likelihood estimates of the Tobit model and corresponding standard errors are presented in the first two columns of Table 1. Of interest is the presence of statistically significant coefficients on the lags of the dependent variable, \( TB_{t-j} \). This persistence suggests that in order to attain the desired target, the Open Market Desk had to exercise pressure on the fed funds market in a gradual manner, on consecutive days. The negative and statistically significant coefficients on lagged temporary sales, \( TS_{t-j} \), confirms Demiralp and Jordà’s (2002) finding that temporary sales have constituted substitutes for temporary purchases. In other words, in the face of a reserve shortage the Open Market Desk could react by conducting temporary purchases and/or delaying temporary sales. The positive and statistically significant coefficients on the \( NEED_{t-1} \times DAY_{tm} \) variables is consistent with an accommodating behavior of the Fed to deviations of the federal funds rate from its target. The Tobit estimates suggest that expectations of target changes were accommodated in the first day of the maintenance period, and did not significantly affect temporary purchases on the remaining days. As for the effect of surprise changes in the target, the estimated coefficients are statistically insignificant. According to Demiralp and Jordà (2002), statistically insignificant coefficients on \( SURPRISE_{t-j} \) can be interpreted as evidence of the announcement effect.\(^1\) This suggest that the Fed did not require temporary purchases to signal the change in the target, once it had been announced (or inferred by the markets).

However, the Tobit estimates will be inconsistent if the error terms are heteroskedastic or non-normal. In fact, this clearly appears to be the case here. A Lagrange multiplier test of heteroskedasticity obtained by assuming \( Var(\varepsilon_t|w_t) = \sigma^2 \exp(\delta'z_t) \), where \( z_t \) is a vector that contains all elements in \( w_t \) but the constant, rejects the null \( H_0 : \delta = 0 \) at the 1% level. In addition, the Jarque-Bera statistics leads us to reject the null that the residuals are normally distributed at a 1% level. This is clearly illustrated in Figure 2, which plots the histogram for the Tobit residuals. Therefore, we consider the Powell’s (1994) CLAD estimator, which

\(^1\)Even though the federal funds target has only been announced since the February 3-4 FOMC meeting, Demiralp and Jordà (forthcoming) provide evidence that, since late 1989, financial markets were able to decode changes in the target from the pattern of open market operations. Furthermore, research by Cook and Hahn (1988) suggest that even in earlier periods, market participants were able to read signals of a target change in the Fed’s behavior.
allows for both heteroskedasticity and nonnormality.

CLAD estimates and corresponding standard errors are reported in the third and fourth column of Table 1, respectively. These estimates imply a reaction of the Fed to reserve needs similar to that obtained from the Tobit model. Nevertheless, the CLAD estimates imply a smaller degree of persistence in temporary purchases, as well as different conclusions regarding the announcement effect, the liquidity effect, and the impact of expected changes in the federal funds target.

First, the CLAD estimates indicate that expected changes in the target where generally accommodated on days other than the first Monday of the maintenance period, not only on the first Thursday as the Tobit estimates suggest. But, more importantly, statistically significant coefficients on the second and third lag of the surprise variable cast some doubts regarding the evidence on the announcement effect. In fact, while the positive coefficient on the second lag is counterintuitive, a negative and statistically significant coefficient on the third lag is consistent with the conventional liquidity effect. Furthermore, clear evidence of the liquidity effect is found if we restrict the sample to the post August 18, 1998 period when lagged reserve accounting was operational, and the federal funds target was being announced. In that subsample, evidence of the announcement effect seems to be only present in the decreased persistence of temporary purchases.

5.4 Conclusions

The estimated Tobit model is broadly consistent with Demiralp and Jordà’s (2002) findings regarding the determinants of temporary open market purchases. However, with the caveat that the variable SURPRISE$t$ takes on nonzero values only for the 115 target changes in the sample, we interpret our CLAD estimates as evidence that the Open Market Desk still reacts to unanticipated changes in the target in a manner broadly consistent with the liquidity effect. All in all, evidence of the announcement effect seems to be only present in the decreased persistence of temporary sales after August 17, 1998.\footnote{Similar conclusions are drawn if we use the futures federal funds rate (Kuttner, 2001) to decompose anticipated and unanticipated changes in policy. The only difference is that autoregressive terms are statistically significant in the 1998-2000 period.}
Table 1: Estimation results for Temporary Open Market Purchases

<table>
<thead>
<tr>
<th>Variable</th>
<th>Tobit estimate</th>
<th>Tobit std.err.</th>
<th>CLAD estimate</th>
<th>CLAD std.err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-18.742**</td>
<td>1.753</td>
<td>-0.504</td>
<td>0.269</td>
</tr>
<tr>
<td>First Thursday</td>
<td>16.962***</td>
<td>2.833</td>
<td>1.896***</td>
<td>0.468</td>
</tr>
<tr>
<td>First Friday</td>
<td>-17.046***</td>
<td>2.917</td>
<td>-1.229*</td>
<td>0.709</td>
</tr>
<tr>
<td>Second Friday</td>
<td>-12.051***</td>
<td>2.683</td>
<td>-3.167***</td>
<td>0.996</td>
</tr>
<tr>
<td>Settlement Wednesday</td>
<td>4.456***</td>
<td>1.553</td>
<td>2.828***</td>
<td>0.431</td>
</tr>
<tr>
<td>TB(-1)</td>
<td>0.264***</td>
<td>0.035</td>
<td>0.047***</td>
<td>0.007</td>
</tr>
<tr>
<td>TB(-2)</td>
<td>0.292***</td>
<td>0.042</td>
<td>0.079***</td>
<td>0.007</td>
</tr>
<tr>
<td>TB(-3)</td>
<td>0.305***</td>
<td>0.049</td>
<td>0.069***</td>
<td>0.007</td>
</tr>
<tr>
<td>TS(-1)</td>
<td>-1.726***</td>
<td>0.611</td>
<td>-6.934***</td>
<td>1.381</td>
</tr>
<tr>
<td>TS(-2)</td>
<td>-0.865*</td>
<td>0.447</td>
<td>-1.058*</td>
<td>0.568</td>
</tr>
<tr>
<td>TS(-3)</td>
<td>-1.895***</td>
<td>0.408</td>
<td>-0.570*</td>
<td>0.273</td>
</tr>
<tr>
<td>PB(-1)</td>
<td>-0.018</td>
<td>0.085</td>
<td>-0.031</td>
<td>0.035</td>
</tr>
<tr>
<td>PB(-2)</td>
<td>-0.065</td>
<td>0.074</td>
<td>0.079***</td>
<td>0.013</td>
</tr>
<tr>
<td>PB(-3)</td>
<td>-0.073</td>
<td>0.072</td>
<td>-0.167***</td>
<td>0.049</td>
</tr>
<tr>
<td>PS(-1)</td>
<td>0.146</td>
<td>0.253</td>
<td>0.178***</td>
<td>0.045</td>
</tr>
<tr>
<td>PS(-2)</td>
<td>-0.151</td>
<td>0.245</td>
<td>0.032</td>
<td>0.045</td>
</tr>
<tr>
<td>PS(-3)</td>
<td>-0.198</td>
<td>0.223</td>
<td>0.039</td>
<td>0.043</td>
</tr>
<tr>
<td>NEED(-1)×Day1</td>
<td>-0.501</td>
<td>2.974</td>
<td>1.582***</td>
<td>0.503</td>
</tr>
<tr>
<td>NEED(-1)×Day2</td>
<td>11.645**</td>
<td>4.691</td>
<td>3.112***</td>
<td>0.727</td>
</tr>
<tr>
<td>NEED(-1)×Day3</td>
<td>24.030***</td>
<td>8.595</td>
<td>3.612*</td>
<td>1.976</td>
</tr>
<tr>
<td>NEED(-1)×Day4</td>
<td>-7.489</td>
<td>7.775</td>
<td>-1.509</td>
<td>1.745</td>
</tr>
<tr>
<td>NEED(-1)×Day5</td>
<td>21.671***</td>
<td>7.612</td>
<td>5.751***</td>
<td>1.848</td>
</tr>
<tr>
<td>NEED(-1)×Day6</td>
<td>5.312</td>
<td>10.941</td>
<td>-8.851***</td>
<td>3.165</td>
</tr>
<tr>
<td>NEED(-1)×Day7</td>
<td>33.429***</td>
<td>10.471</td>
<td>11.853***</td>
<td>2.347</td>
</tr>
<tr>
<td>NEED(-1)×Day8</td>
<td>6.842</td>
<td>7.789</td>
<td>8.894***</td>
<td>2.145</td>
</tr>
<tr>
<td>NEED(-1)×Day9</td>
<td>13.402***</td>
<td>4.953</td>
<td>4.557***</td>
<td>1.171</td>
</tr>
<tr>
<td>NEED(-1)×Day10</td>
<td>3.972*</td>
<td>2.083</td>
<td>1.132**</td>
<td>0.508</td>
</tr>
<tr>
<td>EXPECT(-1)×Day1</td>
<td>121.451**</td>
<td>53.231</td>
<td>38.392***</td>
<td>5.334</td>
</tr>
<tr>
<td>EXPECT(-1)×Day2</td>
<td>48.599</td>
<td>34.018</td>
<td>11.161*</td>
<td>6.105</td>
</tr>
<tr>
<td>EXPECT(-1)×Day3</td>
<td>-14.656</td>
<td>38.828</td>
<td>-27.103***</td>
<td>9.720</td>
</tr>
<tr>
<td>EXPECT(-1)×Day4</td>
<td>-25.528</td>
<td>29.346</td>
<td>4.957</td>
<td>5.303</td>
</tr>
<tr>
<td>EXPECT(-1)×Day5</td>
<td>-54.997*</td>
<td>29.859</td>
<td>4.929</td>
<td>5.134</td>
</tr>
<tr>
<td>EXPECT(-1)×Day6</td>
<td>51.573*</td>
<td>30.636</td>
<td>48.795***</td>
<td>5.201</td>
</tr>
<tr>
<td>EXPECT(-1)×Day7</td>
<td>-37.603</td>
<td>36.721</td>
<td>-16.383</td>
<td>16.266</td>
</tr>
<tr>
<td>EXPECT(-1)×Day8</td>
<td>39.313*</td>
<td>20.515</td>
<td>10.372*</td>
<td>5.052</td>
</tr>
<tr>
<td>EXPECT(-1)×Day10</td>
<td>16.269</td>
<td>16.363</td>
<td>8.230</td>
<td>5.019</td>
</tr>
<tr>
<td>SURPRISE(-1)</td>
<td>-14.046</td>
<td>15.387</td>
<td>-0.647</td>
<td>3.815</td>
</tr>
<tr>
<td>SURPRISE(-2)</td>
<td>11.516</td>
<td>13.044</td>
<td>1.255</td>
<td>3.594</td>
</tr>
<tr>
<td>SURPRISE(-3)</td>
<td>14.619</td>
<td>17.648</td>
<td>19.541***</td>
<td>5.149</td>
</tr>
<tr>
<td>SCALE</td>
<td>1094.099</td>
<td>104.478</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*, ** and *** denote significance at the 10%, 5% and 1% levels respectively.
Figure 1: Temporary Open Market Purchases

Figure 2: Standardized residuals
References


Chavas and Kim (2001), An econometric analysis of the effects of market liberalization on price dynamics and price volatility, mimeo, Department of Agricultural and Applied Economics, University of Wisconsin.

Cook, T. and T. Hahn (1988), The information content of discount rate announcements and
their effect on market interest rates, *Journal of Money, Credit, and Banking* 20(2), 167-180.


Kuttner, Kenneth (2001), Monetary policy surprises and interest rates: evidence from the


**Appendix**

Define $\hat{y}_t^m = 0$ for $m \leq 0$ and $\hat{y}_t^m = \max(0, \eta_t + \sum_{i=1}^p \rho_i \hat{y}_{t-i}^m)$. Therefore, $\hat{y}_t^m$ is the approximation for $y_t$ that presumes $y_{t-m}, \ldots, y_{t-m-p} = 0$. We can obtain a well-defined upper
bound for \( y_t \) and \( \hat{y}_t^{\text{m}} \):

**Lemma 1** If the lag polynomial \((1 - \max(0, \rho_1)B - \ldots - \max(0, \rho_p)B^p)\) has all its roots outside the unit circle and \( \sup_{t \in \mathbb{Z}} E \max(0, \eta_t) < \infty \), then for a well-defined random variable \( f_t = f(\eta_t, \eta_{t-1}, \ldots) = \sum_{j=0}^{\infty} L_j^1 \max(0, \eta_{t-j}) \), and \( L_j^1 \) that are such that \( L_j^1 \leq c_1 \exp(-c_2 j) \) for positive constants \( c_1 \) and \( c_2 \),

\[
\hat{y}_t^{\text{m}} \leq f_t \quad \text{and} \quad y_t \leq f_t.
\]

**Proof of Lemma 1:**

Note that, by successive substitution of the definition of \( y_t \) for the \( y_t \) that has the largest value for \( t \),

\[
\hat{y}_t^{\text{m}} \leq \max(0, \eta_t) + \sum_{i=1}^{p} \max(0, \rho_i) \hat{y}_{t-i}^{\text{m}}
\]

\[
= \max(0, \eta_t) + \sum_{i=1}^{p} L_i^1 \hat{y}_{t-i}^{\text{m}}
\]

\[
\leq \max(0, \eta_t) + \sum_{i=2}^{p} \max(0, \rho_i) \hat{y}_{t-i}^{\text{m}} + \max(0, \rho_1) \max(0, \eta_{t-1}) + \sum_{i=1}^{p} \max(0, \rho_i) \hat{y}_{t-i-1}^{\text{m}}
\]

\[
= \max(0, \eta_t) + L_1^1 \max(0, \eta_{t-1}) + \sum_{i=1}^{p} L_i^2 \hat{y}_{t-i-1}^{\text{m}}
\]

\[
\leq \max(0, \eta_t) + L_1^1 \max(0, \eta_{t-1}) + L_2^1 \max(0, \eta_{t-2}) + \sum_{i=1}^{p} L_i^3 \hat{y}_{t-i-2}^{\text{m}}
\]

\[
\leq \sum_{j=0}^{\infty} L_j^1 \max(0, \eta_{t-j}).
\]

The \( L_j^1 \) satisfy, for \( j \geq 2 \),

\[
L_j^1 = L_{j-1}^1 + \max(0, \rho_1) L_{j-1}^1,
\]

\[
L_1^1 = L_{2}^{-1} + \max(0, \rho_1) L_{1}^{-1},
\]

\[
L_2^1 = L_{3}^{-1} + \max(0, \rho_2) L_{1}^{-1},
\]

19
\[ L_{i-1}^{j} = L_{i}^{j-1} + \max(0, \rho_{i-1}) L_{i}^{j-1}, \]
\[ L_{i}^{j} = \max(0, \rho_{i}) L_{i}^{j-1}. \]

From these equations it follows that we can write, for the backward operator \( B \) that is such that \( B(L_{i}^{j}) = L_{i}^{j-1} \),
\[ (1 - \sum_{j=1}^{p} \max(0, \rho_{j}) B^{j}) L_{i}^{j} = 0. \]

From the fact that the above lag polynomial has all its roots outside the unit circle by assumption, it follows that \( L_{i}^{j} \leq c_{1} \exp(-c_{2} j) \) for positive constants \( c_{1} \) and \( c_{2} \). Also, if \( \sup_{t \in \mathbb{Z}} E \max(0, \eta_{t}) < \infty \), then \( \sum_{j=0}^{\infty} L_{i}^{j} \max(0, \eta_{t-j}) \) is a well-defined random variable. The above reasoning also applies to \( y_{t} \) (instead of \( \hat{y}_{t}^{m} \)). \( \square \)

We are now in a position to prove Theorem 1.

**Proof of Theorem 1:**

The result of Theorem 1 follows by noting that, by Lemma 1,
\[ \| y_{t} \|_{r} \leq \sum_{j=0}^{\infty} L_{i}^{j} \max(0, \eta_{t-j}) \|_{r} < \infty. \]
\( \square \)

We will also need an exponential inequality:

**Lemma 2** If \( x_{t} \) is \( L_{1} \)-near epoch dependent on \( v_{t} \), where \( v_{t} \) is \( \alpha \)-mixing and \( \alpha(m) + \nu(m) \leq C_{1} \exp(-C_{2} m) \) for positive constants \( C_{1} \) and \( C_{2} \), and \( |x_{t}| \leq 1 \), then for all \( \delta > 0 \),
\[ P(\sum_{t=1}^{m} x_{t} > \delta) \leq c_{1} \exp(-c_{2} \delta^{2} m^{1/3}) \]
for positive constants \( c_{1} \) and \( c_{2} \) possibly depending on \( \delta \).
Proof of Lemma 2:

Observe that for all \( k > 0 \),

\[
x_t = \sum_{j=-k}^{k-1} (E(x_t|v_{t-j}, \ldots) - E(x_t|v_{t-j-1}, \ldots)) + E(x_t|v_{t-k}, \ldots) + (x_t - E(x_t|v_{t+k}, \ldots)),
\]

and therefore for all \( k > 0 \),

\[
P(m^{-1} \sum_{t=1}^{m} x_t > \delta) \
\leq \delta^{-1} || E(x_t|v_{t-k-1}, \ldots) ||_1 + \delta^{-1} || x_t - E(x_t|v_{t+k}, \ldots) ||_1 \\
+ \sum_{j=-k}^{k} P(m^{-1} \sum_{t=1}^{m} (E(x_t|v_{t-j}, \ldots) - E(x_t|v_{t-j-1}, \ldots)) > \delta/(2k + 1)).
\]

By the \( L_1 \)-near epoch dependence condition, boundedness and the \( L_1 \)-mixingale property of \( x_t \) (see Andrews (1988)),

\[
\delta^{-1} || E(x_t|v_{t-k-1}, \ldots) ||_1 + \delta^{-1} || x_t - E(x_t|v_{t+k}, \ldots) ||_1 \leq \delta^{-1} C_1 \exp(-C_2 k),
\]

and by Azuma’s inequality (see Azuma (1967)),

\[
P(m^{-1} \sum_{t=1}^{m} (E(x_t|v_{t-j}, \ldots) - E(x_t|v_{t-j-1}, \ldots)) > \delta/(2k))
\leq 2 \exp(-\delta^2 m/8k^2).
\]

By choosing \( k = \lfloor m^{1/3} \rfloor \) and collecting terms, the result now follows. \( \square \)

The following lemma is needed for the stationarity proof of Theorem 2. For \( \zeta > 0 \), let

\[
H^\zeta(x) = -\zeta^{-1} x I(-\zeta \leq x \leq 0) + I(x \leq -\zeta).
\]

\[
I_{tl} = \prod_{j=0}^{p-1} I(\eta_{t-l-j} \leq -\sum_{i=1}^{p} \rho_{j} f_{t-l-j-i})
\]

and

\[
I_{tl}^\zeta = \prod_{j=0}^{p-1} H^\zeta(\eta_{t-l-j} + \sum_{i=1}^{p} \rho_{j} f_{t-l-j-i}).
\]
Lemma 3 Assume that $\eta_t$ is strictly stationary and strong mixing and satisfies $\| \max(0, \eta_t) \|_2 < \infty$. Then for all $t \in \mathbb{Z}$ and $\delta > 0$, as $m \to \infty$,

$$(m - p)^{-1} \sum_{l=1}^{m-p} (I_u^\xi \log(\delta) + \log(1 + \delta)(1 - I_u^\xi)) \xrightarrow{p} E(I_u^\xi \log(\delta) + \log(1 + \delta)(1 - I_u^\xi)).$$

Proof of Lemma 3:

Note that we can write

$$(m - p)^{-1} \sum_{l=1}^{m-p} (I_u^\xi \log(\delta) + \log(1 + \delta)(1 - I_u^\xi))$$

$$= (m - p)^{-1} \sum_{l=1}^{m-p} (I_{t,m-p+1-l}^\xi \log(\delta) + \log(1 + \delta)(1 - I_{t,m-p+1-l}^\xi)).$$

Note that

$$I_{t,m-p+1-l}^\xi = \prod_{j=0}^{p-1} H^\xi(\eta_{t-(m-p+1-l)-j} + \sum_{i=1}^{p} \rho_i f_{t-(m-p+1-l)-j-i}),$$

and for all $t$ and $j$,

$$\eta_{t-(m-p+1-l)-j} + \sum_{i=1}^{p} \rho_i f_{t-(m-p+1-l)-j-i}$$

$$= \eta_{t-(m-p+1-l)-j} + \sum_{i=1}^{p} \sum_{k=0}^{\infty} \rho_i L_i^k \max(0, \eta_{t-(m-p+1-l)-j-i-k})$$

$$= \eta_{t-(m-p+1-l)-j} + \sum_{k=0}^{\infty} \max(0, \eta_{t-(m-p+1-l)-j-i-k}) \sum_{i=1}^{p} \rho_i L_i^{k-1} I(i \leq k) = w_{t-(m-p+1-l)-j}$$

is strictly stationary (as a function of $l$) and $L_2$-near epoch dependent on $\eta_{t-(m-p+1-l)-j}$, and that $\nu(M)$ decays exponentially. This is because for $M \geq 1$,

$$\| w_{t-(m-p+1-l)-j} - E(w_{t-(m-p+1-l)-j} | \eta_{t-(m-p+1-l)-j-M}, \cdots, \eta_{t-(m-p+1-l)-j}) \|_2$$

22
\[
\left\| \max(0, \eta_t) \right\|_2 \sum_{k=M+1}^{\infty} \sum_{i=1}^{p} \rho_i L^i_{\eta_t - l - i} I(i \leq k),
\]
and the last expression converges to 0 as \( M \to \infty \) at exponential rate because \( L^k_1 \) converges to zero at an exponential rate. Therefore, because \( H^\zeta(\cdot) \) is Lipschitz-continuous,

\[
H^\zeta(\eta_t - (m-p+1-l)-j + \sum_{i=1}^{p} \rho_i f_{t-(m-p+1-l)-i})
\]
is also \( L_2 \)-near epoch dependent on \( \eta_t \) with an exponentially decreasing \( \nu(\cdot) \) sequence, and so is

\[
\prod_{j=0}^{p-1} H^\zeta(\eta_t - (m-p+1-l)-j + \sum_{i=1}^{p} \rho_i f_{t-(m-p+1-l)-i}).
\]

See Pötscher and Prucha (1997) for more information about these manipulations with near epoch dependent processes. The result of this lemma then follows from the weak law of large numbers for \( L_2 \)-near epoch dependent processes of Andrews (1988).

\[\square\]

**Lemma 4** If

\[
P[\eta_{t-1} \leq y_1, \ldots, \eta_{t-p} \leq y_p | \eta_{t-p-1}, \eta_{t-p-2}, \ldots] \geq F(y_1, \ldots, y_p) > 0
\]
for all \( (y_1, \ldots, y_p) \in \mathbb{R}^p \), then for all \( \zeta > 0 \),

\[
E \prod_{j=1}^{p} I(\eta_{t-j} + \sum_{i=1}^{p} \rho_i f_{t-i-j} \leq -\zeta) > 0.
\]

**Proof of Lemma 4:**

Note that for all \( K \geq 0 \),

\[
\eta_t + \sum_{i=1}^{p} \rho_i f_{t-i} = \eta_t + \sum_{i=1}^{p} \rho_i \sum_{l=0}^{\infty} \max(0, \eta_{t-i-l}) L^l_1
\]

23
\[ \eta_t + \sum_{k=1}^{K} \max(0, \eta_t - k) \sum_{i=1}^{p} \rho_i L_1^{k-i} I(i \leq k) + \sum_{k=K+1}^{\infty} \max(0, \eta_t - k) \sum_{i=1}^{p} \rho_i L_1^{k-i} I(i \leq k). \]

Now defining
\[ g^K_t = \eta_t + \sum_{k=1}^{K} \max(0, \eta_t - k) \sum_{i=1}^{p} \rho_i L_1^{k-i} I(i \leq k), \]

it follows from our assumption that \((g^{p-1}_t, \ldots, g^0_{t-p})\) satisfies, for some function \(H(\ldots, \ldots)\),
\[
P[g^{p-1}_t \leq y_1, \ldots, g^0_{t-p} \leq y_p | \eta_{t-p-1}, \eta_{t-p-2}, \ldots] \geq H(y_1, \ldots, y_p) > 0.
\]

Therefore,
\[
E \prod_{j=1}^{p} I(\eta_t - j + \sum_{i=1}^{p} \rho_i f_{t-i-j} \leq -\zeta)
\]
\[
= E \prod_{j=1}^{p} I(g^{p-j}_t \leq -\sum_{k=p-j+1}^{\infty} \max(0, \eta_t - k) \sum_{i=1}^{p} \rho_i L_1^{k-i} I(i \leq k) - \zeta)
\]
\[
\geq EH(h_{t-1}, \ldots, h_{t-p})
\]

where
\[
h_{t-j} = -\sum_{k=p-j+1}^{\infty} \max(0, \eta_t - k) \sum_{i=1}^{p} \rho_i L_1^{k-i} I(i \leq k) - \zeta.
\]

Now because \(H(\ldots, \ldots) > 0\) and because the \(h_t\) are well-defined random variables, the result follows. \(\square\)

**Lemma 5** For some well-defined random variable \(y_t\) such that \((y_t, \eta_t)\) is strictly stationary,
\[ \hat{y}_t^m \xrightarrow{as} y_t \quad \text{as} \quad m \to \infty. \]
Proof of Lemma 5:

We will use the Cauchy criterion to show that $\hat{y}_t^m$ converges almost surely, and we will define $y_t$ to be this limit. By the Cauchy criterion, $\hat{y}_t^m$ converges a.s. if $\max_{k \geq m} |\hat{y}_t^k - \hat{y}_t^m| \rightarrow 0$ as $m \rightarrow \infty$. Now, note that for all $m \geq k$,

$$\hat{y}_t^k = \hat{y}_t^m = 0 \text{ if } \eta_t \leq -\sum_{i=1}^{p} \rho_i f_{t-i} \text{ and } \eta_t \leq -\sum_{i=1}^{p} \rho_i \hat{y}_{t-i}^k,$$

so certainly,

$$\hat{y}_t^k = \hat{y}_t^m = 0 \text{ if } \eta_t \leq -\sum_{i=1}^{p} \rho_i f_{t-i},$$

and therefore $\max_{k \geq m} |\hat{y}_t^k - \hat{y}_t^m| = 0$ for all $m > p$ if there can be found $p$ consecutive “small” $\eta_{t-l}$ that are negative and large in absolute value in the range $l = 1, \ldots, m-1$; i.e. if

$$\eta_{t-l} \leq -\sum_{i=1}^{p} \rho_i f_{t-l-i}$$

for all $l \in \{a, a+1, \ldots, a+p-1\}$ for some $a \in \{1, \ldots, m-p\}$. Therefore, for all $1/2 > \delta > 0$, $\zeta > 0$, and $c > 0$,

$$P[\max_{k \geq m} |\hat{y}_t^k - \hat{y}_t^m| > 0] \leq P[\text{there are no } p \text{ consecutive “small” } \eta_t] \leq P[\text{there are no } p \text{ consecutive “small” } \eta_t \text{ starting at } t-l]$$

$$\leq E \prod_{l=1}^{m-p} (1 - I(\text{there are } p \text{ consecutive “small” } \eta_t \text{ starting at } t-l))$$

$$\leq E \prod_{l=1}^{m-p} (1 - \prod_{j=0}^{p-1} I(\eta_{t-l-j} \leq -\sum_{i=1}^{p} \rho_i f_{t-l-j-i}))$$

$$= E \exp((m-p)(m-p)^{-1} \sum_{l=1}^{m-p} \log(1 - \prod_{j=0}^{p-1} I(\eta_{t-l-j} \leq -\sum_{i=1}^{p} \rho_i f_{t-l-j-i})))$$

$$\leq \exp(-(m-p)c) + P[(m-p)^{-1} \sum_{l=1}^{m-p} \log(1 - \prod_{j=0}^{p-1} I(\eta_{t-l-j} \leq -\sum_{i=1}^{p} \rho_i f_{t-l-j-i}) > -c]$$

25
\[
\leq \exp(-(m - p)c)) + P[(m - p)^{-1} \sum_{l=1}^{m-p} (I_{ul} \log(\delta) + \log(1 + \delta)(1 - I_{ul})) > -c]
\]
\[
\leq \exp(-(m - p)c)) + P[(m - p)^{-1} \sum_{l=1}^{m-p} (I_{ul}^{c} \log(\delta) + \log(1 + \delta)(1 - I_{ul}^{c})) > -c],
\]
(19)

where
\[
I_{ul} = \prod_{j=0}^{p-1} I(\eta_{l-j} \leq -\sum_{i=1}^{p} \rho_{j} f_{l-j-i})
\]
and
\[
I_{ul}^{c} = \prod_{j=0}^{p-1} H^{c}(\eta_{l-j} + \sum_{i=1}^{p} \rho_{j} f_{l-j-i})
\]
for
\[
H^{c}(x) = -\zeta^{-1} xI(-\zeta \leq x \leq 0) + I(x \leq -\zeta).
\]
Note that \(I_{ul} \geq I_{ul}^{c}\) because \(I(x \leq 0) \geq H^{c}(x)\). Both terms in Equation (19) now converge to zero as \(m \to \infty\) for a suitable choice of \(\zeta, c\) and \(\delta\) if
\[
E(m - p)^{-1} \sum_{l=1}^{m-p} (I_{ul}^{c} \log(\delta) + \log(1 + \delta)(1 - I_{ul}^{c}))
\]
\[
= E(I_{ul}^{c} \log(\delta) + \log(1 + \delta)(1 - I_{ul}^{c})) < 0
\]
(20)
and
\[
(m - p)^{-1} \sum_{l=1}^{m-p} (I_{ul}^{c} \log(\delta) + \log(1 + \delta)(1 - I_{ul}^{c}))
\]
satisfies a weak law of large numbers as \(m \to \infty\). This weak law of large numbers is proven in Lemma 3. Now if \(EI_{ul}^{c} > 0\), we can pick \(\delta > 0\) small enough to satisfy the requirement of Equation (20). Now,
\[
EI_{ul}^{c} = E \prod_{j=1}^{p} H^{c}(\eta_{l-j} - \sum_{i=1}^{p} \rho_{j} f_{l-j-i})
\]
\[ \geq E \prod_{j=1}^{p} I(\eta_{t-1-j} + \sum_{i=1}^{p} \rho_j f_{t-1-j-i} - \zeta), \]

and the last term is positive by Lemma 4.

Since \( \hat{y}_m^m = f_m(\eta_t, \ldots, \eta_{t-m}) \) is strictly stationary because it depends on a finite numbers of \( \eta_t \), \( \lim_{m \to \infty} (\hat{y}_m^m, \eta_t) = (y_t, \eta_t) \) is also strictly stationary.

\[ \Box \]

**Proof of Theorem 2:**

Noting that \( y_t \) as constructed in Lemma 5 is a solution to the dynamic censored regression model, part (i) of Theorem 2 follows. Also, by the reasoning of Lemma 5 it follows that any \( z_t = f(\eta_t, \eta_{t-1}, \ldots) \) that is a solution to the model also satisfies \( \max_{k \geq m} |\hat{y}_k^m - z_t| = 0 \), implying that \( z_t = y_t \) a.s., thereby showing part (ii) of Theorem 2. To show part (iii), note that, by strict stationarity of \( (y_t, \eta_t) \) and by noting that the conditional expectation is the best \( L_2 \)-approximation,

\[
\sup_{t \in \mathbb{Z}} E|y_t - E(y_t|\eta_{t-m}, \eta_{t-m+1}, \ldots, \eta_t)|^2
\]

\[ = E|y_t - E(y_t|\eta_{t-m}, \eta_{t-m+1}, \ldots, \eta_t)|^2 \]

\[ \leq E|y_t - \hat{y}_m^m|^2, \]

and because \( |y_t| + |\hat{y}_m^m| \leq 2f_t \), it now follows by the dominated convergence theorem that \( y_t \) is \( L_2 \)-near epoch dependent because \( E|f_t|^2 < \infty \) by assumption and by Lemma 1.

In order to obtain the explicit bound for \( \nu(M) \) of the last part of Theorem 2, note that

\[
\nu(m) = E(y_t - E(y_t|\eta_{t-m}, \ldots, \eta_t))^2 \leq E(y_t - \hat{y}_t^m)^2 I(|\hat{y}_t^m - y_t| > 0) \\
\leq (E|2f_t|^{2p})^{1/p} (P(|\hat{y}_t^m - y_t| > 0))^{1/q}
\]

for \( p \) and \( q \) such that \( p^{-1} + q^{-1} = 1 \). Now by choosing \( p \) small enough, \( E|f_t|^{2p} < \infty \) by assumption and by Lemma 1. Therefore, it suffices to show that \( P(|\hat{y}_t^m - y_t| > 0) \) decays as \( c_1 \exp(-c_2m^{1/3}) \) with \( m \). By the earlier reasoning,

\[
P[\max_{k \geq m} |\hat{y}_k^m - y_t| > 0] \\
\leq \exp(-(m-p)c) + P[(m-p)^{-1} \sum_{l=1}^{m-p} (I_l^m \log(\delta) + \log(1 + \delta)(1 - I_l^m)) > -c],
\]

27
and the last probability will decay as $c_1 \exp(-c_2 m^{1/3})$ by Lemma 2. This is because earlier, it was established that the summands are near epoch dependent with an exponentially decreasing $\nu(\cdot)$ sequence. The observations that $L_2$-near epoch dependent processes are also $L_1$-near epoch dependent and that $\alpha(M)$ decays exponentially by assumption now complete the proof. □

The consistency proofs for this paper rest upon the following lemma.

**Lemma 6** Assume that $z_t$ is strictly stationary and $L_2$-near epoch dependent on a strictly stationary strong mixing process $\eta_t$, and assume that $q(z, b)$ is continuous on $\mathbb{R}^a \times B$, where $B$ is a compact subset of $\mathbb{R}^c$. Then if $E \sup_{b \in B} |q(z_t, b)| < \infty$,

$$
\sup_{b \in B} \left| T^{-1} \sum_{t=1}^{T} (q(z_t, b) - Eq(z_t, b)) \right| \overset{p}{\to} 0.
$$

**Proof of Lemma 6:**


**Lemma 7** Under the conditions of Theorem 4,

$$(T - p)^{1/2} (\partial L_T(b) / \partial b) \big|_{b=\beta} \xrightarrow{d} N(0, I).$$

**Proof of Lemma 7:**

Note that by assumption, $E((\partial l_t(b) / \partial b) | y_{t-1}, \ldots, x_t) = 0$ so that $E(\partial l_t(b) / \partial b) |_{b=\beta} = 0$, implying that $(\partial l_t(b) / \partial b) |_{b=\beta}$ is a martingale difference sequence. In particular, by noting that $(y_t, x_t)$ has a “strong mixing base” in Bierens’ (2004) terminology, asymptotic normality now follows from the version of Bierens (2004, Theorem 7.11) of a central limit theorem of McLeish (1974). Applying the information matrix equality yields the result. □
Proof of Theorem 4:
We prove Theorem 4 by checking the conditions of Newey and McFadden (1994, Theorem 3.1). Consistency was shown in Theorem 3. Condition (i) was assumed. Condition (ii), twice differentiability of the log likelihood, follows from the Tobit specification. Condition (iii) was shown in Lemma 7. Note that stationarity and the strong mixing base imply ergodicity. Condition (iv) follows from the result of Lemma 6, the Tobit specification, and the assumption of finite second moments for $|x_t|$ and $\varepsilon_t$. Condition (v) is assumed. □

Proof of Theorem 5:
Under Assumption 4, it follows from the discussion in Powell (1984, p. 318) that $S_T(b)$ is uniquely minimized at $\beta = (\rho', \gamma')'$. From the assumptions of Assumption 4 it follows that $E\sup_{b \in B} |s(y_{t-1}, \ldots, y_{t-p}, x_t, \varepsilon_t, b)| < \infty$, and therefore the uniform law of large numbers of Lemma 6 applies. Therefore, all conditions of the consistency result of Theorem A1 of Wooldridge (1994) are satisfied. □

For the asymptotic normality result, we use the following lemma, which provides a suitable analogue to Powell’s lemma A3. For strictly stationary $(w_t, \varepsilon_t)$, let

$$\lambda(b) = E\psi(w_t, \varepsilon_t, b).$$

**Lemma 8** Assume that $(w_t, \varepsilon_t)$ is strictly stationary and that $|\hat{\beta}_T - \beta| = o_p(1)$. In addition assume that

$$T^{-1/2} \sum_{t=1}^{T} \psi(w_t, \varepsilon_t, \hat{\beta}_T) = o_p(1),$$

and assume that

$$T^{-1/2} \sum_{t=1}^{T} (\psi(w_t, \varepsilon_t, b) - E\psi(w_t, \varepsilon_t, b))$$

is stochastically equicontinuous on $B$. Then

$$T^{1/2} \lambda(b)|_{b=\beta} = -T^{-1/2} \sum_{t=1}^{T} \psi(w_t, \varepsilon_t, \beta) + o_p(1).$$

29
Proof of Lemma 8:

This follows by writing

$$o_p(1) = T^{-1/2} \sum_{t=1}^{T} \psi(w_t, \varepsilon_t, \hat{\beta}_T)$$

$$= T^{-1/2} \sum_{t=1}^{T} (\psi(w_t, \varepsilon_t, \hat{\beta}_T) - E\psi(w_t, \varepsilon_t, \hat{\beta}_T) - \psi(w_t, \varepsilon_t, \beta) + E\psi(w_t, \varepsilon_t, \beta))$$

$$+ T^{-1/2} \sum_{t=1}^{T} \psi(w_t, \varepsilon_t, \beta) + T^{1/2} \lambda(\hat{\beta}_T),$$

and noting that by the stochastic equicontinuity assumption, the first term in the last expression is $o_p(1)$ if $|\hat{\beta}_T - \beta| = o_p(1)$. 

\[\square\]

Remember that $w_t = (y_{t-1}, \ldots, y_{t-p}, x_t')' \text{ and } b = (c', r', s)'$. To show the stochastic equicontinuity of $T^{-1/2} \sum_{t=1}^{T} (\psi(w_t, \varepsilon_t, b) - E\psi(w_t, \varepsilon_t, b))$ and thereby obtain our analogue of Powell's Lemma A3, we first need the following results.

**Lemma 9** Assume $u_{1t} \in \mathbb{R}$ and $u_{2t} \in \mathbb{R}^q$, and assume that $(u_{1t}, u_{2t})'$ is strictly stationary and $L_r$-near epoch dependent, $r \geq 2$, on $\eta_t$ with a $\nu(\cdot)$ sequence satisfying $\nu(m) \leq c_1 \exp(-c_2 m^{1/3})$ for $c_1, c_2 > 0$. Then if $|u_{1t}| \leq 1$ and $\|u_{2t}\|_{r+\delta} < \infty$ for $\delta < r/(r-1)$, $u_{1t}u_{2t}$ is $L_r$-near epoch dependent on $\eta_t$ with a $\nu(\cdot)$ sequence that satisfies $\nu(m) \leq c_1 \exp(-c_2 m^{1/3})$ for constants $c_1, c_2 > 0$ (that are not necessarily identical to the earlier $c_1$ and $c_2$).

**Proof of Lemma 9:**

This follows by noting that

$$\| u_{1t}u_{2t} - E(u_{1t}u_{2t}|\eta_{t-M}, \ldots, \eta_t) \|_r$$

$$\leq \| u_{1t}u_{2t} - E(u_{1t}|\eta_{t-M}, \ldots, \eta_t)E(u_{2t}|\eta_{t-M}, \ldots, \eta_t) \|_r$$

$$\leq \| u_{1t}(u_{2t} - E(u_{2t}|\eta_{t-M}, \ldots, \eta_t)) \|_r + \| E(u_{2t}|\eta_t, \ldots, \eta_{t-m})(u_{1t} - E(u_{1t}|\eta_{t-M}, \ldots, \eta_t)) \|_r$$

$$\leq \| u_{2t} - E(u_{2t}|\eta_{t-M}, \ldots, \eta_t) \|_r$$
\[ + \| E(u_{2t}|\eta_{t-M}, \ldots, \eta_t) \|_{r+\delta} \| u_{1t} - E(u_{1t}|\eta_{t-M}, \ldots, \eta_t) \|_{r(1+r/\delta)} \]
\[ \leq \| u_{2t} - E(u_{2t}|\eta_{t-M}, \ldots, \eta_t) \|_r \]
\[ + \| E(u_{2t}|\eta_{t-M}, \ldots, \eta_t) \|_{r+\delta} \| u_{1t} - E(u_{1t}|\eta_{t-M}, \ldots, \eta_t) \|_{r(1+r/\delta)}, \]
and all terms in the last expression decay with \( m \) as specified.

We also need the following result.

**Lemma 10** Under Assumption 5 and the conditions of Theorem 2, for all \( \eta > 0 \), \( I(b'w_t > 0) \), \( I(b'w_t > -\eta|w_t|) \), \( I(b'w_t \leq \eta|w_t|) \), and \( I(\varepsilon_t + (\beta - b)'w_t > 0) \), \( I(\varepsilon_t + (\beta - b)'w_t > -\eta|w_t|) \), \( I(\varepsilon_t + (\beta - b)'w_t \leq \eta|w_t|) \) are \( L_r \)-near epoch dependent on \( \eta \) with a \( \nu(\cdot) \) sequence satisfying \( \nu(m) \leq c_1 \exp(-c_2m^{1/3}) \) for \( c_1, c_2 > 0 \).

**Proof of Lemma 10:**

We will show this for one case; the other cases are analogous. Note that \( w_t \) is near epoch dependent on \( \eta_t = \gamma'x_t + \varepsilon_t \) with \( \nu(\cdot) \) sequence satisfying \( \nu(m) \leq c_1 \exp(-c_2m^{1/3}) \) for \( c_1, c_2 > 0 \) by Theorem 2. In addition, for any \( \delta > 0 \), let \( T_\delta(\cdot) \) be a continuously differentiable function such that \( T_\delta(x) = I(x > 0) \) for \( |x| > \delta \) and \( \sup_{|x| \leq \delta} |(\partial/\partial x)T(x)| = K/\delta < \infty \). Then

\[ \| I(b'w_t + \eta|w_t| > 0) - E(I(b'w_t + \eta|w_t| > 0)|\eta_{t-M}, \ldots, \eta_t) \|_r \]
\[ \leq 2 \| I(b'w_t + \eta|w_t| > 0) - T_\delta(b'w_t + \eta|w_t|) \|_r \]
\[ + \| T_\delta(b'w_t + \eta|w_t| > 0) - E(T_\delta(b'w_t + \eta|w_t| > 0)|\eta_{t-M}, \ldots, \eta_t) \|_r \]
\[ \leq 2 \| (b'w_t + \eta|w_t| - \delta) \|_q + C_2\delta^{-1} \| b'w_t - E(b'w_t|\eta_{t-M}, \ldots, \eta_t) \|_r \]
\[ + C_3\delta^{-1} \| w_t - E(w_t|\eta_{t-M}, \ldots, \eta_t) \|_r \]
\[ \leq C_1\delta + C_3\nu(M)\delta^{-1} \]

because \( b'w_t + \eta|w_t| \) has a uniformly bounded density, which follows from the assumption that the density of \( \varepsilon_t w_t \) is uniformly bounded. Therefore by setting \( \delta = \nu(M)^{1/2} \), it follows that \( I(b'w_t + \eta|w_t| > 0) \) is near epoch dependent on \( \eta_t \) with a \( \nu(\cdot) \) sequence satisfying \( \nu(m) \leq c_1 \exp(-c_2m^{1/3}) \) for \( c_1, c_2 > 0 \) as well.

**Lemma 11** Under Assumption 5, \( T^{-1/2} \sum_{t=1}^{T} (\psi(w_t, \varepsilon_t, b) - E\psi(w_t, \varepsilon_t, b)) \) is stochastically equicontinuous on \( B \).
Proof of Lemma 11:

Note that for \( \psi(w_t, \varepsilon_t, b) \) we have that
\[
\mu_t(b, \delta) = \sup_{\tilde{b}: |b - \tilde{b}| < \delta} |\psi(w_t, \varepsilon_t, b) - \psi(w_t, \varepsilon_t, \tilde{b})|
\]
satisfies
\[
E \mu_t(b, \delta) \leq C \delta
\]
for some \( C > 0 \) under Assumption 5.2; see Equation (A.22) of Powell (1984) for this result. Therefore, we can cover \( B \) by \( O(\delta^{p+q}) \) balls with center \( b_j \) and radius \( \eta \) and we can define the bracketing functions as
\[
f^L_j(w_t) = I(b'_j w_t > 0)(1/2 - I(\varepsilon_t + (\beta - b) w_t > 0))w_t
\]
\[
- I(b'_j w_t > -\delta|w_t|)1/2 - I(\varepsilon_t + (\beta - b) w_t > -\delta|w_t|)||w_t|
\]
and
\[
f^U_j(w_t) = I(b'_j w_t > 0)(1/2 - I(\varepsilon_t + (\beta - b) w_t > 0))w_t
\]
\[
+ I(b'_j w_t > -\delta|w_t|)1/2 - I(\varepsilon_t + (\beta - b) w_t > -\delta|w_t|)||w_t|
\]
By the result of Lemma 10, the bracketing functions \( f^L_j(\cdot) \) and \( f^U_j(\cdot) \) as well as the \( \psi(w_t, \varepsilon_t, \beta) \) are \( L_r \)-near epoch dependent on \( \eta_t \) with an exponentially decreasing \( \nu(\cdot) \) sequence. By Equation (2) of Andrews (1988), \( L_r \)-near epoch dependent processes are also \( L_r \)-mixingales with mixingale numbers \( \nu(M) + \alpha(M)^{1/r - 1/(2r)} \) and uniformly bounded mixingale numbers.

We will now apply Theorem 3 of Hansen (1996); note that while Hansen’s smoothness condition with respect to the parameter on the function class under consideration does not hold in our situation, his argument will still go through, because his cover number and weak dependence conditions hold in exactly the same way as for his proof. For Hansen’s proof to work, we set Hansen’s constants \( \gamma, q \) and \( s \) equal to \( 1/2 \), \( r \) and \( 2r \) respectively, and we note that the bracketing functions \( f^L_j(\cdot) \) and \( f^U_j(\cdot) \) as well as \( \psi(w_t, \varepsilon_t, \beta) \) are also \( L_r \)-mixingales with mixingale numbers \( \nu(M)^{1/2} + \alpha(M)^{1/(2r) - 1/(4r)} \) and mixingale numbers \( c_t = \| f^L_j(w_t) \|_{2r}^{1/2}, c_t = \| f^L_j(w_t) \|_{2r}^{1/2}, \) or \( c_t = \| f^L_j(w_t) \|_{2r}^{1/2} \) respectively. The condition
\[
\sum_{M=0}^{\infty} (\nu(M)^{1/2} + \alpha(M)^{1/(4r)}) < \infty
\]
now corresponds to Hansen’s (1996) condition 2 of his Assumption 1, and Hansen’s condition \( q = a/(\lambda \gamma) \) now corresponds to, in our notation, \( r > (p + q)/(1/2) = 2(p + q) \).
Proof of Theorem 6:

We follow the asymptotic normality proof of Powell (1984). The strategy of our proof is to replace Powell’s Lemma A3 by the result of Lemma 8, and we note that under the conditions of Theorem 6, the stochastic equicontinuity condition of Lemma 8 follows from the result of Lemma 11. The remainder argument of Powell’s proof can be cast into the current framework in the following manner. It follows from the argument in Powell (1984, p.320) that, under Assumption 4 and 5 (because Powell’s E.1, R.1, and R.2 are met),

$$T^{-1/2} \sum_{t=1}^{T} \psi(w_t, \varepsilon_t, \tilde{\beta}_T) = o_p(1).$$

By Lemma 8 and Lemma 11, for \( \lambda(b) = E\psi(w_t, \varepsilon_t, b) \),

$$T^{1/2} \lambda(b)_{|b=\tilde{\beta}_T} = -T^{-1/2} \sum_{t=1}^{T} \psi(w_t, \varepsilon_t, \beta) + o_p(1).$$

It now follows from Lemma 9 and Lemma 10 that \( \psi(w_t, \varepsilon_t, b) \) is \( L_2 \)-near epoch dependent on \( \eta_t \) with an exponentially decreasing \( \nu(\cdot) \) sequence. Therefore by the central limit theorem of Theorem 2 of de Jong (1997), it follows that

$$T^{-1/2} \sum_{t=1}^{T} \psi(w_t, \varepsilon_t, \beta) \overset{d}{\rightarrow} N(0, \Omega).$$

Since \( \lambda(\beta) = 0 \) by assumption, for some mean value \( \beta_T^* \),

$$T^{1/2} \lambda(b)_{|b=\beta_T^*} = o_p(1) + (\partial/\partial b)\lambda(b)_{|b=\beta_T^*} T^{1/2}(\tilde{\beta}_T - \beta),$$

and identically to the discussion in Powell (1984, p. 320-321, equations A.16-A.19), it now follows that, under Assumptions 4.3, 5.2, and 5.5, (i.e. the analogues of Powell’s E.2, R.1, and R.2),

$$(\partial/\partial b)\lambda(b)_{|b=\beta_T} = o_p(1) + N.$$  

Therefore, it now follows that

$$T^{1/2}(\tilde{\beta}_T - \beta) \overset{d}{\rightarrow} N(0, N^{-1}\Omega N^{-1}),$$

as asserted by the theorem.  \qed