# Dynamic censored regression and the Open Market Desk reaction function 

Robert de Jong* Ana María Herrera ${ }^{\dagger}$

March 9, 2009


#### 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. The central analytical challenge is to prove that the dynamic censored regression model satisfies stationarity and weak dependence properties if a condition on the lag polynomial holds. 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

[^0]of a firm or individual leads to a corner response for a significant proportion of time. In addition, right-censoring may rise due to truncation choices made by the analysts in the process of collecting the data (i.e., top coding). Censored regression models apply to variables that are left-censored at zero, such as the level of open market operations or foreign exchange intervention carried out by a central bank, and in the presence of an intercept in the specification they also apply to time series that are censored at a non-zero point, such as the clearing price in commodity markets where the government imposes price floors, the quantity of imports and exports of goods subject to quotas, and 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). In recent years, asymptotic theory for the dynamic Tobit model in a panel data setting has been established using large- $N$ asymptotics; see Arellano and Honoré (1998) and Honoré and Hu (2004). However, there is no result in the literature that shows stationarity properties of the dynamic censored regression model, leaving the application of cross-section techniques for estimating the dynamic censored regression model in a time series setting formally unjustified. This paper seeks to fill this gap. After all, a justication of standard inference in dynamic nonlinear models requires laws of large numbers and a central limit theorem to hold. Such results require weak dependence and stationarity properties.

While in the case of linear AR models it is well-known that we need the roots of the lag polynomial to lie outside the unit circle in order to have stationarity, no such result is known for nonlinear dynamic models in general and the dynamic regression model in particular. The primary analytical issue addressed in this paper is to show that under some conditions, the dynamic censored regression model as defined below satisfies stationarity and weak dependence properties. This proof is therefore an analogue to well-known proofs of stationarity of ARMA models under conditions on the roots of the AR lag polynomial. The dynamic censored regression model under consideration is

$$
\begin{equation*}
y_{t}=\max \left(0, \sum_{i=1}^{p} \rho_{i} y_{t-i}+\gamma^{\prime} x_{t}+\varepsilon_{t}\right), \tag{1}
\end{equation*}
$$

where $x_{t}$ denotes the regressor, $\varepsilon_{t}$ is a regression error, 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 a linear process (i.e., an MA $(\infty)$ process driven by an i.i.d. vector of disturbances), which means it displays weak dependence and is possibly correlated. 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 is 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 $\rho_{\max }(z)=1-\sum_{i=1}^{p} \max \left(0, \rho_{i}\right) z^{i}$ 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 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 $\operatorname{AR}(p)$ 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 $\max \left(0, \rho_{j}\right)$ shows up in our stationarity requirement.

An alternative formulation for the dynamic censored regression model could be

$$
\begin{equation*}
y_{t}=y_{t}^{*} I\left(y_{t}^{*}>0\right) \quad \text { where } \quad \rho(B) y_{t}^{*}=\gamma^{\prime} x_{t}+\varepsilon_{t} \tag{2}
\end{equation*}
$$

where $B$ denotes the backward operator. This model will not be considered in this paper, and its fading memory properties are 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 today'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.

The literature on the dynamic Tobit model appears to mainly consist of (i) theoretical 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. Three noteworthy contributions to the literature on dynamic Tobit models are Honoré and Hu (2004), Lee (1999), and Wei (1999). Honoré and Hu (2004) considers dynamic Tobit models and deals with the problem of the endogeneity of lagged values of the dependent variable in panel data setting, where the errors are i.i.d., $T$ is fixed and large- $N$ asymptotics are considered. In fact, 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. For a treatment of the dynamic Tobit model in a panel setting, the reader is referred to Arellano and Honoré (1998, section 8.2).

Lee (1999) and Wei (1999) deal with dynamic Tobit models where lags of the latent variable are included as regressors. Lee (1999) considers likelihood simulation for dynamic Tobit models with ARCH disturbances in a time series setting. The central issue in this paper is the simulation of the log likelihood in the case where lags of the latent variable (in contrast to the observed lags of the dependent variable) have been included. Wei (1999) considers dynamic Tobit models in a Bayesian framework. The main contribution of this paper is the development of a sampling scheme for the conditional posterior distributions of the censored data, so as to enable estimation using the Gibbs sampler with a data augmentation algorithm.

In related work, de Jong and Woutersen (2003) consider the dynamic time series binary choice model and derive the weak dependence properties of this model. This paper also considers a formal large- $T$ asymptotic theory when lags of the dependent variable are included as regressors. Both this paper and de Jong and Woutersen (2003) allow 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.

As we mentioned above, a significant body of literature on the dynamic Tobit model consists of applications in a time series setting without providing a formal asymptotic theory. Inference in these papers is either conducted in a classical framework, by assuming the maximum likelihood estimates are asymptotically normal, or by employing Bayesian inference. Papers that estimate censored regression models in time series cover diverse topics. In the financial literature, prices subject to price limits imposed in stock markets, 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 historically intervened to support prices or to impose quotas. An example is provided by Chavas and Kim (2006) 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 open market operations and foreign exchange intervention. 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 $\left(y_{t}, x_{t}\right)$ 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. Section 6 concludes. 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

$$
\begin{equation*}
\sup _{t \in \mathbb{Z}} E\left|y_{t}-E\left(y_{t} \mid \eta_{t-M}, \eta_{t-M+1}, \ldots, \eta_{t+M}\right)\right|^{r}=\nu(M)^{r} \rightarrow 0 \quad \text { as } \quad M \rightarrow \infty \tag{3}
\end{equation*}
$$

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

$$
\begin{equation*}
y_{t}=\max \left(0, \sum_{i=1}^{p} \rho_{i} y_{t-i}+\eta_{t}\right) \tag{4}
\end{equation*}
$$

Later, we will set $\eta_{t}=\gamma^{\prime} x_{t}+\varepsilon_{t}$ 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 $\left(\eta_{t}, \eta_{t-1}, \ldots\right)$ will form a strictly stationary solution to the model.

An alternative way of proceeding to justify inference could be by considering arbitrary initial values $\left(y_{1}, \ldots, y_{p}\right)$ for the process instead of starting values drawn from the stationary distribution, but such an approach will be substantially more complicated.

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 $\rho_{\max }(B)=1-\sum_{i=1}^{p} \max \left(0, \rho_{i}\right) B^{i}$. 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 1 If the linear process $\eta_{t}$ satisfies $\eta_{t}=\sum_{i=0}^{\infty} a_{i} u_{t-i}$, where $a_{0}>0, u_{t}$ is a sequence of i.i.d. random variables with density $f_{u}(),. E\left|u_{t}\right|^{r}<\infty$ for some $r \geq 2$,

$$
\int_{-\infty}^{\infty}\left|f_{u}(y+a)-f_{u}(y)\right| d y \leq M|a|
$$

for some constant $M$ whenever $|a| \leq \delta$ for some $\delta>0, \sum_{t=0}^{\infty} G_{t}^{1 /(1+r)}<\infty$ where $G_{t}=$ $\left(\sum_{j=t}^{\infty} a_{j}^{2}\right)^{r / 2}, \rho_{\max }(z)$ has all its roots outside the unit circle, and for all $x \in \mathbb{R}$,

$$
\begin{equation*}
P\left(u_{t} \leq x\right) \geq F(x)>0 \tag{5}
\end{equation*}
$$

for some function $F($.$) , then (i) there exists a solution y_{t}$ to the model of Equation (4) such that $\left(y_{t}, \eta_{t}\right)$ is strictly stationary; (ii) if $z_{t}=f\left(\eta_{t}, \eta_{t-1}, \ldots\right)$ 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, $a_{i} \leq c_{1} \exp \left(-c_{2} i\right)$ for positive constants $c_{1}$ and $c_{2}$, then the near epoch dependence sequence $\nu(M)$ satisfies $\nu(M) \leq c_{1} \exp \left(-c_{2} M^{1 / 3}\right)$ for positive constants $c_{1}$ and $c_{2}$.

Our proof is based on the probability of $y_{t}$ reaching 0 given the last $p$ values of $\eta_{t}$ always being positive. This property is the key towards our proof and is established using the linear process assumption in combination with the condition of Equation (5). Note that by the results of Davidson (1994, p. 219), our assumption on $\eta_{t}$ implies that $\eta_{t}$ is also strong mixing with $\alpha(m)=O\left(\sum_{t=m+1}^{\infty} G_{t}^{1 /(1+r)}\right)$. Also note that for the dynamic Tobit model where errors are i.i.d. normal and regressors are absent, the condition of the above theorem simplifies to the assumption that $\rho_{\max }(z)$ has all its roots outside the unit circle.

One interesting aspect of the condition on $\rho_{\max }(z)$ 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 $\beta=\left(\rho^{\prime}, \gamma^{\prime}, \sigma\right)^{\prime}$, where $\rho=\left(\rho_{1}, \ldots, \rho_{p}\right)$, and define $b=\left(r^{\prime}, c^{\prime}, s\right)^{\prime}$ 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

$$
\begin{equation*}
L_{T}(b)=L_{T}(c, r, s)=(T-p)^{-1} \sum_{t=p+1}^{T} l_{t}(b), \tag{6}
\end{equation*}
$$

where

$$
\begin{align*}
l_{t}(b)= & I\left(y_{t}>0\right) \log \left(s^{-1} \phi\left(\left(y_{t}-\sum_{i=1}^{p} r_{i} y_{t-i}-c^{\prime} x_{t}\right) / s\right)\right) \\
& +I\left(y_{t}=0\right) \log \left(\Phi\left(\left(-\sum_{i=1}^{p} r_{i} y_{t-i}-c^{\prime} x_{t}\right) / s\right)\right) \tag{7}
\end{align*}
$$

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 linear process assumptions on $\left(x_{t}^{\prime}, \varepsilon_{t}\right)$, by the use of Theorem 1. It should be noted that even under the assumption of i.i.d. errors, no results regarding stationarity of the dynamic Tobit model have been derived in the literature thus far.

Let $\hat{\beta}_{T}$ denote a maximizer of $L_{T}(b)$ over $b \in B$. Define $w_{t}=\left(y_{t-1}, \ldots, y_{t-p}, x_{t}^{\prime}, 1\right)^{\prime}$. The " 1 " at the end of the definition of $w_{t}$ allows us to write " $b^{\prime} w_{t}$ ". For showing consistency, we need the following two assumptions. Below, let $|$.$| denote the usual matrix norm defined as$ $|M|=\left(\operatorname{tr}\left(M^{\prime} M\right)\right)^{1 / 2}$, and let $\|X\|_{r}=\left(E|X|^{r}\right)^{1 / r}$.

Assumption 1 The linear process $z_{t}=\left(x_{t}^{\prime}, \varepsilon_{t}\right)^{\prime}$ satisfies $z_{t}=\sum_{j=0}^{\infty} \Pi_{j} v_{t-j}$, where the $v_{t}$ are i.i.d. $(k \times 1)$ vectors, $\left\|v_{t}\right\|_{r}<\infty$ for some $r \geq 1$, the coefficient matrices $\Pi_{j}$ satisfy and $\sum_{t=0}^{\infty} G_{t}^{1 /(1+r)}<\infty$ where $G_{t}=\left(\sum_{j=t}^{\infty}\left|\Pi_{j}\right|^{2}\right)^{r / 2}, x_{t} \in \mathbb{R}^{q}$, and

$$
\begin{equation*}
y_{t}=\max \left(0, \sum_{i=1}^{p} \rho_{i} y_{t-i}+\gamma^{\prime} x_{t}+\varepsilon_{t}\right) . \tag{8}
\end{equation*}
$$

## Assumption 2

1. The linear process $z_{t}=\left(x_{t}^{\prime}, \varepsilon_{t}\right)^{\prime}$ satisfies $z_{t}=\sum_{j=0}^{\infty} \Pi_{j} v_{t-j}$, where the $v_{t}$ are i.i.d., $\left\|v_{t}\right\|_{r}<\infty$, and the coefficient matrices $\Pi_{j}$ satisfy $\sum_{t=0}^{\infty} G_{t}^{1 /(1+r)}<\infty$ where $G_{t}=$ $\left(\sum_{j=t}^{\infty}\left|\Pi_{j}\right|^{2}\right)^{r / 2}$.
2. Conditional on $\left(x_{1}, \ldots, x_{T}\right)$, $\varepsilon_{t}$ is independently normally distributed with mean zero and variance $\sigma^{2}>0$.
3. $\beta \in B$, where $B$ is a compact subset of $\mathbb{R}^{p+q+1}$, and $B=\Gamma \times R \times \Sigma$ where $\inf \Sigma>0$.
4. $E w_{t} w_{t}^{\prime} I\left(\sum_{i=1}^{p} \rho_{i} y_{t-i}+\gamma^{\prime} x_{t}>\delta\right)$ is positive definite for some positive $\delta$.

Theorem 2 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. $I=E(\partial / \partial b) l_{t}(\beta)\left(\partial / \partial b^{\prime}\right) l_{t}(\beta)=-E(\partial / \partial b)\left(\partial / \partial b^{\prime}\right) l_{t}(\beta)$ is invertible.

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

## 4 Powell's LAD for dynamic censored regression

For this section, define $\beta=\left(\rho^{\prime}, \gamma^{\prime}\right)^{\prime}$, where $\rho=\left(\rho_{1}, \ldots, \rho_{p}\right)$, define $b=\left(r^{\prime}, c^{\prime}\right)^{\prime}$ where $r$ is a $(p \times 1)$ vector and $c$ is a $(q \times 1)$ vector, and $w_{t}=\left(y_{t-1}, \ldots, y_{t-p}, x_{t}^{\prime}\right)^{\prime}$. 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

$$
\begin{align*}
S_{T}(b) & =S_{T}(c, r, s)=(T-p)^{-1} \sum_{t=p+1}^{T} s\left(y_{t-1}, \ldots, y_{t-p}, x_{t}, \varepsilon_{t}, b\right) \\
& =(T-p)^{-1} \sum_{t=p+1}^{T}\left|y_{t}-\max \left(0, \sum_{i=1}^{p} r_{i} y_{t-i}+c^{\prime} x_{t}\right)\right| \tag{9}
\end{align*}
$$

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 $F\left(\varepsilon_{t} \mid w_{t}\right)$ satisfies $F\left(0 \mid w_{t}\right)=1 / 2$, and $f\left(\varepsilon \mid w_{t}\right)=(\partial / \partial \varepsilon) F(\varepsilon \mid w)$ is continuous in $\varepsilon$ on a neighborhood of 0 and satisfies $c_{2} \geq f\left(0 \mid w_{t}\right) \geq c_{1}>0$ for constants $c_{1}, c_{2}>0$.
3. $E\left|x_{t}\right|^{3}<\infty$, and $E w_{t} w_{t}^{\prime} I\left(\sum_{i=1}^{p} \rho_{i} y_{t-i}+\gamma^{\prime} x_{t}>\delta\right)$ is nonsingular for some positive $\delta$.

Theorem 4 Under Assumptions 1 and 4, $\tilde{\beta}_{T} \xrightarrow{p} \beta$.
For asymptotic normality, we need the following additional assumption. Below, let

$$
\begin{equation*}
\psi\left(w_{t}, \varepsilon_{t}, b\right)=I\left(b^{\prime} w_{t}>0\right)\left(1 / 2-I\left(\varepsilon_{t}+(\beta-b)^{\prime} w_{t}>0\right)\right) w_{t} \tag{10}
\end{equation*}
$$

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

## Assumption 5

1. $\beta$ is in the interior of $B$.
2. Defining $G(z, b, r)=E I\left(\left|w_{t}^{\prime} b\right| \leq\left|w_{t}\right| z\right)\left|w_{t}\right|^{r}$, we have for $z$ near 0 , for $r=0,1,2$,

$$
\begin{equation*}
\sup _{|b-\beta|<\zeta_{0}}|G(z, b, r)| \leq K_{1} z \tag{11}
\end{equation*}
$$

3. The matrix

$$
\begin{equation*}
\Omega=\lim _{T \rightarrow \infty} E\left(T^{-1 / 2} \sum_{t=1}^{T} \psi\left(w_{t}, \varepsilon_{t}, \beta\right)\right)\left(T^{-1 / 2} \sum_{t=1}^{T} \psi\left(w_{t}, \varepsilon_{t}, \beta\right)\right)^{\prime} \tag{12}
\end{equation*}
$$

is well-defined, and $N=E f\left(0 \mid w_{t}\right) I\left(w_{t}^{\prime} \beta>0\right) w_{t} w_{t}^{\prime}$ is invertible.
4. For some $r \geq 2, E\left|x_{t}\right|^{2 r}<\infty, E\left|\varepsilon_{t}\right|^{2 r}<\infty$, and $\left|\Pi_{j}\right| \leq c_{1} \exp \left(-c_{2} j\right)$ for positive constants $c_{1}$ and $c_{2}$.
5. The conditional density $f\left(\varepsilon \mid w_{t}\right)$ satisfies, for a nonrandom Lipschitz constant $L_{0}$,

$$
\begin{equation*}
\left|f\left(\varepsilon \mid w_{t}\right)-f\left(\tilde{\varepsilon} \mid w_{t}\right)\right| \leq L_{0}|\varepsilon-\tilde{\varepsilon}| . \tag{13}
\end{equation*}
$$

Theorem 5 Under Assumptions 1, 4 and $5, T^{1 / 2}\left(\tilde{\beta}_{T}-\beta\right) \xrightarrow{d} N\left(0, N^{-1} \Omega N^{-1}\right)$.
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 5 imposes moment conditions of order 4 or higher. The conditions imposed by Theorem 5 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 5.

## 5 Simulations

In this section, we evaluate the consistency of the Tobit and CLAD estimators of the dynamic censored regression model. We consider the data generating process

$$
y_{t}=\max \left(0, \gamma_{1}+\gamma_{2} x_{t}+\sum_{i=1}^{p} \rho_{i} y_{t-i}+\varepsilon_{t}\right)
$$

where

$$
x_{t}=\alpha_{1}+\alpha_{2} x_{t-1}+v_{t},
$$

$\varepsilon_{t} \sim N\left(0, \sigma_{\varepsilon}^{2}\right)$, and $v_{t} \sim N\left(0, \sigma_{v}^{2}\right)$. For our simulations, we consider the cases $p=1$ and $p=2$. Many configurations for $\alpha_{1}, \alpha_{2}, \gamma_{1}, \gamma_{2}, \sigma_{v}^{2}$, and $\sigma_{\varepsilon}^{2}$ were considered. To conserve space, we only report results for $\gamma_{1}=1, \gamma_{2}=1, \alpha_{1}=\alpha_{2}=0.5, \sigma_{\varepsilon}^{2}=\sigma_{v}^{2}=1$.

For $p=1$, simulations were conducted for $\rho \in\{0,0.3,0.6,0.9\}$, whereas for $p=2$ we conducted simulations for $\left(\rho_{1}, \rho_{2}\right) \in\{(0.2,0.1),(0.5,0.1),(0.8,0.1),(0,-0.3),(0.3,-0.3)$, $(0.6,-0.3),(0.9,-0.3)\}$. Note that, in contrast with Honore and Hu (2004), in our simulations the values of $\rho_{i}$ are not restricted to be non-negative. For both the case $p=1$ and the case $p=2$ the number of replications used to compute the bias reported in the tables is 10,000.

Tables 4 reports simulation results for the dynamic Tobit model where estimates of $\beta=\left(\rho^{\prime}, \gamma^{\prime}, \sigma_{\varepsilon}\right)$, with $\rho^{\prime}=\rho$ if $p=1, \rho^{\prime}=\left(\rho_{1}, \rho_{2}\right)$ if $p=2$, and $\gamma^{\prime}=\left(\gamma_{1}, \gamma_{2}\right)$ are obtained via maximum likelihood. Results for Powell's LAD estimates of the dynamic censored regression model are reported in Table 5 where $\beta=\left(\rho^{\prime}, \gamma^{\prime}\right)$, with $\rho^{\prime}=\rho$ if $p=1, \rho^{\prime}=\left(\rho_{1}, \rho_{2}\right)$ if $p=2$, and $\gamma^{\prime}=\left(\gamma_{1}, \gamma_{2}\right)$. As we mentioned in section 4, because Powell's LAD estimator does not provide a first-round estimator of $\sigma_{\varepsilon}$ we redefine $\beta$ as to not include $\sigma_{\varepsilon}$. Estimates in this
case are obtained using the BRCENS algorithm proposed by Fitzenberger (1997a,b). We report results for $T=100,300,600,1000,2000$.

The simulations reveal that the maximum likelihood estimator for the dynamic Tobit model and Powell's LAD estimator of the dynamic censored regression model perform well for $T \geq 300$ (see Tables 4 and 5). As expected, the bias decreases as the sample size increases. For the same number of observations, the bias does not seem to vary much over the different values for $\rho$ and $\left(\rho_{1}, \rho_{2}\right)$ that were considered.

## 6 Empirical Application

In a time series framework, censored variables can arise when the dynamic optimization behavior of a firm, individual or policy maker leads to a corner response for a significant proportion of time. Thus, is not surprising that dynamic Tobit models have been estimated to study a number of variables such as open market operations (Demiralp and Jordà, 2002) and central bank intervention in foreign exchange markets (Kim and Sheen, 2002). What is surprising is that inference is conducted using the t-statistic critical values, without having considered formal issues of stationarity. As we have noted before, strict stationarity and ergodicity of the dynamic censored regression model is required to show asymptotic normality and consistency of the maximum likelihood estimator of the dynamic Tobit model and of Powell's LAD estimator of the dynamic censored regression.

In what follows we discuss an application of the dynamic censored regression model to the Open Market Desk reaction function. Although there are a significant number of papers that model and estimate the Federal Open Market Committee's reaction function (e.g. Feinman, 1993, Demiralp and Farley, 2005), we are only aware of a recent study where lags of the dependent variable (i.e. open market operations) are included among the regressors. Without having considered formal issues of stationarity, Demiralp and Jordà (2002) estimated 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 the following section we re-evaluate their findings.

### 6.1 Data and Summary of Previous Results

The data used by Demiralp and Jordà (2002) to study the announcement effect on the Open Market Desk reaction function are daily and span the period between April 25, 1984 and August 14, 2000. They divide the sample in three subsamples. The first subsample corresponds to the period preceding the Fed decision to publicly announce changes in the
federal fund rate target on February 4, 1994. The second period spans the days between February 4, 1994 and the decision to shift from contemporaneous reserve accounting (CRA) to lagged reserves accounting (LRA) system in August 17, 1998. The last subsample covers the period following the shift to the CRA system.

Demiralp and Jordà (2002) classify open market operations in six groups according to whether they inject or drain liquidity and to the permanence of the operation. Operations that add liquidity can be grouped into overnight reversible repurchase agreements (OB), term repurchase agreements (TB), and permanent purchases (PB), which include T-bill purchases and coupon purchases. Operations that drain liquidity can be grouped into overnight sales (OS), term matched-sale purchases (TS), and permanent sales (PS), which comprise T-bill sales and coupon sales.

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 maintenanceperiod average is the object of attention of the Open Market 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.

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

$$
\begin{align*}
& N E E D_{t}=f_{t}-\left[f_{m(t)-1}^{*}+w_{t} E_{m(t)-1}\left(\Delta f_{m(t)}^{*}\right)\right]  \tag{14}\\
& E X P E C T_{t}=E_{m(t)-1}\left(\Delta f_{m(t)}^{*}\right)  \tag{15}\\
& S U R P R I S E E_{t}=\Delta f_{t}^{*}-E_{m(t)-1}\left(\Delta f_{m(t)}^{*}\right) \tag{16}
\end{align*}
$$

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 to the one to which observation $t$ belongs; $E_{m(t)-1}\left(\Delta f_{m(t)}^{*}\right)$ 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, $E_{m(t)-1}\left(\Delta f_{m(t)}^{*}\right)$, 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 or drain liquidity in order to accommodate shocks to the demand for reserves; (2) to accommodate expectations of future changes in the target; and (3) to adjust to a new target level. Thus, $N E E D_{t}$ represents a proxy for the projected reserve need, and changes in the federal funds rate are separated into an expected component, $E X P E C T_{t}$, and a surprise component, $S U R P R I S E_{t}$. 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.

Because the Desk engages in open market operations only on $60 \%$ of the days in the sample, and even the most common operation only takes place on $35 \%$ of the days (i.e., the data is censored at zero during a large number of days), Demiralp and Jordà (2002) use a Tobit model to analyze the reaction function of the Open Market Desk. Furthermore, to allow for a different response of sales and purchases -with varying degrees of permanenceto changes in the explanatory variables they estimate separate regressions for each of the six types of operation and each of the periods of interest. Because very few term and permanent sales were carried out during the 1998-2000 and 1984-1994 periods, no regressions are estimated for this type of operation in these subsamples. Hence, a total of sixteen regression are estimated. Demiralp and Jordà (2002) use the following dynamic Tobit model to describe open market operations carried out by the Open Market Desk:

$$
\begin{align*}
y_{t}= & \max \left(0, \sum_{m=1}^{10} \gamma_{m}^{\alpha} D A Y_{t m}+\sum_{j=1}^{3} \rho_{j} y_{t-j}+\sum_{j=1}^{3} v_{j}^{\prime} \mathbf{z}_{t-j}+\sum_{m=1}^{10} \gamma_{m}^{N} N E E D_{t-m} \times D A Y_{t m}\right. \\
& \left.+\sum_{m=1}^{10} \gamma_{m}^{E} E X P E C T_{t-m} \times D A Y_{t m}+\sum_{j=0}^{3} \gamma_{j}^{S} S U R P R I S E_{t-j}+\varepsilon_{t}\right) \tag{17}
\end{align*}
$$

where $y_{t}$ denotes one of the open market operation of interest, that is, $y_{t}$ equals either overnight purchases $\left(O B_{t}\right)$, term purchases $\left(T B_{t}\right)$, permanent purchases $\left(P B_{t}\right)$, overnight sales $\left(O S_{t}\right)$, term sales $\left(T S_{t}\right)$, or permanent sales $\left(P S_{t}\right)$. $\mathbf{z}_{t}$ denotes a vector containing the remaining five types of operations. For instance, if $y_{t}=O B_{t}$ (overnight purchases), then $\mathbf{z}_{t}=\left[T B_{t}, P B_{t}, O S_{t}, T S_{t}, P S_{t}\right] . D A Y_{t m}$ denotes a vector of maintenance-day dummies, and $\varepsilon_{t}$ is a stochastic disturbance.

We thus start our empirical analysis by re-estimating Demiralp and Jordà's (2002) specifications under the assumption of normality. That is, we follow their lead in assuming the dynamic Tobit model is correctly specified. We report the coefficient estimates for the lags of the dependent variable in Table 1. See Tables A. 1 and A. 2 in the Appendix for the complete set of parameter estimates. Because we are interested in whether the roots of the polynomial $\rho_{\max }(z)=1-\sum_{i=1}^{3} \max \left(0, \rho_{i}\right) z^{i}$ are outside the unit circle we report the smallest of the moduli of the roots of this lag polynomial.

Note that 10 out of the 16 regressions estimated by Demiralp and Jordà (2002) appear to have at least one root that falls on or inside the unit circle. One may wonder whether this result stems from nonstationarity issues or from misspecification in the error distribution. To investigate this issue, we test for normality of the Tobit residuals and report the Jarque-Bera statistics in Table 1; these results lead us to reject the null that the underlying disturbances are normally distributed. Thus, we proceed in the following section to estimate the Open Desk's reaction function using Powell's LAD estimator, which is robust to unknown error
distributions. If the problem is one of nonstationarity, one would then expect the roots of the $\rho_{\max }($.$) polynomial to be on or inside the unit circle. However, in what follows we will$ see that CLAD estimates for temporary Open Market purchases indicate that the roots of the $\rho_{\max }($.$) polynomial are outside the unit circle. That finding suggests that the results$ of Demiralp and Jordà (2002) suffer from misspecification in the error distribution, but not from nonstationarity issues.

### 6.2 Model and estimation procedure

From here on we will restrict our attention to the Desk's reaction function for temporary open market purchases over the whole 1984-2000 sample. We focus on temporary purchases because overnight and term RPs are the most common operations carried out by the Open Market Desk; thus, they are informative regarding the Desk's reaction function. In fact, daily values of temporary purchases plotted in Figure 1 reveal that the Open Market Desk engaged in temporary purchases on $37 \%$ of the days between April 25, 1984 and August 14, 2000. In contrast, permanent purchases, temporary sales, and permanent sales were carried out, respectively, on $24 \%, 7 \%$, and $2 \%$ of the days in the sample.

Thus, in contrast with Demiralp and Jordà (2002) we re-classify open market operations in four groups: (a) temporary purchases, which comprise overnight reversible repurchase agreements (RP) and term RP, $O T B_{t}=O B_{t}+T B_{t}$; (b) permanent purchases, which include T-bill purchases and coupon purchases, $P B_{t}$; (c) temporary sales, which include overnight and term matched sale-purchases, $O T S_{t}=O S_{t}+T S_{t}$; and (d) permanent sales, which comprise T-bill sales and coupon sales, $P S_{t}$. In brief, we group overnight and term operations and restrict our analysis to the change in the maintenance-period-average level of reserves brought about by temporary purchases of the Open Market Desk, OTB $\boldsymbol{B}_{t}$.

We employ the following dynamic censored regression model to describe temporary purchases by the Open Market Desk:

$$
\begin{align*}
O T B_{t} & =\max \left(0, \gamma+\sum_{m=1}^{4} \gamma_{m}^{\alpha} D_{t m}+\sum_{j=1}^{3} \rho_{j} O T B_{t-j}+\sum_{j=1}^{3} \gamma_{j}^{T S} O T S_{t-j}+\sum_{j=1}^{3} \gamma_{j}^{P B} P B_{t-j}\right. \\
& +\sum_{j=1}^{3} \gamma_{j}^{P S} P S_{t-j}+\sum_{m=1}^{10} \gamma_{m}^{N} N E E D_{t-m} \times D A Y_{t m}+\sum_{m=1}^{10} \gamma_{m}^{E} E X P E C T_{t-m} \times D A Y_{t m} \\
& \left.+\sum_{j=0}^{3} \gamma_{j}^{S} S U R P R I S E_{t-j}+\varepsilon_{t}\right) \tag{18}
\end{align*}
$$

where $O T B_{t}$ denotes temporary purchases, $O T S_{t}$ denotes temporary sales, $P B_{t}$ denotes permanent purchases, $P S_{t}$ denotes permanent sales, $D A Y_{t m}$ denotes a vector of maintenanceday dummies, $D_{t m}$ is such that $D_{t 1}=D A Y_{t 1}$ (First Thursday), $D_{t 2}=D A Y_{t 2}$ (First Friday), $D_{t 3}=D A Y_{t 7}$ (Second Friday), and $D_{t 4}=D A Y_{t, 10}$ (Settlement Wednesday), and $\varepsilon_{t}$ is a stochastic disturbance.

This model is a restricted version of (17) 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 certain days of the maintenance period where the target level of reserves is expected to be different from the average (see Demiralp and Farley, 2005). 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.

Regarding the estimation procedure, Tobit estimates $\widehat{b}$ are obtained in the usual manner via maximum likelihood estimation, whereas the CLAD estimates $\widetilde{b}$ are obtained by using the BRCENS algorithm proposed by Fitzenberger (1997a,b). Extensive Monte Carlo simulations by Fitzenberger (1997a) suggest that this algorithm, which is an adaptation of the BarrodaleRoberts algorithm for the censored quantile regression, performs better than the iterative linear programming algorithm (ILPA) of Buchinsky (1994) and the modified ILPA algorithm (MILPA) of Fitzenberger (1994), in terms of the percentage of times it detects the global minimum of a censored quantile regression. In fact, for our application, a grid search over 1000 points in the neighborhood of the estimates $\widetilde{b}$ indicates both the ILPA and MILPA algorithms converge to a local minimum. In contrast, the BRCENS algorithm is stable and appears to converge to a global minimum.

Because the CLAD does not provide a first-round estimate for the variance, $N^{-1} \Omega N^{-1}$, we compute it in the following manner. $\widetilde{\Omega}$ is calculated as the long-run variance of $\widetilde{\psi}\left(w_{t}, \widetilde{b}\right)=$ $I\left(\widetilde{b^{\prime}} w_{t}>0\right)\left[\frac{1}{2}-I\left(y_{t}<\widetilde{b}^{\prime} w_{t}\right)\right] w_{t}$, following the suggestions of Andrews (1991) to select the bandwidth for the Bartlett kernel. To compute $\widetilde{N}$, we estimate $f\left(0 \mid w_{t}\right)$ using a higher-order Gaussian kernel with the order and bandwidth selected according to Hansen (2003, 2004).

### 6.3 Estimation Results

Maximum likelihood estimates of the dynamic Tobit model for the entire sample, and corresponding standard errors are presented in the first two columns of Table 2. ${ }^{1}$ Before we comment on the estimation results, it is important to inspect whether the roots of the lag polynomial $\rho_{\max }(z)$ lie outside the unit circle. The three roots of $\widehat{\rho}_{\max }(z)=$ $1-0.2639 z-0.2916 z^{2}-0.3054 z^{3}$ lie all outside the unit circle, and the smallest modulus

[^1]of these roots equals 1.075. Because this root is near the unit circle and because we do not have the tools to test if it is statistically greater than one, we should proceed with caution.

Of interest is the presence of statistically significant coefficients on the lags of the dependent variable, $T B_{t-j}$. 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, $T S_{t-j}$, imply that temporary sales 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 $N E E D_{t-1} \times D A Y_{t m}$ 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 days of the maintenance period, and did not significantly affect temporary purchases on most of 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 $S U R P R I S E_{t-j}$ can be interpreted as evidence of the announcement effect. ${ }^{2}$ This suggests 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, it is well known that the Tobit estimates are inconsistent if the underlying disturbances are heteroskedastic or non-normal (Greene, 2000). Thus, to assess whether the Tobit specification of the reaction function is appropriate, we conduct tests for homoskedasticity and normality. A Lagrange multiplier test of heteroskedasticity obtained by assuming $\operatorname{Var}\left(\varepsilon_{t} \mid w_{t}\right)=\sigma^{2} \exp \left(\delta^{\prime} z_{t}\right)$, 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 statistic 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.

Summing up, the finding of a root that is close to the unit circle in conjunction with the rejection of the normality and homoskedasticity assumptions suggest that the Tobit estimates could be biased. Hence, our finding of a root near the unit circle may stem either from misspecification of the error term or from non-stationarity of the dynamic Tobit model. To further investigate this issue, we consider the CLAD estimator, which is robust to heteroskedasticity and nonnormality and is consistent in the presence of weakly dependent errors (see Section 4). Finding a root close to unity for the CLAD estimates would be indica-

[^2]tive of nonstationarity in the dynamic censored regression model driving the test results. In contrast, finding roots that are outside the unit circle would point towards misspecification of the error distribution being the cause of the bias in the Tobit estimates.

CLAD estimates and corresponding standard errors are reported in the third and fourth column of Table 2, respectively ${ }^{3}$. Notice that, in this case, the smallest root of the lag polynomial $\widetilde{\rho}_{\max }(z)=1-0.068 z-0.093 z^{2}-0.073 z^{3}$ appears to be clearly outside the unit circle. Here the smallest modulus of the roots equals 1.928. Given that the roots are far from the unit circle, standard inference techniques seem to be asymptotically justified. Furthermore, this suggest that our finding of roots that are near the unit circle for the Tobit model is a consequence of misspecification in the error term as normal and homoskedastic.

Comparing the CLAD and the Tobit estimates reveals some differences regarding the Desk's reaction function. First, the CLAD estimates imply a considerably smaller degree of persistence in temporary purchases. The magnitude of the $\rho_{j}, j=1,2,3$, parameter estimates is at most $1 / 3$ of the Tobit estimates. Consequently, the roots of the lag polynomial $\widetilde{\rho}_{\max }(z)$ implied by the CLAD estimates are larger, giving us confidence regarding stationarity of the censored regression model.

Second, although both estimates imply a similar reaction of the Fed to reserve needs, there are some differences in the magnitude and statistical significance of the parameters. In particular, the CLAD estimates suggest a pattern in which the Fed is increasingly less reluctant to intervene during the first three days of the maintenance period; then, no significant response is apparent for the following four days (with the exception of Day5, the first Wednesday); and finally, the response to reserve needs becomes positive and significant for the last three days of the period. Furthermore, on Mondays (Day3 and Day8), the Desk appears to be more willing to accommodate shocks in the demand for reserves in order to maintain the federal funds rate aligned with the target.

The expectation of a change in the target seldom triggers temporary open market purchases. The coefficient on $E X P E C T$ is statistically significant on the first and eight day of the maintenance period, and marginally significant on the second and sixth day. This suggest the Fed in only seldom willing to accommodate (or profit) from anticipated changes in the target. Moreover, even though both estimation methods point towards a larger effect on the first day, the CLAD estimate (40.5) suggest an impact that is about $67 \%$ smaller than the Tobit estimate (121.5).

The majority of the coefficients on the contemporaneous and lagged SURPRISE have a negative sign, which is consistent with the liquidity effect. That is, in order to steer the

[^3]federal funds rate towards a new lower target level the Desk would add liquidity by using temporary purchases. Yet, the fact that none of the coefficients are statistically significant suggests that once the target was announced (or inferred by the financial markets) little additional pressure was needed to enforce the new target.

Finally, to further explore the "announcement effect" we redo the estimation using only the observations after the Fed started announcing the new target level in February 4, 1994 (see Table 3) ${ }^{4}$. There appears to be a somewhat smaller but still significant degree of persistence in temporary open market purchases after the decision to announce the target level. This suggest the Fed still had to exert some pressure on the market to drive the fed funds rate towards the new target level. ${ }^{5}$ In brief, we find the coefficients on the current and lagged values of SURPRISE to be statistically insignificant whether we include or not the observations that predate the Fed decision to announce the target level.

## 7 Conclusions

This paper shows stationarity properties of the dynamic censored regression model in a time series framework. It then provides a formal justification for maximum likelihood estimation of the dynamic Tobit model and for Powell's LAD estimation of the dynamic censored regression model, showing consistency and asymptotic normality of both estimators. Two important features of the treatment of the censored regression model in this paper is that no assumption is made on the lag polynomial other than that $\rho_{\max }(z)=1-\sum_{i=1}^{p} \max \left(0, \rho_{i}\right) z^{i}$ has its roots outside the unit circle and that the error term, $\varepsilon_{t}$, is itself allowed to be potentially correlated. Hence, in terms of the conditions on $\rho_{\max }(z)$ and the dependence allowed for $\varepsilon_{t}$, this paper analyzes the dynamic censored regression model on a level of generality that is comparable to the level of generality under which results for the linear model $\operatorname{AR}(p)$ model can be derived.

The censored regression model is then applied to study the Open Market Desk's reaction function. Robust estimates for temporary purchases using Powell's CLAD suggest that maximum likelihood estimates of the dynamic Tobit model may lead to overestimating the persistence of temporary purchases, as well as the effect of demand for reserves and expectations of future changes in the federal funds target on temporary purchases. Moreover, a comparison of the Tobit and CLAD estimates suggests that temporary purchases are stationary, but that the error normality assumed in the Tobit specification does not hold.

[^4]
## References

Amemiya, T. (1973), Regression analysis when the dependent variable is truncated normal, Econometrica 41, 997-1016.

Amemiya, T. (1985) Advanced Econometrics. Oxford: Basil Blackwell.
Andrews, D.W.K. (1987), Consistency in nonlinear econometric models: a generic uniform law of large numbers, Econometrica 55, 1465-1471.

Andrews, D.W.K. (1988), Laws of large numbers for dependent non-identically distributed random variables, Econometric Theory 4, 458-467.
Andrews, D.W.K. (1991), Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation, Econometrica 59, 817-858.

Arellano, M. and B. Honoré (1998), Panel data models: some recent developments, Chapter 53 in Handbook of Econometrics, Vol. 5, ed. by James J. Heckman and Edward Leamer. Amsterdam: North-Holland.

Azuma, K. (1967), Weighted sums of certain dependent variables, Tokohu Mathematical Journal 19, 357-367.

Bierens, H. J. (2004) Introduction to the mathematical and statistical foundations of Econometrics. Cambridge University Press, forthcoming, available at http://econ.la.psu.edu/~hbierens/CHAPTER7.PDF .

Bougerol, P. and N. Picard (1992), Strict stationarity of generalized autoregressive processes, Annals of Probability 20, 1714-1730.

Buchinsky, M. (1994), Changes in the U.S. wage structure 1963-1987: application of quantile regression, Econometrica 62, 405-458.
Chavas, J-P. and Kim (2006), An econometric analysis of the effects of market liberalization on price dynamics and price volatility, Empirical Economics 31, 65-82.

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, 167-180.

Chernozhukov, V. and H. Hong (2002), Three-step censored quantile regression and extramarital affairs, Journal of the American Statistical Association 97 (459), 872-882.

Davidson, J. (1994) Stochastic Limit Theory. Oxford: Oxford University Press.
de Jong, R.M. (1997), Central limit theorems for dependent heterogeneous random variables,

Econometric Theory 13, 353-367.
de Jong, R.M. and T. Woutersen, Dynamic time series binary choice, forthcoming, Econometric Theory.

Demiralp, S. and D. Farley (2005), Declining required reserves, funds rate volatility, and open market operations, Journal of Banking and Finance 29, 1131-1152.
Demiralp, S. and O. Jordà (2002), The announcement effect: evidence from open market desk data, Economic Policy Review, 29-48.

Demiralp, S. and O. Jordà (2004), The response of term rates to Fed announcements, Journal of Money, Credit, and Banking 36, 387-406.
Feinman, J. (1993), Estimating the Open Market Desk's daily reaction function, Journal of Money, Credit, and Banking 25, 231-247.
Frenkel, M., Pierdzioch, C. and G. Stadtmann (2003), Modeling coordinated foreign exchange market interventions: The case of the Japanese and U.S. interventions in the 1990s, Review of World Economics, 139, 709-729.

Fitzenberger, B. (1994), A note on estimating censored quantile regressions, Discussion Paper 14, University of Konstanz.
Fitzenberger, B. (1997a), "Computational Aspects of Censored Quantile Regression," in Proceedings of the 3rd International Conference on Statistical Data Analysis Based on the L1-Norm and Related Methods, ed. Y. Dodge, Hayward, CA: IMS, 171-186.
Fitzenberger, B. (1997b), "A Guide to Censored Quantile Regressions," in Handbook of Statistics, Vol. 14, Robust Inference, Amsterdam: North-Holland, pp. 405-437.

Gallant, A.R. and H. White (1988) A unified theory of estimation and inference for nonlinear dynamic models. New York: Basil Blackwell.

Greene, W. (2000) Econometric Analysis, 5th Edition. New Jersey: Prentice-Hall.
Hamilton, J.D. and O. Jordà (2002), A Model for the Federal Funds Rate Target, Journal of Political Economy, 110, 1135-1167.
Hansen, B. E. (1996), Stochastic equicontinuity for unbounded dependent heterogeneous arrays, Econometric Theory, 12, 347-359.

Hansen, B.E. (2003), Exact mean integrated square error of higher-order kernel estimators, Working Paper, University of Wisconsin.
Hansen, B.E. (2004), Bandwidth selection for nonparametric kernel estimation, Working Pa-
per, University of Wisconsin.
Kim, S. and J. Sheen (2002), The determinants of foreign exchange intervention by central banks: evidence from Australia, Journal of International Money and Finance, 21, 619-649.
Kodres, L. E. (1988), Test of unbiasedness in the foreign exchange futures markets: The effects of price limits, Review of Futures Markets 7, 139-166.
Kodres, L.E. (1993), Test of unbiasedness in the foreign exchange futures makets: An Examination of Price Limits and Conditional Heteroscedasticity, Journal of Business 66, 463-490.
Kuttner, Kenneth (2001), Monetary policy surprises and interest rates: evidence from the Fed funds futures market, Journal of Monetary Economics 47, 523-544.

McLeish, D. L. (1974), Dependent central limit theorems and invariance principles, Annals of Probability 2, 620-628.
Meyn, S.P. and R.L. Tweedie (1994) Markov Chains and Stochastic Stability. Berlin: SpringerVerlag, second edition.
Newey, W. K., and D. McFadden (1994), Large sample estimation and hypothesis testing, in Handbook of Econometrics, Vol. 4, ed. by R. F. Engle and D. MacFadden. Amsterdam: North-Holland.

Persitiani, S. (1994), An empirical investigation of the determinants of discount window borrowing: a disaggregate analysis, Journal of Banking and Finance 18, 183-197.
Phillips, P. and V. Solo (1992), Asymptotics for linear processes, Annals of Statistics, Vol. 20, No. 2, p. 971-1001.
Pötscher, B.M. and I.R. Prucha (1986), A class of partially adaptive one-step M-estimators for the nonlinear regression model with dependent observations, Journal of Econometrics 32, 219-251.

Pötscher, B.M. and I.R. Prucha (1997) Dynamic nonlinear econometric models. Berlin: Springer-Verlag.
Powell, J.L. (1984), Least absolute deviations estimation for the censored regression model, Journal of Econometrics 25, 303-325.
Wei, S.X. (1999), A Bayesian approach to dynamic Tobit models, Econometric Reviews 18, 417-439.

Wei, S.X. (2002), A censored-GARCH model of asset returns with price limits, Journal of Empirical Finance 9, 197-223.
Wooldridge, J. (1994), Estimation and inference for dependent processes, in Handbook of

Econometrics, volume 4, ed. by R. F. Engle and D. MacFadden. Amsterdam: NorthHolland.
Zangari, P.J. and Tsurumi, H. (1996), A Bayesian analysis of censored autocorrelated data on exports of Japanese passenger cars to the United States, Advances in Econometrics, Volume 11, Part A, 111-143.


Figure 2: Standardized residuals

Table 1
Demiralp and Jordà (2002) Tobit regression for Open Market Operations Coefficient Estimates for Lags of the Dependent Variable


[^5]Table 2
Tobit and CLAD Estimates for Open Market Temporary Purchases 1986-2000

| Variable | Tobit |  | CLAD |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Estimate | Std. Err. | Estimate | Std. Err. |
| Constant | -18.742 *** | 1.753 | -1.252 *** | 0.341 |
| First Thursday | 16.962 *** | 2.833 | 1.814 *** | 0.502 |
| First Friday | -17.046 *** | 2.917 | -3.595 *** | 1.393 |
| Second Friday | -12.051 *** | 2.683 | -3.466 *** | 1.262 |
| Settlement Wednesday | 4.456 *** | 1.553 | 2.811 *** | 0.464 |
| OTB(-1) | 0.264 *** | 0.035 | 0.068 *** | 0.008 |
| OTB(-2) | 0.292 *** | 0.042 | 0.093 *** | 0.007 |
| OTB(-3) | 0.305 *** | 0.049 | 0.073 *** | 0.007 |
| OTS(-1) | $-1.726^{* * *}$ | 0.611 | -3.941 | 3.666 |
| OTS(-2) | -0.865 * | 0.447 | -1.025 | 1.554 |
| OTS(-3) | -1.895 *** | 0.408 | -0.558 * | 0.359 |
| PB(-1) | -0.018 | 0.085 | -0.164 *** | 0.065 |
| PB(-2) | -0.065 | 0.074 | -0.045 | 0.038 |
| PB(-3) | -0.073 | 0.072 | -0.022 | 0.031 |
| PS(-1) | 0.146 | 0.253 | -0.033 | 0.116 |
| PS(-2) | -0.151 | 0.245 | -0.067 | 0.092 |
| PS(-3) | -0.198 | 0.223 | 0.065 * | 0.045 |
| SURPRISE | -14.046 | 15.387 | -0.659 | 4.441 |
| SURPRISE(-1) | 11.516 | 13.044 | -0.268 | 3.898 |
| SURPRISE(-2) | -14.619 | 17.648 | 0.735 | 3.849 |
| SURPRISE(-3) | -7.259 | 14.832 | -1.755 | 4.08 |
| NEED(-1)*Day1 | -0.501 | 2.974 | 1.742 *** | 0.526 |
| NEED(-1)*Day2 | 11.645 ** | 4.691 | 3.433 *** | 0.794 |
| NEED(-1)*Day3 | 24.030 *** | 8.595 | 9.163 *** | 2.108 |
| NEED(-1)*Day4 | -7.489 | 7.775 | -1.868 | 1.822 |
| NEED(-1)*Day5 | 21.671 *** | 7.612 | 5.884 *** | 2.012 |
| NEED(-1)*Day6 | 5.312 | 10.941 | -0.593 | 2.478 |
| NEED(-1)*Day7 | 33.429 *** | 10.471 | 1.367 | 2.362 |
| NEED(-1)*Day8 | 6.842 | 7.789 | 9.616 *** | 2.213 |
| NEED(-1)*Day9 | 13.402 *** | 4.953 | 5.206 *** | 1.248 |
| NEED(-1)*Day10 | 3.972 * | 2.083 | 1.144 ** | 0.52 |
| EXPECT*Day1 | 121.451 ** | 53.231 | 40.461 *** | 5.499 |
| EXPECT*Day2 | 48.599 | 34.018 | 13.107 * | 8.277 |
| EXPECT*Day3 | -14.656 | 38.828 | 1.884 | 6.077 |
| EXPECT*Day4 | -25.528 | 29.346 | 3.586 | 5.711 |
| EXPECT*Day5 | -54.997 * | 29.859 | -1.924 | 7.271 |
| EXPECT*Day6 | 51.573 * | 30.636 | 8.856 * | 5.449 |
| EXPECT*Day ${ }^{\text {a }}$ | -37.603 | 36.721 | 1.632 | 12.704 |
| EXPECT*Day8 | 39.313 * | 20.515 | 12.535 ** | 5.454 |
| EXPECT*Day9 | -48.249 * | 24.666 | -11.71 | 9.445 |
| EXPECT*Day10 | 16.269 | 16.363 | 7.958 * | 5.168 |
| SCALE | 1094.099 *** | 104.478 |  |  |
| Smallest root | 1. 0750 |  | 2. 0315 |  |

Note: ***, ** and * denote significance at the 1, 5 and 10\% level, respectively. "Smallest root" denotes the smallest modulus of the roots of the $\rho_{\text {max }}(B)$ lag polynomial.

Table 3
Tobit and CLAD Estimates for Open Market Temporary Purchases 1994-2000

| Variable | Tobit |  | CLAD |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Estimate | Std. err. | Estimate | Std. Err. |
| Constant | -3.856 ** | 1.904 | 3.318 *** | 0.379 |
| First Thursday | 24.899 *** | 4.520 | 10.248 *** | 0.766 |
| First Friday | -13.723 *** | 4.013 | -4.162 *** | 1.145 |
| Second Friday | -9.781 *** | 3.389 | -3.006 *** | 0.885 |
| Settlement Wednesday | 1.914 | 1.771 | 2.443 *** | 0.750 |
| OTB(-1) | 0.146 *** | 0.040 | 0.015 * | 0.010 |
| OTB(-2) | 0.142 *** | 0.043 | 0.037 *** | 0.009 |
| OTB(-3) | 0.219 *** | 0.053 | 0.023 *** | 0.009 |
| OTS(-1) | -0.904 | 0.563 | -3.435 * | 2.091 |
| OTS(-2) | -0.561 | 0.440 | -1.144 ** | 0.570 |
| OTS(-3) | -2.068 *** | 0.710 | -3.851 *** | 1.233 |
| PB(-1) | -0.109 | 0.117 | -0.028 | 0.042 |
| PB(-2) | 0.030 | 0.108 | 0.008 | 0.027 |
| PB(-3) | -0.082 | 0.103 | -0.049 * | 0.036 |
| PS(-1) | 0.135 | 0.255 | 0.335 *** | 0.092 |
| PS(-2) | 0.398 | 0.244 | 0.586 *** | 0.085 |
| PS(-3) | -0.212 | 0.270 | -0.008 | 0.081 |
| SURPRISE | 19.230 * | 10.688 | 1.898 | 5.521 |
| SURPRISE(-1) | -18.989 | 15.698 | -1.063 | 5.790 |
| SURPRISE(-2) | -42.791 | 34.819 | -0.654 | 8.035 |
| SURPRISE(-3) | -42.019 * | 23.901 | -12.702 | 12.600 |
| NEED(-1)*Day1 | -1.260 | 11.661 | 7.145 *** | 1.624 |
| NEED(-1)*Day2 | 24.978 | 15.438 | 11.032 *** | 4.263 |
| NEED(-1)*Day3 | 22.241 ** | 10.014 | 26.737 *** | 5.221 |
| NEED(-1)*Day4 | -11.121 | 10.667 | 2.232 | 3.159 |
| NEED(-1)*Day5 | 42.804 *** | 12.427 | 17.637 *** | 4.020 |
| NEED(-1)*Day6 | -14.788 | 19.240 | -28.351 *** | 6.376 |
| NEED(-1)*Day7 | 20.849 | 19.487 | 10.726 *** | 3.831 |
| NEED(-1)*Day8 | 8.298 | 14.080 | 11.127 *** | 4.187 |
| NEED(-1)*Day9 | 0.540 | 7.724 | 2.384 | 2.055 |
| NEED(-1)*Day10 | 10.502 * | 5.679 | 0.945 | 2.493 |
| EXPECT*Day1 | 93.199 | 68.418 | 33.119 *** | 7.166 |
| EXPECT*Day2 | 27.979 | 38.112 | 0.933 | 10.344 |
| EXPECT*Day3 | -12.791 | 38.854 | -3.585 | 7.963 |
| EXPECT*Day4 | -43.710 | 31.953 | -13.921 ** | 7.604 |
| EXPECT*Day5 | -77.641 ** | 33.233 | -18.657 ** | 9.407 |
| EXPECT*Day6 | 7.185 | 30.834 | 9.525 * | 6.507 |
| EXPECT*Day7 | -51.883 | 37.008 | -8.995 | 17.458 |
| EXPECT*Day8 | -9.313 | 21.748 | 6.363 | 6.739 |
| EXPECT*Day9 | -98.467 *** | 29.943 | -18.863 ** | 11.178 |
| EXPECT*Day10 | -15.402 | 16.353 | 0.078 | 7.091 |
| SCALE | 983.513 *** | 132.711 |  |  |
| Smallest root | 1.353 |  | 3.001 |  |

Note: ***, ** and * denote significance at the 1, 5 and $10 \%$ level, respectively. "Smallest root" denotes the smallest modulus of the roots of the $\rho_{\max }(B)$ lag polynomial.

Table 4
Bias: Tobit model with $\mathrm{p}=1$

|  |  | Observations |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | T=100 | T=300 | T=600 | $\mathrm{T}=1000$ | $\mathrm{T}=2000$ |
| $V_{1}$ | 1 | 0.0157 | 0.0046 | 0.0030 | 0.0017 | 0.0011 |
| $V_{2}$ | 1 | 0.0076 | 0.0027 | 0.0007 | 0.0007 | 0.0005 |
| $\rho$ | 0 | -0.0098 | -0.0031 | -0.0013 | -0.0009 | -0.0007 |
| $\sigma^{2}$ | 1 | -0.0307 | -0.0099 | -0.0049 | -0.0035 | -0.0016 |
| $V_{1}$ | 1 | 0.0285 | 0.0074 | 0.0042 | 0.0023 | 0.0012 |
| $V_{2}$ | 1 | 0.0099 | 0.0037 | 0.0019 | 0.0013 | 0.0007 |
| $\rho$ | 0.3 | -0.0105 | -0.0033 | -0.0019 | -0.0011 | -0.0007 |
| $\sigma^{2}$ | 1 | -0.0294 | -0.0105 | -0.0038 | -0.0035 | -0.0014 |
| $\gamma_{1}$ | 1 | 0.0428 | 0.0111 | 0.0072 | 0.0051 | 0.0019 |
| $\gamma_{2}$ | 1 | 0.0131 | 0.0040 | 0.0025 | 0.0009 | 0.0003 |
| $\rho$ | 0.6 | -0.0086 | -0.0025 | -0.0015 | -0.0008 | -0.0005 |
| $\sigma^{2}$ | 1 | -0.0302 | -0.0095 | -0.0058 | -0.0030 | -0.0015 |
| $V_{1}$ | 1 | 0.1091 | 0.0324 | 0.0156 | 0.0091 | 0.0101 |
| $V_{2}$ | 1 | 0.0077 | 0.0018 | 0.0018 | 0.0006 | 0.0004 |
| $\rho$ | 0.9 | -0.0036 | -0.0010 | -0.0006 | -0.0003 | -0.0005 |
| $\sigma^{2}$ | 1 | -0.0295 | -0.0095 | -0.0050 | -0.0031 | -0.0017 |

Model: $\mathrm{y}_{\mathrm{t}}=\max \left(0, \gamma_{1}+\gamma_{2}{ }^{*} \mathrm{x}_{\mathrm{t}}+\rho{ }^{*} \mathrm{y}_{\mathrm{t}-1}+\varepsilon_{\mathrm{t}}\right)$

$$
x_{t}=0.5+0.5 x_{t-1}+v_{t}
$$

Table 5
Bias: Tobit model with $\mathrm{p}=2$

|  |  |  | Observations |  |  |  |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
|  |  | $\mathrm{T}=100$ | $\mathrm{~T}=300$ | $\mathrm{~T}=600$ | $\mathrm{~T}=1000$ | $\mathrm{~T}=2000$ |
|  |  |  |  |  |  |  |
| $\gamma_{1}$ | 1 | 0.0394 | 0.0125 | 0.0071 | 0.0044 | 0.0018 |
| $\gamma_{2}$ | 1 | 0.0070 | 0.0027 | 0.0004 | 0.0007 | 0.0000 |
| $\rho_{1}$ | 0.2 | -0.0062 | -0.0018 | -0.0008 | -0.0006 | -0.0002 |
| $\rho_{2}$ | 0.1 | -0.0099 | -0.0034 | -0.0018 | -0.0012 | -0.0004 |
| $\sigma^{2}$ | 1 | -0.0408 | -0.0129 | -0.0067 | -0.0044 | -0.0020 |
|  |  |  |  |  |  |  |
| $\gamma_{1}$ | 1 | 0.0613 | 0.0169 | 0.0095 | 0.0053 | 0.0028 |
| $\gamma_{2}$ | 1 | 0.0058 | 0.0022 | 0.0012 | 0.0009 | 0.0004 |
| $\rho_{1}$ | 0.5 | -0.0072 | -0.0023 | -0.0011 | -0.0009 | -0.0004 |
| $\rho_{2}$ | 0.1 | -0.0062 | -0.0015 | -0.0011 | -0.0004 | -0.0003 |
| $\sigma^{2}$ | 1 | -0.0394 | -0.0138 | -0.0054 | -0.0044 | -0.0018 |
|  |  |  |  |  |  |  |
| $\gamma_{1}$ | 1 | 0.2140 | 0.0647 | 0.0330 | 0.0230 | 0.0087 |
| $\gamma_{2}$ | 1 | 0.0076 | 0.0022 | 0.0016 | 0.0002 | 0.0005 |
| $\rho_{1}$ | 0.8 | -0.0091 | -0.0024 | -0.0016 | -0.0005 | -0.0005 |
| $\rho_{2}$ | 0.1 | -0.0020 | -0.0010 | -0.0001 | -0.0006 | 0.0001 |
| $\sigma^{2}$ | 1 | -0.0411 | -0.0131 | -0.0076 | -0.0041 | -0.0020 |
|  |  |  |  |  |  |  |
| $\rho_{1}$ | 1 | 0.0161 | 0.0054 | 0.0023 | 0.0025 | 0.0014 |
| $\sigma_{2}$ | -0.3 | -0.0001 | -0.0001 | 0.0002 | 0.0000 | 0.0000 |
| $\gamma_{1}$ | 1 | -0.0391 | -0.0130 | -0.0056 | -0.0042 | -0.0023 |
| $\gamma_{2}$ | 1 | 0.0058 | -0.0001 | 0.0015 | -0.0001 | 0.0001 |
| $\rho_{2}$ | 1 | 0.0332 | 0.0128 | 0.0055 | 0.0030 | 0.0011 |
| $\rho_{1}$ | 0 | -0.0068 | -0.0006 | -0.0011 | -0.0002 | -0.0007 |
| $\rho_{2}$ | -0.3 | -0.0069 | -0.0026 | -0.0016 | -0.0010 | -0.0002 |
| $\sigma^{2}$ | 1 | -0.0380 | -0.0121 | -0.0067 | -0.0043 | -0.0022 |
|  |  | 0.0074 | 0.0026 | 0.0014 | 0.0003 | 0.0005 |
| $\gamma_{1}$ | 1 | 0.0167 | 0.0066 | 0.0017 | 0.0024 | 0.0012 |
| $\gamma_{2}$ | 1 | 0.0061 | 0.0022 | 0.0006 | 0.0007 | 0.0003 |
| $\rho_{1}$ | 0.3 | -0.0064 | -0.0029 | -0.0012 | -0.0009 | -0.0004 |
| $\rho_{2}$ | -0.3 | -0.0055 | -0.0016 | -0.0004 | -0.0007 | -0.0002 |
| $\sigma^{2}$ | 1 | -0.0369 | -0.0139 | -0.0061 | -0.0040 | -0.0017 |
|  |  |  |  |  |  |  |
| $\gamma_{1}$ | 1 | 0.0220 | 0.0060 | 0.0035 | 0.0016 | 0.0010 |
| $\gamma_{2}$ | 1 | 0.0075 | 0.0016 | 0.0008 | 0.0010 | 0.0003 |
|  | 0.6 | -0.0072 | -0.0020 | -0.0008 | -0.0007 | -0.0005 |
|  | -0.0030 | -0.0007 | -0.0007 | -0.0004 | 0.0001 |  |
|  |  | -0.0124 | -0.0059 | -0.0044 | -0.0020 |  |
|  |  |  |  |  |  |  |

Model: $y_{t}=\max \left(0, \gamma_{1}+\gamma_{2}{ }^{*} x_{t}+\rho_{1}{ }^{*} y_{t-1}+\rho_{2}{ }^{*} y_{\mathrm{t}-2}+\varepsilon_{\mathrm{t}}\right)$

$$
x_{t}=0.5+0.5 x_{t-1}+v_{t}
$$

Table 6 Bias: CLAD with $\mathrm{p}=1$

|  |  | Observations |  |  |  |  |  |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: | :---: |
|  |  | $\mathrm{T}=100$ | $\mathrm{~T}=300$ | $\mathrm{~T}=600$ | $\mathrm{~T}=1000$ | $\mathrm{~T}=2000$ |  |
|  |  |  |  |  |  |  |  |
| $\gamma_{1}$ | 1 | -0.0027 | 0.0008 | 0.0004 | 0.0005 | 0.0008 |  |
| $\gamma_{2}$ | 1 | 0.0112 | 0.0043 | 0.0026 | 0.0009 | 0.0001 |  |
| $\rho$ | 0 | -0.0071 | -0.0032 | -0.0021 | -0.0011 | -0.0007 |  |
|  |  |  |  |  |  |  |  |
| $\gamma_{1}$ | 1 | 0.0130 | 0.0072 | 0.0024 | 0.0020 | 0.0020 |  |
| $\gamma_{2}$ | 1 | -0.0012 | 0.0034 | 0.0015 | 0.0003 | 0.0005 |  |
| $\rho$ | 0.3 | -0.0090 | -0.0039 | -0.0012 | -0.0010 | -0.0007 |  |
|  |  |  |  |  |  |  |  |
| $\gamma_{1}$ | 1 | 0.0414 | 0.0168 | 0.0063 | 0.0045 | 0.0030 |  |
| $\gamma_{2}$ | 1 | 0.0077 | 0.0033 | 0.0005 | 0.0005 | 0.0006 |  |
| $\rho$ | 0.6 | -0.0100 | -0.0040 | -0.0014 | -0.0011 | -0.0008 |  |
|  |  |  |  |  |  |  |  |
| $\gamma_{1}$ | 1 | 0.1939 | 0.0592 | 0.0248 | 0.0189 | 0.0086 |  |
| $\gamma_{2}$ | 1 | 0.0063 | 0.0029 | 0.0002 | 0.0007 | 0.0004 |  |
| $\rho$ | 0.9 | -0.0100 | -0.0031 | -0.0013 | -0.0010 | -0.0004 |  |

Model: $\mathrm{y}_{\mathrm{t}}=\max \left(0, \gamma_{1}+\gamma_{2}{ }^{*} \mathrm{x}_{\mathrm{t}}+\rho{ }^{*} \mathrm{y}_{\mathrm{t}-1}+\varepsilon_{\mathrm{t}}\right)$
$\mathrm{x}_{\mathrm{t}}=0.5+0.5 \mathrm{x}_{\mathrm{t}-1}+\mathrm{v}_{\mathrm{t}}$

Table 7
Bias: CLAD with $\mathrm{p}=2$

|  |  | Observations |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | $\mathrm{T}=100$ | $\mathrm{T}=300$ | $\mathrm{T}=600$ | $\mathrm{T}=1000$ | $\mathrm{T}=2000$ |
| $V_{1}$ | 1 | 0.0309 | 0.0101 | 0.0036 | 0.0031 | 0.0020 |
| $V_{2}$ | 1 | 0.0078 | 0.0037 | 0.0009 | 0.0008 | -0.0003 |
| $\rho_{1}$ | 0.2 | -0.0050 | -0.0018 | -0.0005 | -0.0007 | -0.0002 |
| $\rho_{2}$ | 0.1 | -0.0091 | -0.0031 | -0.0014 | -0.0007 | -0.0004 |
| $V_{1}$ | 1 | 0.0588 | 0.0170 | 0.0101 | 0.0045 | 0.0029 |
| $V_{2}$ | 1 | 0.0075 | 0.0027 | 0.0002 | 0.0006 | -0.0002 |
| $\rho_{1}$ | 0.5 | -0.0091 | -0.0015 | -0.0011 | -0.0007 | -0.0003 |
| $\rho_{2}$ | 0.1 | -0.0041 | -0.0024 | -0.0010 | -0.0003 | -0.0003 |
| $V_{1}$ | 1 | 0.2206 | 0.0657 | 0.0323 | 0.0181 | 0.0094 |
| $\gamma_{2}$ | 1 | 0.0065 | 0.0029 | 0.0003 | 0.0007 | -0.0002 |
| $\rho_{1}$ | 0.8 | -0.0107 | -0.0025 | -0.0015 | -0.0010 | -0.0003 |
| $\rho_{2}$ | 0.1 | -0.0007 | -0.0009 | -0.0001 | 0.0000 | -0.0002 |
| $V_{1}$ | 1 | -0.0035 | 0.0008 | 0.0008 | 0.0009 | 0.0002 |
| $V_{2}$ | 1 | 0.0195 | 0.0055 | 0.0017 | 0.0013 | 0.0008 |
| $\rho_{1}$ | 0 | -0.0063 | -0.0015 | -0.0009 | -0.0008 | -0.0001 |
| $\rho_{2}$ | -0.3 | -0.0115 | -0.0039 | -0.0018 | -0.0011 | -0.0005 |
| $V_{1}$ | 1 | 0.0066 | 0.0029 | 0.0022 | 0.0010 | 0.0008 |
| $V_{2}$ | 1 | 0.0126 | 0.0044 | 0.0011 | 0.0009 | -0.0002 |
| $\rho_{1}$ | 0.3 | -0.0078 | -0.0012 | -0.0008 | -0.0007 | -0.0001 |
| $\rho_{2}$ | -0.3 | -0.0041 | -0.0027 | -0.0014 | -0.0004 | -0.0003 |
| $V_{1}$ | 1 | 0.0142 | 0.0054 | 0.0027 | 0.0015 | 0.0013 |
| $V_{2}$ | 1 | 0.0093 | 0.0036 | 0.0016 | 0.0008 | -0.0002 |
| $\rho_{1}$ | 0.6 | -0.0078 | -0.0014 | -0.0010 | -0.0008 | -0.0003 |
| $\rho_{2}$ | -0.3 | -0.0012 | -0.0019 | -0.0007 | -0.0001 | -0.0001 |
| $\gamma_{1}$ | 1 | 0.0321 | 0.0117 | 0.0039 | 0.0030 | 0.0018 |
| $\gamma_{2}$ | 1 | 0.0072 | 0.0031 | 0.0010 | 0.0007 | -0.0003 |
| $\rho_{1}$ | 0.9 | -0.0077 | -0.0023 | -0.0017 | -0.0009 | -0.0003 |
| $\rho_{2}$ | -0.3 | -0.0003 | -0.0006 | 0.0005 | 0.0001 | 0.0000 |

Model: $\mathrm{y}_{\mathrm{t}}=\max \left(0, \gamma_{1}+\gamma_{2}{ }^{*} \mathrm{x}_{\mathrm{t}}+\rho_{1}{ }^{*} \mathrm{y}_{\mathrm{t}-1}+\rho_{2}{ }^{*} \mathrm{y}_{\mathrm{t}-2}+\varepsilon_{\mathrm{t}}\right)$

$$
x_{t}=0.5+0.5 x_{t-1}+v_{t}
$$

## Appendix

Define $\hat{y}_{t}^{m}=0$ for $m \leq 0$ and $\hat{y}_{t}^{m}=\max \left(0, \eta_{t}+\sum_{i=1}^{p} \rho_{i} \hat{y}_{t-i}^{m-i}\right)$. 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 an almost surely finite upper bound for $y_{t}$ and $\hat{y}_{t}^{m}$ :

Lemma 1 If the lag polynomial $\left(1-\max \left(0, \rho_{1}\right) B-\ldots-\max \left(0, \rho_{p}\right) B^{p}\right)$ has all its roots outside the unit circle and $\sup _{t \in \mathbb{Z}} E \max \left(0, \eta_{t}\right)<\infty$, then for an almost surely finite random variable $f_{t}=f\left(\eta_{t}, \eta_{t-1}, \ldots\right)=\sum_{j=0}^{\infty} L_{1}^{j} \max \left(0, \eta_{t-j}\right)$, and $L_{1}^{j}$ that are such that $L_{1}^{j} \leq c_{1} \exp \left(-c_{2} j\right)$ for positive constants $c_{1}$ and $c_{2}$,

$$
\hat{y}_{t}^{m} \leq f_{t}
$$

## Proof of Lemma 1:

Note that, by successive substitution of the definition of $\hat{y}_{t}^{m}$ for the $\hat{y}_{t}^{m}$ that has the largest value for $t$,

$$
\begin{aligned}
& \hat{y}_{t}^{m} \leq \max \left(0, \eta_{t}\right)+\sum_{i=1}^{p} \max \left(0, \rho_{i}\right) \hat{y}_{t-i}^{m-i} \\
& =\max \left(0, \eta_{t}\right)+\sum_{i=1}^{p} L_{i}^{1} \hat{y}_{t-i}^{m-i} \\
& \leq \max \left(0, \eta_{t}\right)+\sum_{i=2}^{p} \max \left(0, \rho_{i}\right) \hat{y}_{t-i}^{m-i}+\max \left(0, \rho_{1}\right)\left(\max \left(0, \eta_{t-1}\right)+\sum_{i=1}^{p} \max \left(0, \rho_{i}\right) \hat{y}_{t-i-1}^{m-i-1}\right) \\
& =\max \left(0, \eta_{t}\right)+L_{1}^{1} \max \left(0, \eta_{t-1}\right)+\sum_{i=1}^{p} L_{i}^{2} \hat{y}_{t-i-1}^{m-i-1} \\
& \leq \max \left(0, \eta_{t}\right)+L_{1}^{1} \max \left(0, \eta_{t-1}\right)+L_{1}^{2} \max \left(0, \eta_{t-2}\right)+\sum_{i=1}^{p} L_{i}^{3} \hat{y}_{t-i-2}^{m-i-2} \\
& \leq \sum_{j=0}^{\infty} L_{1}^{j} \max \left(0, \eta_{t-j}\right) .
\end{aligned}
$$

The $L_{i}^{j}$ satisfy, for $j \geq 2$,

$$
\begin{aligned}
& L_{1}^{j}=L_{2}^{j-1}+\max \left(0, \rho_{1}\right) L_{1}^{j-1}, \\
& L_{2}^{j}=L_{3}^{j-1}+\max \left(0, \rho_{2}\right) L_{1}^{j-1}, \\
& \vdots \\
& L_{p-1}^{j}=L_{p}^{j-1}+\max \left(0, \rho_{p-1}\right) L_{1}^{j-1}, \\
& L_{p}^{j}=\max \left(0, \rho_{p}\right) L_{1}^{j-1}
\end{aligned}
$$

From these equations it follows that we can write, for the backward operator $B$ that is such that $B\left(L_{i}^{j}\right)=L_{i}^{j-1}$,

$$
\left(1-\sum_{j=1}^{p} \max \left(0, \rho_{j}\right) B^{j}\right) L_{1}^{j}=0
$$

From the fact that the above lag polynomial has all its roots outside the unit circle by assumption, it follows that $L_{1}^{j} \leq c_{1} \exp \left(-c_{2} j\right)$ for positive constants $c_{1}$ and $c_{2}$. Also, if $\sup _{t \in \mathbb{Z}} E \max \left(0, \eta_{t}\right)<\infty$, then $\sum_{j=0}^{\infty} L_{1}^{j} \max \left(0, \eta_{t-j}\right)$ is an a.s. finite random variable.

We will first proceed by deriving a moment bound for $y_{t}$. The following theorem provides such a result:

Lemma 2 If $\eta_{t}$ is strictly stationary, $\rho_{\max }(B)$ has all its roots outside the unit circle, and $\left\|\max \left(0, \eta_{t}\right)\right\|_{r}<\infty$ for some $r \geq 1$, then $\sup _{t \in \mathbb{Z}}\left\|f_{t}\right\|_{r}<\infty$.

## Proof of Lemma 2:

The result Lemma 2 follows by noting that, by Lemma 1,

$$
\left\|f_{t}\right\|_{r} \leq \sum_{j=0}^{\infty} L_{1}^{j}\left\|\max \left(0, \eta_{t-j}\right)\right\|_{r}<\infty
$$

We will also need an exponential inequality:

Lemma 3 If $x_{t}$ is $L_{1}$-near epoch dependent on $v_{t}$, where $v_{t}$ is $\alpha$-mixing and $\alpha(m)+\nu(m) \leq$ $C_{1} \exp \left(-C_{2} m\right)$ for positive constants $C_{1}$ and $C_{2}$, and $\left|x_{t}\right| \leq 1$, then for all $\delta>0$,

$$
P\left(\left|m^{-1} \sum_{t=1}^{m} x_{t}\right|>\delta\right) \leq c_{1} \exp \left(-c_{2} \delta^{2} m^{1 / 3}\right)
$$

for positive constants $c_{1}$ and $c_{2}$ possibly depending on $\delta$.

## Proof of Lemma 3:

Observe that for all $k>0$,

$$
x_{t}=\sum_{j=-k}^{k-1}\left(E\left(x_{t} \mid v_{t-j}, \ldots\right)-E\left(x_{t} \mid v_{t-j-1,}, \ldots\right)\right)+E\left(x_{t} \mid v_{t-k}, \ldots\right)+\left(x_{t}-E\left(x_{t} \mid v_{t+k}, \ldots\right)\right),
$$

and therefore for all $k>0$,

$$
\begin{aligned}
& P\left(\left|m^{-1} \sum_{t=1}^{m} x_{t}\right|>\delta\right) \\
& \leq \delta^{-1}\left\|E\left(x_{t} \mid v_{t-k-1}, \ldots\right)\right\|_{1}+\delta^{-1}\left\|x_{t}-E\left(x_{t} \mid v_{t+k}, \ldots\right)\right\|_{1} \\
& +\sum_{j=-k}^{k} P\left(m^{-1} \sum_{t=1}^{m}\left(E\left(x_{t} \mid v_{t-j}, \ldots\right)-E\left(x_{t} \mid v_{t-j-1,}, \ldots\right)\right)>\delta /(2 k+1)\right)
\end{aligned}
$$

By the $L_{1}$-near epoch dependence condition, boundedness and the $L_{1}$-mixingale property of $x_{t}$ (see Andrews (1988) ),

$$
\delta^{-1}\left\|E\left(x_{t} \mid v_{t-k-1}, \ldots\right)\right\|_{1}+\delta^{-1}\left\|x_{t}-E\left(x_{t} \mid v_{t+k}, \ldots\right)\right\|_{1} \leq \delta^{-1} C_{1} \exp \left(-C_{2} k\right)
$$

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

$$
\begin{aligned}
& P\left(m^{-1} \sum_{t=1}^{m}\left(E\left(x_{t} \mid v_{t-j}, \ldots\right)-E\left(x_{t} \mid v_{t-j-1, \ldots}, .\right)\right)>\delta /(2 k)\right) \\
& \leq 2 \exp \left(-\delta^{2} m / 8 k^{2}\right)
\end{aligned}
$$

By choosing $k=\left[m^{1 / 3}\right]$ and collecting terms, the result now follows.

The following lemma is needed for the stationarity proof of Theorem 1 . For $\zeta>0$, let

$$
\begin{aligned}
& H^{\zeta}(x)=-\zeta^{-1} x I(-\zeta \leq x \leq 0)+I(x \leq-\zeta) . \\
& I_{t l}=\prod_{j=0}^{p-1} I\left(\eta_{t-l-j} \leq-\sum_{i=1}^{p} \rho_{j} f_{t-l-j-i}\right)
\end{aligned}
$$

and

$$
I_{t l}^{\zeta}=\prod_{j=0}^{p-1} H^{\zeta}\left(\eta_{t-l-j}+\sum_{i=1}^{p} \rho_{j} f_{t-l-j-i}\right) .
$$

Lemma 4 Assume that $\eta_{t}$ is strictly stationary and strong mixing and satisfies $\left\|\max \left(0, \eta_{t}\right)\right\|_{2}<$ $\infty$. Then for all $t \in \mathbb{Z}$ and $\delta>0$, as $m \rightarrow \infty$,

$$
(m-p)^{-1} \sum_{l=1}^{m-p}\left(I_{t l}^{\zeta} \log (\delta)+\log (1+\delta)\left(1-I_{t l}^{\zeta}\right)\right) \xrightarrow{p} E\left(I_{t l}^{\zeta} \log (\delta)+\log (1+\delta)\left(1-I_{t l}^{\zeta}\right)\right) .
$$

## Proof of Lemma 4:

Note that we can write

$$
\begin{aligned}
& (m-p)^{-1} \sum_{l=1}^{m-p}\left(I_{t l}^{\zeta} \log (\delta)+\log (1+\delta)\left(1-I_{t l}^{\zeta}\right)\right) \\
& =(m-p)^{-1} \sum_{l=1}^{m-p}\left(I_{t, m-p+1-l}^{\zeta} \log (\delta)+\log (1+\delta)\left(1-I_{t, m-p+1-l}^{\zeta}\right)\right)
\end{aligned}
$$

Note that

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

and for all $t$ and $j$,

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

$$
\begin{aligned}
& =\eta_{t-(m-p+1-l)-j}+\sum_{i=1}^{p} \sum_{k=0}^{\infty} \rho_{i} L_{1}^{k} \max \left(0, \eta_{t-(m-p+1-l)-j-i-k}\right) \\
& =\eta_{t-(m-p+1-l)-j}+\sum_{k=0}^{\infty} \max \left(0, \eta_{t-(m-p+1-l)-j-i-k}\right) \sum_{i=1}^{p} \rho_{i} L_{1}^{k-i} I(i \leq k)=w_{t-(m-p+1-l)-j}
\end{aligned}
$$

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 $M \geq 1$,

$$
\begin{aligned}
& \left\|w_{t-(m-p+1-l)-j}-E\left(w_{t-(m-p+1-l)-j} \mid \eta_{t-(m-p+1-l)-j-M}, \ldots, \eta_{t-(m-p+1-l)-j}\right)\right\|_{2} \\
& \leq\left\|\max \left(0, \eta_{t}\right)\right\|_{2} \sum_{k=M+1}^{\infty} \sum_{i=1}^{p} \rho_{i} L_{1}^{k-i} I(i \leq k),
\end{aligned}
$$

and the last expression converges to 0 as $M \rightarrow \infty$ at exponential rate because $L_{1}^{k}$ converges to zero at an exponential rate. Therefore, because $H^{\zeta}(\cdot)$ is Lipschitz-continuous,

$$
H^{\zeta}\left(\eta_{t-(m-p+1-l)-j}+\sum_{i=1}^{p} \rho_{i} f_{t-(m-p+1-l)-j-i}\right)
$$

is also $L_{2}$-near epoch dependent on $\eta_{t}$ with an exponentially decreasing $\nu(\cdot)$ sequence, and so is

$$
\prod_{j=0}^{p-1} H^{\zeta}\left(\eta_{t-(m-p+1-l)-j}+\sum_{i=1}^{p} \rho_{i} f_{t-(m-p+1-l)-j-i}\right)
$$

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).

Lemma 5 Under the assumptions of Theorem 1, for all $\zeta>0$,

$$
E \prod_{j=1}^{p} I\left(\eta_{t-j}+\sum_{i=1}^{p} \rho_{i} f_{t-i-j} \leq-\zeta\right)>0
$$

## Proof of Lemma 5:

Note that, under our assumptions, for some $c>0$ and $a \in(0,1)$,

$$
\eta_{t}+\sum_{i=1}^{p} \rho_{i} f_{t-i} \leq u_{t}+c \sum_{j=0}^{\infty}|a|^{j}\left|u_{t-j}\right| .
$$

Noting that

$$
\begin{aligned}
& E \prod_{j=1}^{p} I\left(\eta_{t-j}+\sum_{i=1}^{p} \rho_{i} f_{t-i-j} \leq-\zeta\right)>0 \\
& =E\left(P\left(\eta_{t-1}+\sum_{i=1}^{p} \rho_{i} f_{t-1-j} \leq-\zeta\right) \prod_{j=2}^{p} I\left(\eta_{t-j}+\sum_{i=1}^{p} \rho_{i} f_{t-i-j} \leq-\zeta\right)\right) \\
& \geq E\left(P\left(u_{t-1}+c \sum_{j=1}^{\infty}|a|^{j}\left|u_{t-1-j}\right| \leq-\zeta\right) \prod_{j=2}^{p} I\left(\eta_{t-j}+\sum_{i=1}^{p} \rho_{i} f_{t-i-j} \leq-\zeta\right)\right) \\
& =E\left(F\left(-c \sum_{j=0}^{\infty}|a|^{j}\left|u_{t-j}\right|-\zeta\right) \prod_{j=2}^{p} I\left(\eta_{t-j}+\sum_{i=1}^{p} \rho_{i} f_{t-i-j} \leq-\zeta\right)\right)
\end{aligned}
$$

where $F($.$) is as defined in Equation (5), and observing that for random variables X$ such that $X \geq 0$, we can have $E X=0$ only if $P(X=0)=1$, it can be seen that it suffices to show that

$$
F\left(-c \sum_{j=0}^{\infty}|a|^{j}\left|u_{t-j}\right|-\zeta\right) \prod_{j=2}^{p} I\left(\eta_{t-j}+\sum_{i=1}^{p} \rho_{i} f_{t-i-j} \leq-\zeta\right)
$$

exceeds zero with probability 1 . Because of positivity of $F($.$) , this means it suffices that$

$$
E \prod_{j=2}^{p} I\left(\eta_{t-j}+\sum_{i=1}^{p} \rho_{i} f_{t-i-j} \leq-\zeta\right)>0
$$

(that is, the product is over $j=2$ instead of $j=1$ now.) By repeating this reasoning $p$ times and because of the finiteness of random variables such as $\sum_{j=0}^{\infty}|a|^{j}\left|w_{t-j}\right|$, it now follows that for all $\zeta>0$,

$$
E \prod_{j=1}^{p} I\left(\eta_{t-j}+\sum_{i=1}^{p} \rho_{i} f_{t-i-j} \leq-\zeta\right)>0
$$

Lemma 6 For some almost surely finite random variable $y_{t}$ such that $\left(y_{t}, \eta_{t}\right)$ is strictly stationary,

$$
\hat{y}_{t}^{m} \xrightarrow{\text { as }} y_{t} \quad \text { as } \quad m \rightarrow \infty .
$$

## Proof of Lemma 6:

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 $\max _{k \geq m}\left|\hat{y}_{t}^{k}-\hat{y}_{t}^{m}\right|$ converges to zero in probability as $m \rightarrow \infty$. Now, note that for all $m \geq k$,

$$
\hat{y}_{t}^{k}=\hat{y}_{t}^{m}=0 \quad \text { if } \quad \eta_{t} \leq-\sum_{i=1}^{p} \rho_{i} y_{t-i} \quad \text { and } \quad \eta_{t} \leq-\sum_{i=1}^{p} \rho_{i} \hat{y}_{t-i}^{k-i}
$$

so certainly,

$$
\hat{y}_{t}^{k}=\hat{y}_{t}^{m}=0 \quad \text { if } \quad \eta_{t} \leq-\sum_{i=1}^{p} \rho_{i} f_{t-i},
$$

and therefore $\max _{k \geq m}\left|\hat{y}_{t}^{k}-\hat{y}_{t}^{m}\right|=0$ 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 $c>0$,

$$
\begin{aligned}
& P\left[\max _{k \geq m}\left|\hat{y}_{t}^{k}-\hat{y}_{t}^{m}\right|>0\right] \\
& \leq P\left[\text { there are no } p \text { consecutive "small" } \eta_{t}\right] \\
& \leq E \prod_{l=1}^{m-p}\left(1-I\left(\text { there are } p \text { consecutive "small" } \eta_{t} \text { starting at } t-l\right)\right)
\end{aligned}
$$

$$
\begin{align*}
& \leq E \prod_{l=1}^{m-p}\left(1-\prod_{j=0}^{p-1} I\left(\eta_{t-l-j} \leq-\sum_{i=1}^{p} \rho_{j} f_{t-l-j-i}\right)\right) \\
& =E \exp \left[(m-p)(m-p)^{-1} \sum_{l=1}^{m-p} \log \left(1-\prod_{j=0}^{p-1} I\left(\eta_{t-l-j} \leq-\sum_{i=1}^{p} \rho_{i} f_{t-l-j-i}\right)\right)\right] \\
& \leq \exp (-(m-p) c))+P\left[(m-p)^{-1} \sum_{l=1}^{m-p} \log \left(1-\prod_{j=0}^{p-1} I\left(\eta_{t-l-j} \leq-\sum_{i=1}^{p} \rho_{i} f_{t-l-j-i}\right)\right)>-c\right] \\
& \leq \exp (-(m-p) c))+P\left[(m-p)^{-1} \sum_{l=1}^{m-p}\left(I_{t l} \log (\delta)+\log (1+\delta)\left(1-I_{t l}\right)\right)>-c\right] \\
& \leq \exp (-(m-p) c))+P\left[(m-p)^{-1} \sum_{l=1}^{m-p}\left(I_{t l}^{\zeta} \log (\delta)+\log (1+\delta)\left(1-I_{t l}^{\zeta}\right)\right)>-c\right] \tag{19}
\end{align*}
$$

where

$$
I_{t l}=\prod_{j=0}^{p-1} I\left(\eta_{t-l-j} \leq-\sum_{i=1}^{p} \rho_{j} f_{t-l-j-i}\right)
$$

and

$$
I_{t l}^{\zeta}=\prod_{j=0}^{p-1} H^{\zeta}\left(\eta_{t-l-j}+\sum_{i=1}^{p} \rho_{j} f_{t-l-j-i}\right)
$$

for

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

Note that $I_{t l} \geq I_{t l}^{\zeta}$ because $I(x \leq 0) \geq H^{\zeta}(x)$. Both terms in Equation (19) now converge to zero as $m \rightarrow \infty$ for a suitable choice of $\zeta, c$ and $\delta$ if

$$
\begin{align*}
& E(m-p)^{-1} \sum_{l=1}^{m-p}\left(I_{t l}^{\zeta} \log (\delta)+\log (1+\delta)\left(1-I_{t l}^{\zeta}\right)\right) \\
& =E\left(I_{t l}^{\zeta} \log (\delta)+\log (1+\delta)\left(1-I_{t l}^{\zeta}\right)\right)<0 \tag{20}
\end{align*}
$$

and

$$
(m-p)^{-1} \sum_{l=1}^{m-p}\left(I_{t l}^{\zeta} \log (\delta)+\log (1+\delta)\left(1-I_{t l}^{\zeta}\right)\right)
$$

satisfies a weak law of large numbers as $m \rightarrow \infty$. This weak law of large numbers is proven in Lemma 4. Now if $E I_{t l}^{\zeta}>0$, we can pick $\delta>0$ small enough to satisfy the requirement of Equation (20). Now,

$$
\begin{aligned}
& E I_{t l}^{\zeta}=E \prod_{j=1}^{p} H^{\zeta}\left(\eta_{t-l-j}-\sum_{i=1}^{p} \rho_{j} f_{t-l-j-i}\right) \\
& \geq E \prod_{j=1}^{p} I\left(\eta_{t-l-j}+\sum_{i=1}^{p} \rho_{j} f_{t-l-j-i} \leq-\zeta\right)
\end{aligned}
$$

and the last term is positive by Lemma 5 .
Since $\hat{y}_{t}^{m}=f_{m}\left(\eta_{t}, \ldots, \eta_{t-m}\right)$ is strictly stationary because it depends on a finite numbers of $\eta_{t}, \lim _{m \rightarrow \infty}\left(\hat{y}_{t}^{m}, \eta_{t}\right)=\left(y_{t}, \eta_{t}\right)$ is also strictly stationary.

## Proof of Theorem 1:

Noting that $y_{t}$ as constructed in Lemma 6 is a solution to the dynamic censored regression model, part (i) of Theorem 1 follows. Also, by the reasoning of Lemma 6 it follows that any $z_{t}=f\left(\eta_{t}, \eta_{t-1}, \ldots\right)$ that is a solution to the model also satisfies $\max _{k \geq m}\left|\hat{y}_{t}^{k}-z_{t}\right|=0$, implying that $z_{t}=y_{t}$ a.s., thereby showing part (ii) of Theorem 1. To show part (iii), note that, by strict stationarity of $\left(y_{t}, \eta_{t}\right)$ and by noting that the conditional expectation is the best $L_{2}$-approximation,

$$
\begin{aligned}
& \sup _{t \in \mathbb{Z}} E\left|y_{t}-E\left(y_{t} \mid \eta_{t-m}, \eta_{t-m+1}, \ldots, \eta_{t}\right)\right|^{2} \\
& =E\left|y_{t}-E\left(y_{t} \mid \eta_{t-m}, \eta_{t-m+1}, \ldots, \eta_{t}\right)\right|^{2} \\
& \leq E\left|y_{t}-\hat{y}_{t}^{m}\right|^{2}
\end{aligned}
$$

and because $\left|y_{t}\right|+\left|\hat{y}_{t}^{m}\right| \leq 2 f_{t}$, it now follows by the dominated convergence theorem that $y_{t}$ is $L_{2}$-near epoch dependent because $E\left|f_{t}\right|^{2}<\infty$ by assumption and by Lemma 1 .

In order to obtain the explicit bound for $\nu(M)$ of the last part of Theorem 1, note that

$$
\nu(m)=E\left(y_{t}-E\left(y_{t} \mid \eta_{t-m}, \ldots, \eta_{t}\right)\right)^{2} \leq E\left(y_{t}-\hat{y}_{t}^{m}\right)^{2} I\left(\left|\hat{y}_{t}^{m}-y_{t}\right|>0\right)
$$

$$
\leq\left(E\left|2 f_{t}\right|^{2 p}\right)^{1 / p}\left(P\left(\left|\hat{y}_{t}^{m}-y_{t}\right|>0\right)\right)^{1 / q}
$$

for $p \geq 1$ and $q \geq 1$ such that $p^{-1}+q^{-1}=1$. Now by choosing $p$ small enough, $E\left|f_{t}\right|^{2 p}<\infty$ by assumption and by Lemma 1. Therefore, it suffices to show that $P\left(\left|\hat{y}_{t}^{m}-y_{t}\right|>0\right)$ decays as $c_{1} \exp \left(-c_{2} m^{1 / 3}\right)$ with $m$. By the earlier reasoning,

$$
\begin{aligned}
& P\left[\max _{k \geq m}\left|\hat{y}_{t}^{k}-y_{t}\right|>0\right] \\
& \leq \exp (-(m-p) c))+P\left[(m-p)^{-1} \sum_{l=1}^{m-p}\left(I_{t l}^{\zeta} \log (\delta)+\log (1+\delta)\left(1-I_{t l}^{\zeta}\right)\right)>-c\right]
\end{aligned}
$$

and the last probability will decay as $c_{1} \exp \left(-c_{2} m^{1 / 3}\right)$ 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 7 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 $E \sup _{b \in B}\left|q\left(z_{t}, b\right)\right|<\infty$,

$$
\sup _{b \in B}\left|T^{-1} \sum_{t=1}^{T}\left(q\left(z_{t}, b\right)-E q\left(z_{t}, b\right)\right)\right| \xrightarrow{p} 0 .
$$

## Proof of Lemma 7:

See Lemma A. 2 of Pötscher and Prucha (1986).

## Proof of Theorem 2:

For convergence in probability, we check the conditions of Theorem 2.7 of Newey and McFadden (1994). As is well-known (see e.g. Greene (2000)), the objective function is concave
in $(b / s, r / s, 1 / s)$. The assumptions of Theorem 1 are met, and the conclusion of the theorem implies ergodicity; see White (2001, theorem 3.34). This and the moment conditions of Assumption 2 imply the pointwise convergence of the objective function. Assumption 2.4 and the condition that $\inf \Sigma>0$ imply that the limit objective function is uniquely maximized. Therefore, all conditions of Theorem 2.7 of Newey and McFadden (1994) are satisfied, and consistency follows.

Lemma 8 Under the conditions of Theorem 3,

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

## Proof of Lemma 8:

Note that by assumption, $E\left(\left.\left(\partial l_{t}(b) / \partial b\right)\right|_{b=\beta} \mid y_{t-1}, \ldots, x_{t}\right)=0$ so that $\left.E\left(\partial l_{t}(b) / \partial b\right)\right|_{b=\beta}=0$, implying that $\left.\left(\partial l_{t}(b) / \partial b\right)\right|_{b=\beta}$ is a martingale difference sequence. In particular, by noting that $\left(y_{t}, x_{t}\right)$ 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 3:

We prove Theorem 3 by checking the conditions of Newey and McFadden (1994, Theorem 3.1). Consistency was shown in Theorem 2. Condition (i) was assumed. Condition (ii), twice differentiability of the loglikelihood, follows from the Tobit specification. Condition (iii) was shown in Lemma 8. Note that stationarity and the strong mixing base imply ergodicity. Condition (iv) follows from the result of Lemma 7, the Tobit specification, and the assumption of finite second moments for $\left|x_{t}\right|$ and $\varepsilon_{t}$. Condition (v) is assumed.

## Proof of Theorem 4:

Under Assumption 4, it follows from the discussion in Powell (1984, p. 318) that $S_{T}(b)$ is uniquely minimized at $\beta=\left(\rho^{\prime}, \gamma^{\prime}\right)^{\prime}$. From Assumption 4 it follows that $E \sup _{b \in B}\left|s\left(y_{t-1}, \ldots, y_{t-p}, x_{t}, \varepsilon_{t}, b\right)\right|<$
$\infty$, and therefore the uniform law of large numbers of Lemma 7 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 $\left(w_{t}, \varepsilon_{t}\right)$, let

$$
\lambda(b)=E \psi\left(w_{t}, \varepsilon_{t}, b\right)
$$

Lemma 9 Assume that $\left(w_{t}, \varepsilon_{t}\right)$ is strictly stationary and that $\left|\hat{\beta}_{T}-\beta\right|=o_{p}(1)$. In addition assume that

$$
T^{-1 / 2} \sum_{t=1}^{T} \psi\left(w_{t}, \varepsilon_{t}, \hat{\beta}_{T}\right)=o_{p}(1),
$$

and assume that

$$
T^{-1 / 2} \sum_{t=1}^{T}\left(\psi\left(w_{t}, \varepsilon_{t}, b\right)-E \psi\left(w_{t}, \varepsilon_{t}, b\right)\right)
$$

is stochastically equicontinuous on $B$. Then

$$
\left.T^{1 / 2} \lambda(b)\right|_{b=\hat{\beta}_{T}}=-T^{-1 / 2} \sum_{t=1}^{T} \psi\left(w_{t}, \varepsilon_{t}, \beta\right)+o_{p}(1)
$$

## Proof of Lemma 9:

This follows by writing

$$
\begin{aligned}
& o_{p}(1)=T^{-1 / 2} \sum_{t=1}^{T} \psi\left(w_{t}, \varepsilon_{t}, \hat{\beta}_{T}\right) \\
& =T^{-1 / 2} \sum_{t=1}^{T}\left(\psi\left(w_{t}, \varepsilon_{t}, \hat{\beta}_{T}\right)-E \psi\left(w_{t}, \varepsilon_{t}, \hat{\beta}_{T}\right)-\psi\left(w_{t}, \varepsilon_{t}, \beta\right)+E \psi\left(w_{t}, \varepsilon_{t}, \beta\right)\right) \\
& \quad+T^{-1 / 2} \sum_{t=1}^{T} \psi\left(w_{t}, \varepsilon_{t}, \beta\right)+T^{1 / 2} \lambda\left(\hat{\beta}_{T}\right),
\end{aligned}
$$

and noting that by the stochastic equicontinuity assumption, the first term in the last expression is $o_{p}(1)$ if $\left|\hat{\beta}_{T}-\beta\right|=o_{p}(1)$.

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

Lemma 10 Assume $u_{1 t} \in$ and $u_{2 t} \in \mathbb{R}^{q}$, and assume that $\left(u_{1 t}, u_{2 t}^{\prime}\right)^{\prime}$ is strictly stationary and $L_{r}$-near epoch dependent, $r \geq 2$, on $\eta_{t}$ with a $\nu(\cdot)$ sequence satisfying $\nu(m) \leq c_{1} \exp \left(-c_{2} m^{1 / 3}\right)$ for $c_{1}, c_{2}>0$. Then if $\left|u_{1 t}\right| \leq 1$ and $\left\|u_{2 t}\right\|_{r+\delta}<\infty$ for $\delta<r /(r-1)$, $u_{1 t} u_{2 t}$ is $L_{r}$-near epoch dependent on $\eta_{t}$ with a $\nu(\cdot)$ sequence that satisfies $\nu(m) \leq c_{1} \exp \left(-c_{2} m^{1 / 3}\right)$ for constants $c_{1}, c_{2}>0$ (that are not necessarily identical to the earlier $c_{1}$ and $c_{2}$ ).

## Proof of Lemma 10:

This follows by noting that

$$
\begin{aligned}
& \left\|u_{1 t} u_{2 t}-E\left(u_{1 t} u_{2 t} \mid \eta_{t-M}, \ldots, \eta_{t}\right)\right\|_{r} \\
& \quad \leq\left\|u_{1 t} u_{2 t}-E\left(u_{1 t} \mid \eta_{t-M}, \ldots, \eta_{t}\right) E\left(u_{2 t} \mid \eta_{t-M}, \ldots, \eta_{t}\right)\right\|_{r} \\
& \leq\left\|u_{1 t}\left(u_{2 t}-E\left(u_{2 t} \mid \eta_{t-M}, \ldots, \eta_{t}\right)\right)\right\|_{r}+\left\|E\left(u_{2 t} \mid \eta_{t}, \ldots, \eta_{t-m}\right)\left(u_{1 t}-E\left(u_{1 t} \mid \eta_{t-M}, \ldots, \eta_{t}\right)\right)\right\|_{r} \\
& \quad \leq\left\|u_{2 t}-E\left(u_{2 t} \mid \eta_{t-M}, \ldots, \eta_{t}\right)\right\|_{r} \\
& \quad \quad+\left\|E\left(u_{2 t} \mid \eta_{t-M}, \ldots, \eta_{t}\right)\right\|_{r+\delta}\left\|u_{1 t}-E\left(u_{1 t} \mid \eta_{t-M}, \ldots, \eta_{t}\right)\right\|_{r(1+r / \delta)} \\
& \quad \leq\left\|u_{2 t}-E\left(u_{2 t} \mid \eta_{t-M}, \ldots, \eta_{t}\right)\right\|_{r} \\
& \quad \quad+\left\|E\left(u_{2 t} \mid \eta_{t-M}, \ldots, \eta_{t}\right)\right\|_{r+\delta}\left\|u_{1 t}-E\left(u_{1 t} \mid \eta_{t-M}, \ldots, \eta_{t}\right)\right\|_{r}^{r /(1+r / \delta)}
\end{aligned}
$$

and all terms in the last expression decay with $m$ as specified.

We also need the following result.
Lemma 11 Under Assumption 5 and the conditions of Theorem 1, for all $\eta>0, I\left(b^{\prime} w_{t}>\right.$ 0), $I\left(b^{\prime} w_{t}>-\eta\left|w_{t}\right|\right)$, $I\left(b^{\prime} w_{t} \leq \eta\left|w_{t}\right|\right)$, and $\left.I\left(\varepsilon_{t}+(\beta-b)^{\prime} w_{t}>0\right)\right), I\left(\varepsilon_{t}+(\beta-b)^{\prime} w_{t}>-\eta\left|w_{t}\right|\right)$, $I\left(\varepsilon_{t}+(\beta-b)^{\prime} w_{t} \leq \eta\left|w_{t}\right|\right)$ are $L_{r}$-near epoch dependent on $\eta_{t}$ with $a \nu(\cdot)$ sequence satisfying $\nu(m) \leq c_{1} \exp \left(-c_{2} m^{1 / 3}\right)$ for $c_{1}, c_{2}>0$.

## Proof of Lemma 11:

We will show this for one case; the other cases are analogous. Note that $w_{t}$ is near epoch dependent on $\eta_{t}=\gamma^{\prime} x_{t}+\varepsilon_{t}$ with $\nu(\cdot)$ sequence satisfying $\nu(m) \leq c_{1} \exp \left(-c_{2} m^{1 / 3}\right)$ for $c_{1}, c_{2}>$ 0 by Theorem 1. 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

$$
\begin{aligned}
& \left\|I\left(b^{\prime} w_{t}+\eta\left|w_{t}\right|>0\right)-E\left(I\left(b^{\prime} w_{t}+\eta\left|w_{t}\right|>0\right) \mid \eta_{t-M}, \ldots, \eta_{t}\right)\right\|_{r} \\
& \leq 2\left\|I\left(b^{\prime} w_{t}+\eta\left|w_{t}\right|>0\right)-T_{\delta}\left(b^{\prime} w_{t}+\eta\left|w_{t}\right|\right)\right\|_{r} \\
& \quad+\left\|T_{\delta}\left(b^{\prime} w_{t}+\eta\left|w_{t}\right|>0\right)-E\left(T_{\delta}\left(b^{\prime} w_{t}+\eta\left|w_{t}\right|>0\right) \mid \eta_{t-M}, \ldots, \eta_{t}\right)\right\|_{r} \\
& \leq 2\left\|I\left(\left|b^{\prime} w_{t}+\eta\right| w_{t} \mid \| \leq \delta\right)\right\|_{q}+C_{2} \delta^{-1}\left\|b^{\prime} w_{t}-E\left(b^{\prime} w_{t} \mid \eta_{t-M}, \ldots, \eta_{t}\right)\right\|_{r} \\
& \quad+C_{2} \delta^{-1}\left\|w_{t}-E\left(w_{t} \mid \eta_{t-M}, \ldots, \eta_{t}\right)\right\|_{r} \\
& \leq C_{1} \delta+C_{3} \nu(M) \delta^{-1}
\end{aligned}
$$

because $b^{\prime} w_{t}+\eta\left|w_{t}\right|$ has a uniformly bounded density, which follows from the assumption that the density of $\varepsilon_{t} \mid w_{t}$ is uniformly bounded. Therefore by setting $\delta=\nu(M)^{1 / 2}$, it follows that $I\left(b^{\prime} w_{t}+\eta\left|w_{t}\right|>0\right)$ is near epoch dependent on $\eta_{t}$ with a $\nu(\cdot)$ sequence satisfying $\nu(m) \leq c_{1} \exp \left(-c_{2} m^{1 / 3}\right)$ for $c_{1}, c_{2}>0$ as well.

Lemma 12 Under Assumption 5, $T^{-1 / 2} \sum_{t=1}^{T}\left(\psi\left(w_{t}, \varepsilon_{t}, b\right)-E \psi\left(w_{t}, \varepsilon_{t}, b\right)\right)$ is stochastically equicontinuous on $B$.

## Proof of Lemma 12:

Note that for $\psi\left(w_{t}, \varepsilon_{t}, b\right)$ we have that

$$
\mu_{t}(b, \delta)=\sup _{\tilde{b}:|b-\tilde{b}|<\delta}\left|\psi\left(w_{t}, \varepsilon_{t}, b\right)-\psi\left(w_{t}, \varepsilon_{t}, \tilde{b}\right)\right|
$$

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 $O\left(\delta^{p+q}\right)$ balls with center $b_{j}$ and radius $\eta$ and we can define the bracketing functions as

$$
f_{j}^{L}\left(w_{t}\right)=I\left(b_{j}^{\prime} w_{t}>0\right)\left(1 / 2-I\left(\varepsilon_{t}+\left(\beta-b_{j}\right)^{\prime} w_{t}>0\right)\right) w_{t}
$$

$$
-I\left(b_{j}^{\prime} w_{t}>-\delta\left|w_{t}\right|\right)\left|1 / 2-I\left(\varepsilon_{t}+\left(\beta-b_{j}\right)^{\prime} w_{t}>-\delta\left|w_{t}\right|\right)\right|\left|w_{t}\right|
$$

and

$$
\begin{aligned}
& f_{j}^{U}\left(w_{t}\right)=I\left(b_{j}^{\prime} w_{t}>0\right)\left(1 / 2-I\left(\varepsilon_{t}+\left(\beta-b_{j}\right)^{\prime} w_{t}>0\right)\right) w_{t} \\
& \quad+I\left(b_{j}^{\prime} w_{t}>-\delta\left|w_{t}\right|\right)\left|1 / 2-I\left(\varepsilon_{t}+\left(\beta-b_{j}\right)^{\prime} w_{t}>-\delta\left|w_{t}\right|\right) \| w_{t}\right| .
\end{aligned}
$$

By the result of Lemma 11, the bracketing functions $f_{j}^{L}(\cdot)$ and $f_{j}^{U}(\cdot)$ as well as the $\psi\left(w_{t}, \varepsilon_{t}, \beta\right)$ 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 $\nu(M)+\alpha(M)^{1 / r-1 /(2 r)}$ 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 $2 r$ respectively, and we note that the bracketing functions $f_{j}^{L}(\cdot)$ and $f_{j}^{U}(\cdot)$ as well as $\psi\left(w_{t}, \varepsilon_{t}, \beta\right)$ are also $L_{r}$-mixingales with mixingale numbers $\nu(M)^{1 / 2}+\alpha(M)^{1 /(2 r)-1 /(4 r)}$ and mixingale numbers $c_{t}=\left\|f_{j}^{L}\left(w_{t}\right)\right\|_{2 r}^{1 / 2}, c_{t}=\left\|f_{j}^{L}\left(w_{t}\right)\right\|_{2 r}^{1 / 2}$, or $c_{t}=\left\|f_{j}^{L}\left(w_{t}\right)\right\|_{2 r}^{1 / 2}$ respectively. The condition

$$
\sum_{M=0}^{\infty}\left(\nu(M)^{1 / 2}+\alpha(M)^{1 /(4 r)}\right)<\infty
$$

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 5:

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 9, and we note that under the conditions of Theorem 5, the stochastic equicontinuity condition of Lemma 9 follows from the result of Lemma 12. 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),

$$
T^{-1 / 2} \sum_{t=1}^{T} \psi\left(w_{t}, \varepsilon_{t}, \tilde{\beta}_{T}\right)=o_{p}(1) .
$$

By Lemma 9 and Lemma 12, for $\lambda(b)=E \psi\left(w_{t}, \varepsilon_{t}, b\right)$,

$$
\left.T^{1 / 2} \lambda(b)\right|_{b=\tilde{\beta}_{T}}=-T^{-1 / 2} \sum_{t=1}^{T} \psi\left(w_{t}, \varepsilon_{t}, \beta\right)+o_{p}(1)
$$

It now follows from Lemma 10 and Lemma 11 that $\psi\left(w_{t}, \varepsilon_{t}, b\right)$ 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

$$
T^{-1 / 2} \sum_{t=1}^{T} \psi\left(w_{t}, \varepsilon_{t}, \beta\right) \xrightarrow{d} N(0, \Omega) .
$$

Since $\lambda(\beta)=0$ by assumption, for some mean value $\beta_{T}^{*}$,

$$
\left.T^{1 / 2} \lambda(b)\right|_{b=\tilde{\beta}_{T}}=o_{p}(1)+\left.(\partial / \partial b) \lambda(b)\right|_{b=\beta_{T}^{*}} T^{1 / 2}\left(\tilde{\beta}_{T}-\beta\right),
$$

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),

$$
\left.(\partial / \partial b) \lambda(b)\right|_{b=\beta_{T}^{*}}=o_{p}(1)+N .
$$

Therefore, it now follows that

$$
T^{1 / 2}\left(\tilde{\beta}_{T}-\beta\right) \xrightarrow{d} N\left(0, N^{-1} \Omega N^{-1}\right),
$$

as asserted by the theorem.
Demiralp and Jordà (2002) Tobit regression for Open Market Purchases

|  | OB |  |  |  |  |  | TB |  |  |  |  |  | PB |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 1984-1994 |  | 1994-1998 |  | 1998-2000 |  | 1984-1994 |  | 1994-1998 |  | 1998-2000 |  | 1984-1994 |  | 1994-1998 |  | 1998-2000 |  |
|  | Estimate | Std. Err. | Estimate | Std. Err. | Estimate | Std. Err. | Estimate | Std. Err. | Estimate | Std. Err. | Estimate | Std. Err. | Estimate | Std. Err. | Estimate | Std. Err. | Estimate | Std. Err. |
| Day1 | -24.80 | 2.55 | -10.75 | 1.65 | 4.14 | 1.56 | -39.09 | 5.05 | -9.54 | 6.05 | 23.12 | 10.62 | 3.27 | 1.43 | -11.34 | 5.64 | -1.52 | 4.67 |
| Day2 | -24.51 | 2.70 | -10.92 | 1.91 | 3.91 | 1.89 | -132.62 | 20.09 | -64.39 | 10.38 | -42.82 | 15.29 | -9.38 | 1.77 | -39.24 | 8.06 | -16.05 | 6.81 |
| Day 3 | -38.84 | 4.89 | -9.82 | 1.65 | 1.13 | 1.96 | -40.99 | 5.35 | -17.50 | 6.54 | 4.90 | 14.36 | -14.56 | 1.74 | -36.29 | 7.56 | -13.60 | 7.24 |
| Day4 | -32.17 | 3.60 | -12.73 | 1.77 | 1.69 | 1.97 | -67.94 | 7.59 | -34.28 | 7.59 | -31.79 | 15.16 | -12.54 | 1.79 | -36.23 | 7.21 | -6.79 | 6.34 |
| Day5 | -28.10 | 3.08 | -7.11 | 1.43 | 0.69 | 1.78 | -89.18 | 9.25 | -52.72 | 8.10 | -36.56 | 14.11 | -11.43 | 1.57 | -43.89 | 7.30 | -24.81 | 8.26 |
| Day6 | -26.92 | 2.77 | -9.47 | 1.52 | 1.50 | 1.62 | -34.28 | 4.57 | -11.69 | 5.41 | -12.24 | 11.52 | -8.61 | 1.44 | -34.34 | 6.01 | -10.49 | 5.14 |
| Day7 | -24.86 | 2.74 | -7.04 | 1.50 | -1.16 | 2.01 | -81.43 | 8.16 | -54.50 | 8.42 | -25.28 | 13.28 | -13.03 | 1.55 | -48.30 | 8.43 | -15.46 | 5.76 |
| Day8 | -27.56 | 2.91 | -8.57 | 1.50 | 2.86 | 1.73 | -33.64 | 4.69 | -12.12 | 5.39 | -7.12 | 11.82 | -13.12 | 1.59 | -44.00 | 7.48 | -14.08 | 5.55 |
| Day9 | -29.02 | 2.97 | -10.59 | 1.59 | -1.09 | 1.93 | -40.71 | 5.07 | -40.23 | 6.83 | -52.13 | 15.12 | -13.49 | 1.65 | -37.91 | 6.85 | -20.75 | 6.68 |
| Day10 | -8.15 | 1.63 | 2.15 | 1.08 | 5.56 | 1.62 | -111.84 | 13.23 | -89.90 | 13.74 | -101.96 | 23.87 | -13.73 | 1.58 | -44.24 | 7.08 | -19.65 | 6.02 |
| $\mathrm{OB}(-1)$ | 0.63 | 0.14 | 0.51 | 0.09 | 0.06 | 0.08 | 1.45 | 0.34 | 0.98 | 0.41 | 0.79 | 0.58 | 0.09 | 0.12 | -0.57 | 0.51 | -0.25 | 0.31 |
| $\mathrm{OB}(-2)$ | 0.27 | 0.15 | 0.30 | 0.09 | -0.02 | 0.07 | -1.03 | 0.47 | 0.28 | 0.46 | -0.73 | 0.56 | -0.14 | 0.13 | 0.08 | 0.43 | -0.47 | 0.31 |
| $\mathrm{OB}(-3)$ | 0.02 | 0.15 | 0.24 | 0.09 | 0.09 | 0.07 | 0.24 | 0.43 | 0.30 | 0.48 | -1.26 | 0.71 | -0.04 | 0.12 | -0.95 | 0.56 | -0.28 | 0.31 |
| TB(-1) | 0.14 | 0.04 | -0.01 | 0.02 | 0.00 | 0.02 | 0.27 | 0.15 | 0.13 | 0.10 | -0.02 | 0.12 | 0.06 | 0.04 | -0.19 | 0.13 | -0.07 | 0.07 |
| TB( -2 ) | -0.01 | 0.06 | 0.04 | 0.02 | 0.02 | 0.02 | 0.59 | 0.10 | 0.08 | 0.09 | -0.04 | 0.11 | 0.07 | 0.04 | 0.18 | 0.09 | -0.08 | 0.06 |
| TB(-3) | 0.08 | 0.04 | 0.10 | 0.02 | 0.01 | 0.02 | 0.40 | 0.10 | 0.15 | 0.10 | 0.11 | 0.12 | 0.02 | 0.04 | 0.02 | 0.09 | -0.05 | 0.06 |
| PB(-1) | -0.04 | 0.07 | -0.09 | 0.06 | -0.15 | 0.08 | 0.57 | 0.21 | 0.08 | 0.26 | 0.05 | 0.55 | 0.14 | 0.04 | 0.22 | 0.17 | -0.07 | 0.28 |
| $\mathrm{PB}(-2)$ | -0.15 | 0.16 | 0.03 | 0.05 | -0.13 | 0.08 | -0.17 | 0.16 | -0.20 | 0.22 | 0.10 | 0.53 | 0.11 | 0.05 | 0.33 | 0.16 | 0.74 | 0.21 |
| PB(-3) | -0.03 | 0.09 | 0.02 | 0.05 | -0.17 | 0.09 | 0.20 | 0.15 | -0.12 | 0.25 | 0.83 | 0.56 | 0.10 | 0.05 | 0.28 | 0.17 | 0.42 | 0.23 |
| OS(-1) | -1.94 | 1.18 | -2.61 | 1.22 | -0.72 | 0.42 | -3.26 | 1.62 | -2.47 | 1.37 | -0.81 | 2.37 | -0.67 | 0.38 | -115.60 | 1.81E+07 | -1.85 | 1.34 |
| OS(-2) | -0.99 | 0.79 | 0.22 | 0.37 | -0.69 | 0.44 | 0.31 | 1.32 | -1.06 | 1.45 | 3.16 | 2.72 | -0.95 | 0.45 | -83.07 | $1.52 \mathrm{E}+07$ | -1.39 | 1.34 |
| OS(-3) | -0.27 | 0.44 | 0.00 | 0.26 | -0.56 | 0.40 | -1.37 | 1.54 | 0.77 | 1.41 | -10.61 | 6.01 | -1.10 | 0.43 | -4.44 | 5.12 | -1.64 | 1.82 |
| TS(-1) | -25.91 | 9.10E+06 | -22.26 | $1.55 \mathrm{E}+06$ | -0.18 | 0.44 | -62.74 | $1.50 \mathrm{E}+07$ | -91.14 | $3.84 \mathrm{E}+06$ | -8.30 | 5.44E+05 | -1.51 | 0.55 | -64.65 | $1.11 \mathrm{E}+07$ | -1.48 | $2.22 \mathrm{E}+06$ |
| TS(-2) | -1.73 | 1.24 | -0.35 | 0.34 | 0.51 | 0.40 | -73.66 | $1.67 \mathrm{E}+07$ | -110.88 | $4.20 \mathrm{E}+06$ | -41.52 | $1.27 \mathrm{E}+06$ | -1.16 | 0.38 | -71.41 | 1.22E+07 | -15.32 | 5.23E+06 |
| TS(-3) | -0.81 | 0.71 | -0.14 | 0.27 | -6.60 | $2.75 \mathrm{E}+05$ | -4.12 | 2.82 | -83.72 | $4.15 \mathrm{E}+06$ | -36.26 | $1.10 \mathrm{E}+06$ | -0.78 | 0.34 | -23.91 | 8.70E+06 | -13.54 | $4.66 \mathrm{E}+06$ |
| PS(-1) | 0.10 | 0.20 | 0.18 | 0.32 | 0.39 | 0.12 | 0.25 | 0.63 | -3.28 | 3.41 | -7.11 | 6.93 | -0.04 | 0.15 | -2.70 | 2.95 | -1.13 | 1.37 |
| PS(-2) | -0.01 | 0.32 | 0.00 | 0.33 | -0.09 | 0.12 | -10.06 | 6.06 | -0.22 | 1.44 | 0.36 | 0.75 | 0.07 | 0.15 | -3.77 | 3.59 | 0.28 | 0.32 |
| PS(-3) | 0.33 | 0.18 | -1.62 | 1.09 | 0.02 | 0.12 | -5.93 | 3.31 | 1.25 | 1.19 | 0.08 | 1.14 | -0.35 | 0.26 | -219.99 | 3.44E+07 | 0.60 | 0.34 |
| NEED(-1)*Day1 | 3.25 | 1.67 | 1.71 | 3.08 | 10.46 | 6.50 | -1.97 | 4.97 | 24.31 | 10.56 | -27.57 | 39.86 | -2.40 | 1.72 | -16.67 | 11.68 | 1.31 | 18.21 |
| NEED(-1)*Day2 | 8.44 | 2.00 | 23.61 | 5.74 | -33.18 | 8.10 | 13.26 | 8.63 | 5.51 | 40.79 | 94.03 | 60.87 | -8.99 | 5.18 | 14.51 | 25.40 | 16.10 | 42.54 |
| NEED(-1)*Day3 | 24.33 | 13.39 | -6.22 | 8.76 | -1.22 | 6.66 | 14.95 | 19.72 | -6.78 | 31.05 | 130.04 | 69.32 | -3.29 | 7.17 | -15.44 | 43.78 | 16.67 | 39.30 |
| NEED(-1)*Day4 | 23.09 | 9.07 | 8.10 | 4.80 | 5.50 | 9.32 | 25.77 | 23.59 | -105.39 | 51.50 | -116.42 | 73.14 | -6.04 | 6.35 | 3.98 | 23.19 | -45.55 | 30.47 |
| NEED(-1)*Day5 | 12.48 | 9.81 | 13.41 | 6.17 | 10.26 | 9.79 | 99.72 | 25.48 | 46.28 | 31.34 | -1.89 | 90.02 | -2.19 | 6.14 | -21.80 | 32.03 | -43.33 | 58.01 |
| NEED (-1)*Day6 | 22.90 | 8.43 | 10.44 | 10.12 | 7.35 | 9.70 | 3.83 | 18.34 | -62.10 | 38.79 | -49.78 | 68.10 | -1.52 | 6.58 | 20.08 | 33.39 | -27.42 | 30.14 |
| NEED(-1)*Day7 | 22.42 | 6.97 | 5.40 | 5.58 | 41.15 | 10.57 | 15.11 | 20.91 | -23.26 | 56.32 | 2.09 | 66.43 | -7.37 | 5.97 | 10.24 | 27.15 | -0.01 | 28.51 |
| NEED( -1 * ${ }^{\text {Day }} 8$ | 15.56 | 8.14 | 9.84 | 7.41 | 16.32 | 8.71 | -12.84 | 16.46 | 21.91 | 26.66 | -47.13 | 54.76 | -1.63 | 5.77 | 27.33 | 30.29 | -0.68 | 28.54 |
| NEED(-1)*Day9 | 21.90 | 5.19 | 5.08 | 3.14 | 14.42 | 8.05 | -1.14 | 12.84 | -8.36 | 18.57 | -71.68 | 73.45 | 2.25 | 4.20 | -8.48 | 19.65 | -19.43 | 32.72 |
| NEED(-1)*Day10 | 2.69 | 1.43 | 0.67 | 3.44 | 7.61 | 7.01 | 0.18 | 9.15 | 25.60 | 29.72 | -85.46 | 69.69 | 0.95 | 1.49 | -4.68 | 21.46 | 46.04 | 42.05 |
| EXPECT(-1)*Day 1 | -17.30 | 39.44 | -45.50 | 21.78 | 19.90 | 21.60 | 52.45 | 83.47 | 94.46 | 45.08 | 576.88 | 133.90 | 4.25 | 25.76 | -40.08 | 42.05 | -21.21 | 63.86 |
| EXPECT(-1)* ${ }^{\text {²y }}$ 2 | -25.50 | 42.35 | 4.40 | 13.20 | 51.33 | 24.81 | 469.49 | 204.27 | 23.87 | 72.48 | 387.22 | 167.29 | -21.30 | 29.67 | 37.91 | 46.26 | -728.74 | 986.54 |
| EXPECT(-1)*Day3 | 23.02 | 73.19 | -22.05 | 16.35 | -46.00 | 29.07 | -109.39 | 96.81 | 20.29 | 48.58 | 123.81 | 153.54 | -5.97 | 31.00 | -11.67 | 53.56 | -28.06 | 108.68 |
| EXPECT(-1)*Day 4 | -37.16 | 51.02 | 8.30 | 12.94 | 12.76 | 23.10 | 25.95 | 105.05 | 10.25 | 56.67 | 73.49 | 169.67 | -29.98 | 30.96 | 76.48 | 41.56 | 12.94 | 73.23 |
| EXPECT(-1)*Day5 | 0.19 | 42.93 | -40.36 | 18.38 | -2.68 | 22.06 | -152.51 | 124.22 | 27.25 | 63.97 | 102.51 | 160.34 | -23.80 | 29.57 | 52.71 | 50.98 | -1826.98 | 978.21 |
| EXPECT(-1)*Day 6 | -25.11 | 38.99 | -35.96 | 20.25 | -33.79 | 24.68 | 62.84 | 80.30 | 23.22 | 47.74 | -63.79 | 153.04 | -10.98 | 27.53 | 76.30 | 40.65 | -59.91 | 89.17 |
| EXPECT(-1)*Day 7 | -11.37 | 41.68 | -7.77 | 13.29 | -16.04 | 27.81 | -283.58 | 163.53 | -237.32 | 191.75 | 168.48 | 146.45 | -12.79 | 29.48 | 58.97 | 54.58 | -455.14 | 551.67 |
| EXPECT(-1)*Day8 | 44.23 | 42.16 | -22.21 | 16.31 | -150.37 | 84.91 | -37.89 | 87.40 | 9.64 | 48.39 | -49.01 | 153.97 | -25.07 | 30.39 | -16.16 | 66.40 | -360.66 | 473.80 |
| EXPECT(-1)*Day9 | -17.52 | 40.94 | -10.01 | 15.47 | -33.88 | 30.55 | 14.99 | 89.55 | -56.43 | 70.62 | 160.67 | 156.52 | 2.63 | 29.82 | -7.58 | 54.76 | -238.04 | 422.71 |
| EXPECT(-1)*Day10 | 1.94 | 29.18 | -4.98 | 10.04 | 0.47 | 21.25 | 159.85 | 217.37 | 46.46 | 99.31 | 177.40 | 268.13 | 9.74 | 30.14 | 70.77 | 45.16 | -119.49 | 180.78 |
| SURPRISE | -28.58 | 11.43 | 15.24 | 8.75 | 1.94 | 11.51 | -53.66 | 29.53 | 21.15 | 68.18 | -2.86 | 116.51 | -14.40 | 9.26 | 30.69 | 33.98 | 7.73 | 47.59 |
| SURPRISE(-1) | -2.80 | 14.70 | -21.42 | 11.66 | -8.03 | 11.88 | 72.78 | 38.42 | 27.17 | 50.74 | -126.67 | 131.85 | 15.20 | 10.83 | 66.78 | 39.02 | -19.06 | 54.39 |
| SURPRISE(-2) | -42.29 | 14.56 | -18.47 | 17.64 | -7.51 | 13.39 | 5.28 | 35.11 | -56.71 | 59.00 | -93.72 | 79.92 | 2.13 | 11.02 | 42.45 | 38.61 | 4.11 | 38.59 |
| SURPRISE(-3) | -19.09 | 12.96 | -24.12 | 18.61 | -8.50 | 13.44 | -12.63 | 40.45 | -29.72 | 87.75 | -167.15 | 114.22 | -6.00 | 10.37 | 49.76 | 38.08 | 0.20 | 53.99 |
| SCALE | 16.92 | 0.90 | 9.04 | 0.42 | 9.02 | 0.40 | 43.59 | 2.00 | 40.56 | 1.90 | 55.96 | 3.41 | 17.17 | 0.45 | 29.45 | 2.09 | 20.16 | 1.83 |





[^0]:    *Department of Economics, Ohio State University, 429 Arps Hall, Columbus, OH 43210, email dejong@econ.ohio-state.edu
    ${ }^{\dagger}$ 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.

[^1]:    ${ }^{1}$ The reported standard errors for the Tobit estimates are the quasi-maximum likelihood standard errors.

[^2]:    ${ }^{2}$ Even though the federal funds target has only been announced since the February 3-4 FOMC meeting, Demiralp and Jordà (2004) 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) suggests that even in earlier periods, market participants were able to read signals of a target change in the Fed's behavior.

[^3]:    ${ }^{3}$ Our estimate for $\Omega$ was calculated with an HAC estimator using the Bartlett kernel, and the bandwidth was selected using the data-driven method suggested by Andrews (1991). The density function $f(0 \mid w t)$ was estimated using a higher-order Gaussian kernel with the order and bandwidth selected according to Hansen (2003, 2004).

[^4]:    ${ }^{4}$ Due to the stationarity issues discussed in the previous section, we do not use the same subsamples as Demiralp and Jordà (2002).
    ${ }^{5}$ Similar conclusions are drawn if we use the futures federal funds rate (Kuttner, 2001) to decompose anticipated and unanticipated changes in policy.

[^5]:    Note: Standard errors reported in parenthesis. "Smallest root" denotes the smallest modulus of the roots of the $\rho_{\max }(B)$ lag polynomial.

