Dynamic censored regression and the Open Market Desk reaction function

Robert de Jong*   Ana María Herrera†

April 20, 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

*Department of Economics, Ohio State University, 429 Arps Hall, Columbus, OH 43210, email dejong@econ.ohio-state.edu

†Department of Economics, Wayne State University, 656 W Kirby, 2095 FAB, Detroit, MI 48202, email amherrera@wayne.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.

of a firm or individual leads to a corner response for a significant proportion of time. In addition, right-censoring may arise 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

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

where x_t denotes the regressor, ε_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 ε_t is itself allowed to be a linear process (i.e., an MA(∞) 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 $\rho_{max}(z) = 1 - \sum_{i=1}^{p} \max(0, \rho_i) z^i$ has its roots outside the unit circle. Therefore, in terms of the conditions on $\rho_{max}(z)$ and the dependence allowed for ε_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 AR(p) model can be derived. Note that intuitively, negative values for ρ_j can never be problematic when considering the stationarity properties of y_t, since they "pull y_t back to zero". This intuition is formalized by the fact that only $\max(0, \rho_j)$ shows up in our stationarity requirement.

An alternative formulation for the dynamic censored regression model could be

$$y_t = y_t^* I(y_t^* > 0) \quad \text{where} \quad \rho(B) y_t^* = \gamma' x_t + \varepsilon_t, \tag{2}$$

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 determinants of 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 presents our weak dependence results for (y_t, x_t) in the censored regression model. In Section 3, we show the asymptotic validity of the dynamic Tobit procedure. Powell's (1984) 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.

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 η_t is defined as follows:

Definition 1 *Random variables y_t are called L_r-*near epoch dependent *on η_t if*

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

The base process η_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 (α-) and uniform (ϕ-) 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 η_t.

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

$$y_t = \max(0, \sum_{i=1}^{p} \rho_i y_{t-i} + \eta_t). \tag{4}$$

Later, we will set $\eta_t = \gamma' x_t + \varepsilon_t$ in order to obtain weak dependence results for the general dynamic censored regression model that contains regressors.

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

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

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

Theorem 1 *If the linear process* η_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* $f_u(.)$, $E|u_t|^r < \infty$ *for some* $r \geq 2$,

$$\int_{-\infty}^{\infty} |f_u(y+a) - f_u(y)| dy \leq M|a|$$

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

$$P(u_t \leq x) \geq F(x) > 0 \tag{5}$$

for some function $F(.)$, *then (i) there exists a solution* y_t *to the model of Equation (4) such that* (y_t, η_t) *is strictly stationary; (ii) if* $z_t = f(\eta_t, \eta_{t-1}, \ldots)$ *is a solution to the model, then* $y_t = z_t$ *a.s.; and (iii)* y_t *is* L_2*-near epoch dependent on* η_t*. If in addition,* $a_i \leq c_1 \exp(-c_2 i)$ *for positive constants* c_1 *and* c_2, *then the near epoch dependence sequence* $\nu(M)$ *satisfies* $\nu(M) \leq c_1 \exp(-c_2 M^{1/3})$ *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 η_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 η_t implies that η_t is also strong mixing with $\alpha(m) = O(\sum_{t=m+1}^{\infty} G_t^{1/(1+r)})$. 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 ρ_i are not affecting the strict stationarity of the model. The intuition is that because $y_t \geq 0$ a.s., negative ρ_i can only "pull y_t back to zero" and because the model has the trivial lower bound of 0 for y_t, unlike the linear model, this model does not have the potential for y_t to tend to minus infinity.

3 The dynamic Tobit model

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

$$L_T(b) = L_T(c, r, s) = (T-p)^{-1} \sum_{t=p+1}^{T} l_t(b), \tag{6}$$

where

$$l_t(b) = I(y_t > 0) \log(s^{-1}\phi((y_t - \sum_{i=1}^{p} r_i y_{t-i} - c'x_t)/s))$$
$$+ I(y_t = 0) \log(\Phi((-\sum_{i=1}^{p} r_i y_{t-i} - c'x_t)/s)). \tag{7}$$

In order for the loglikelihood function to be maximized at the true parameter β, it appears hard to achieve more generality than to assume that ε_t is distributed normally given $y_{t-1}, \ldots, y_{t-p}, x_t$. This assumption is close to assuming that ε_t given x_t and all lagged y_t is normally distributed, which would then imply that ε_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 β^*. Such a result can be established under general linear process assumptions on (x_t', ε_t), 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 $w_t = (y_{t-1}, \ldots, y_{t-p}, x_t', 1)'$. The "1" at the end of the definition of w_t allows us to write "$b'w_t$". For showing consistency, we need the following two assumptions. Below, let $|.|$ denote the usual matrix norm defined as $|M| = (\mathrm{tr}(M'M))^{1/2}$, and let $\| X \|_r = (E|X|^r)^{1/r}$.

Assumption 1 *The linear process $z_t = (x_t', \varepsilon_t)'$ satisfies $z_t = \sum_{j=0}^{\infty} \Pi_j v_{t-j}$, where the v_t are i.i.d. $(k \times 1)$ vectors, $\| v_t \|_r < \infty$ for some $r \geq 1$, the coefficient matrices Π_j satisfy and $\sum_{t=0}^{\infty} G_t^{1/(1+r)} < \infty$ where $G_t = (\sum_{j=t}^{\infty} |\Pi_j|^2)^{r/2}$, $x_t \in \mathbb{R}^q$, and*

$$y_t = \max(0, \sum_{i=1}^{p} \rho_i y_{t-i} + \gamma' x_t + \varepsilon_t). \tag{8}$$

Assumption 2

1. *The linear process* $z_t = (x_t', \varepsilon_t)'$ *satisfies* $z_t = \sum_{j=0}^{\infty} \Pi_j v_{t-j}$, *where the* v_t *are i.i.d.*, $\| v_t \|_r < \infty$, *and the coefficient matrices* Π_j *satisfy* $\sum_{t=0}^{\infty} G_t^{1/(1+r)} < \infty$ *where* $G_t = (\sum_{j=t}^{\infty} |\Pi_j|^2)^{r/2}$.

2. *Conditional on* $(x_1, \ldots, x_T)$, ε_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. $E w_t w_t' I(\sum_{i=1}^{p} \rho_i y_{t-i} + \gamma' x_t > \delta)$ *is positive definite for some positive* δ.

Theorem 2 *Under Assumption 1 and 2,* $\hat{\beta}_T \xrightarrow{p} \beta$.

The proofs of this and the theorems to follow (i.e., all proofs except for that of Theorem 1) can be found in a full length version of this paper that is available on the websites of both authors (`http://www.clas.wayne.edu/herrera` and `http://www.econ.ohio-state.edu/dejong`).

For asymptotic normality, we need the following additional assumption.

Assumption 3

1. β *is in the interior of* B.

2. $I = E(\partial/\partial b) l_t(\beta)(\partial/\partial b') l_t(\beta) = -E(\partial/\partial b)(\partial/\partial b') l_t(\beta)$ *is invertible.*

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

4 Powell's LAD for dynamic censored regression

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

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

$$= (T-p)^{-1} \sum_{t=p+1}^{T} |y_t - \max(0, \sum_{i=1}^{p} r_i y_{t-i} + c'x_t)| \tag{9}$$

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

Assumption 4

1. *$\beta \in B$, where B is a compact subset of $\mathbb{R}^{p+q}$.*
2. *The conditional distribution $F(\varepsilon_t|w_t)$ satisfies $F(0|w_t) = 1/2$, and $f(\varepsilon|w_t) = (\partial/\partial\varepsilon)F(\varepsilon|w)$ is continuous in ε on a neighborhood of 0 and satisfies $c_2 \geq f(0|w_t) \geq c_1 > 0$ for constants $c_1, c_2 > 0$.*
3. *$E|x_t|^3 < \infty$, and $Ew_t w_t' I(\sum_{i=1}^{p} \rho_i y_{t-i} + \gamma' x_t > \delta)$ is nonsingular for some positive δ.*

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

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

$$\psi(w_t, \varepsilon_t, b) = I(b'w_t > 0)(1/2 - I(\varepsilon_t + (\beta - b)'w_t > 0))w_t. \tag{10}$$

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

Assumption 5

1. *β is in the interior of B.*
2. *Defining $G(z, b, r) = EI(|w_t'b| \leq |w_t|z)|w_t|^r$, we have for z near 0, for $r = 0, 1, 2$,*

$$\sup_{|b-\beta|<\zeta_0} |G(z, b, r)| \leq K_1 z. \tag{11}$$

3. *The matrix*

$$\Omega = \lim_{T\to\infty} E(T^{-1/2} \sum_{t=1}^{T} \psi(w_t, \varepsilon_t, \beta))(T^{-1/2} \sum_{t=1}^{T} \psi(w_t, \varepsilon_t, \beta))' \tag{12}$$

is well-defined, and $N = Ef(0|w_t)I(w_t'\beta > 0)w_t w_t'$ is invertible.

4. *For some* $r \geq 2$, $E|x_t|^{2r} < \infty$, $E|\varepsilon_t|^{2r} < \infty$, *and* $|\Pi_j| \leq c_1 \exp(-c_2 j)$ *for positive constants* c_1 *and* c_2.

5. *The conditional density* $f(\varepsilon|w_t)$ *satisfies, for a nonrandom Lipschitz constant* L_0,

$$|f(\varepsilon|w_t) - f(\tilde{\varepsilon}|w_t)| \leq L_0|\varepsilon - \tilde{\varepsilon}|. \tag{13}$$

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

Assumption 5.1 is identical to Powell's Assumption P.2, and Assumption 5.2 is the same as Powell's Assumption R.2. Theorem 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

$$y_t = \max(0, \gamma_1 + \gamma_2 x_t + \sum_{i=1}^{p} \rho_i y_{t-i} + \varepsilon_t)$$

where

$$x_t = \alpha_1 + \alpha_2 x_{t-1} + v_t,$$

$\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$, and $v_t \sim N(0, \sigma_v^2)$. For our simulations, we consider the cases $p = 1$ and $p = 2$. Many configurations for α_1, α_2, γ_1, γ_2, σ_v^2, and σ_ε^2 were considered. To conserve space, we only report results for $p = 2$, $\gamma_1 = 1$, $\gamma_2 = 1$, $\alpha_1 = \alpha_2 = 0.5$, $\sigma_\varepsilon^2 = \sigma_v^2 = 1$. We conducted simulations for $(\rho_1, \rho_2) \in \{(0.2, 0.1), (0.5, 0.1), (0.8, 0.1), (0, -0.3), (0.3, -0.3), (0.6, -0.3), (0.9, -0.3)\}$. Note that, in contrast with Honoré and Hu (2004), in our simulations the values of ρ_i are not restricted to be non-negative. The number of replications used to compute the bias reported in the tables is 10,000.

Table 1 reports the simulation results. For the dynamic Tobit model estimates of $\beta = (\rho', \gamma', \sigma_\varepsilon)$, with $\rho' = (\rho_1, \rho_2)$, and $\gamma' = (\gamma_1, \gamma_2)$ are obtained via maximum likelihood. Powell's LAD estimates of the dynamic censored regression model where $\beta = (\rho', \gamma')$, with $\rho' = (\rho_1, \rho_2)$, and $\gamma' = (\gamma_1, \gamma_2)$ are obtained using the BRCENS algorithm proposed by Fitzenberger (1997a,b). As we mentioned in section 4, because Powell's LAD estimator does not provide a first-round estimator of σ_ε we redefine β as to not include σ_ε. 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 Table 1). As expected, the bias decreases as the sample size increases.

6 Empirical Application

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 what follows we re-evaluate their findings.

6.1 Data and Summary of Previous Results

The data used by Demiralp and Jordà (2002) are daily and span the period between April 25, 1984 and August 14, 2000. They divide the sample in three subsamples: (i) the period preceding the Fed decision to publicly announce changes in the federal fund rate target on February 4, 1994; (ii) 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; and (iii) the period following the shift to the CRA system.

Open market operations are classified in six groups. Operations that add liquidity are overnight reversible repurchase agreements, term repurchase agreements, and permanent purchases (i.e., T-bill purchases and coupon purchases). Operations that drain liquidity are overnight sales, term matched-sale purchases, and permanent sales (i.e., 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 maintenance-period 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: $NEED_t = f_t - [f^*_{m(t)-1} + w_t E_{m(t)-1}(\Delta f^*_{m(t)})]$; $EXPECT_t = E_{m(t)-1}(\Delta f^*_{m(t)})$; $SURPRISE_t = \Delta f^*_t - E_{m(t)-1}(\Delta f^*_{m(t)})$. $m(t)$ denotes the maintenance period to which observation in day t belongs, f_t is the federal funds rate in day t; $f^*_{m(t)-1}$ is the value of the target in the previous maintenance period; $E_{m(t)-1}(\Delta f^*_{m(t)})$ is the expectation of a target change in day t, conditional on the information available at the beginning of the maintenance period; and w_t is the probability of a target change on date t. ($E_{m(t)-1}(\Delta f^*_{m(t)})$, and w_t are both calculated using the ACH model of Hamilton and Jordà, 2002). This decomposition reflects 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, $NEED_t$ represents a proxy for the projected reserve need, and changes in the federal funds rate are separated into an expected component, $EXPECT_t$, and a surprise component, $SURPRISE_t$.

Because the Open Market Desk engaged in open market operations on 60% of the days in the sample (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. To allow for a different response of sales and purchases –with varying degrees of permanence– to changes in the explanatory variables they estimate separate regressions for each of the six types of operation and each of the periods of interest. Very few term and permanent sales were carried out during the 1998-2000 and 1984-1994 periods respectively, thus no regressions are estimated for this type of operation in these subsamples. Demiralp and Jordà (2002) estimate the following model:

$$y_t = \max(0, \sum_{m=1}^{10} \gamma^{\alpha}_m DAY_{tm} + \sum_{j=1}^{3} \rho_j y_{t-j} + \sum_{j=1}^{3} \upsilon'_j \mathbf{z}_{t-j} + \sum_{m=1}^{10} \gamma^N_m NEED_{t-m} \times DAY_{tm}$$
$$+ \sum_{m=1}^{10} \gamma^E_m EXPECT_{t-m} \times DAY_{tm} + \sum_{j=0}^{3} \gamma^S_j SURPRISE_{t-j} + \varepsilon_t) \quad (14)$$

where y_t denotes one of the open market operation of interest, that is, y_t equals either overnight purchases (OB_t), term purchases (TB_t), permanent purchases (PB_t), overnight sales (OS_t), term sales (TS_t), or permanent sales (PS_t). $\mathbf{z}_t$ denotes a vector containing the remaining five types of operations. For instance, if $y_t = OB_t$ (overnight purchases), then $\mathbf{z}_t = [TB_t, PB_t, OS_t, TS_t, PS_t]$. DAY_{tm} denotes a vector of maintenance-day dummies, and ε_t is a stochastic disturbance.

We 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 2. (For the complete set of parameter estimates, see Tables A.1 and A.2 in the full version of this paper at `http://www.clas.wayne.edu/herrera/` and `http://www.econ.ohio-state.edu/dejong`.) Because we are interested in whether the roots of the polynomial $\rho_{max}(z) = 1 - \sum_{i=1}^{3} \max(0, \rho_i) z^i$ 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 Market 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.

6.2 Model and estimation procedure

From here on we will restrict our attention to the Open Market 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; thus, they are informative regarding the Open Market Desk's reaction function. 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.

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, $OTB_t = OB_t + TB_t$; (b) permanent purchases, which include T-bill purchases and coupon purchases, PB_t; (c) temporary sales, which include overnight and term matched sale-purchases, $OTS_t = OS_t + TS_t$; and (d) permanent sales, which comprise T-bill sales and coupon sales, PS_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_t.

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

$$OTB_t = \max(0, \gamma + \sum_{m=1}^{4} \gamma_m^{\alpha} D_{tm} + \sum_{j=1}^{3} \rho_j OTB_{t-j} + \sum_{j=1}^{3} \gamma_j^{TS} OTS_{t-j} + \sum_{j=1}^{3} \gamma_j^{PB} PB_{t-j}$$

$$+\sum_{j=1}^{3}\gamma_j^{PS}PS_{t-j}+\sum_{m=1}^{10}\gamma_m^{N}NEED_{t-m}\times DAY_{tm}+\sum_{m=1}^{10}\gamma_m^{E}EXPECT_{t-m}\times DAY_{tm}$$

$$+\sum_{j=0}^{3}\gamma_j^{S}SURPRISE_{t-j}+\varepsilon_t) \qquad (15)$$

where OTB_t denotes temporary purchases, OTS_t denotes temporary sales, PB_t denotes permanent purchases, PS_t denotes permanent sales, DAY_{tm} denotes a vector of maintenance-day dummies, D_{tm} is such that $D_{t1} = DAY_{t1}$ (First Thursday), $D_{t2} = DAY_{t2}$ (First Friday), $D_{t3} = DAY_{t7}$ (Second Friday), and $D_{t4} = DAY_{t,10}$ (Settlement Wednesday), and ε_t is a stochastic disturbance.

This model is a restricted version of (14) 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).

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 Barrodale-Roberts 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 $\widetilde{\psi}(w_t,\widetilde{b}) = I(\widetilde{b}'w_t > 0)[\frac{1}{2} - I(y_t < \widetilde{b}'w_t)]w_t$, following the suggestions of Andrews (1991) to select the bandwidth for the Bartlett kernel. To compute $\widetilde{N}$, we estimate $f(0|w_t)$ 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 and corresponding quasi-maximum likelihood standard errors are presented in the first two columns of Table 3. 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 $\widehat{\rho}_{max}(z) = 1 - 0.2639z - 0.2916z^2 - 0.3054z^3$ lie all outside the unit circle, and the smallest modulus 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, TB_{t-j}. This persistence suggests that in order to attain the desired target, the Open Market Desk had to exercise pressure on the fed funds market in a gradual manner, on consecutive days. The negative and statistically significant coefficients on lagged temporary sales, TS_{t-j}, 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 $NEED_{t-1} \times DAY_{tm}$ variables is consistent with an accommodating behavior of the Fed to deviations of the federal funds rate from its target. The Tobit estimates suggest that expectations of target changes were accommodated in the first 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 $SURPRISE_{t-j}$ can be interpreted as evidence of the announcement effect. 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; see Demiralp and Jordà, 2004).

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 $Var(\varepsilon_t|w_t) = \sigma^2 \exp(\delta' z_t)$, where z_t is a vector that contains all elements in w_t but the constant, rejects the null $H_0 : \delta = 0$ at the 1% level. In addition, the Jarque-Bera statistic leads us to reject the null that the residuals are normally distributed at a 1% level.

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 indicative 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 3, respectively. Notice that, in this case, the smallest root of the lag polynomial $\widetilde{\rho}_{max}(z) = 1 - 0.068z - 0.093z^2 - 0.073z^3$ 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 Open Market Desk's reaction function. First, the CLAD estimates imply a considerably smaller degree of persistence in temporary purchases. The magnitude of the ρ_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); 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 Open Market 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 $EXPECT$ is only statistically significant on the first and eight day of the maintenance period. This suggests the Fed is only seldom willing to accommodate (or profit) from anticipated changes in the target. Although both estimation methods reveal 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).

Most of the coefficients on the contemporaneous and lagged $SURPRISE$ are negative, which is consistent with the liquidity effect. That is, in order to steer the federal funds rate towards a new lower target level the Open Market 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.

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 $\rho_{max}(z) = 1 - \sum_{i=1}^{p} \max(0, \rho_i) z^i$ has its roots outside the unit circle and that the error term, ε_t, is itself allowed to be potentially correlated. Hence, in terms of the conditions on $\rho_{max}(z)$ and the dependence allowed for ε_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 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.

References

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

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.

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.

Davidson, J. (1994) *Stochastic Limit Theory.* Oxford: Oxford University Press.

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.

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.

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.

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 Paper, University of Wisconsin.

Honoré, B. and L. Hu (2004), Estimation of cross sectional and panel data censored regression models with endogeneity, *Journal of Econometrics* 122, 293-316.

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.

Lee, L-F. (1999), Estimation of dynamic and ARCH Tobit models, *Journal of Econometrics* 92, 355-390.

Meyn, S.P. and R.L. Tweedie (1994) *Markov Chains and Stochastic Stability.* Berlin: Springer-Verlag, second edition.

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.

Zangari, P.J. and H. Tsurumi (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.

Table 1
Simulation results for censored regression model with p=2

Tobit							CLAD						
	True parameter values	Bias						True parameter values	Bias				
		T=100	T=300	T=600	T=1000	T=2000			T=100	T=300	T=600	T=1000	T=2000
γ_1	1	0.0394	0.0125	0.0071	0.0044	0.0018	γ_1	1	0.0309	0.0101	0.0036	0.0031	0.0020
γ_2	1	0.0070	0.0027	0.0004	0.0007	0.0000	γ_2	1	0.0078	0.0037	0.0009	0.0008	-0.0003
ρ_1	0.2	-0.0062	-0.0018	-0.0008	-0.0006	-0.0002	ρ_1	0.2	-0.0050	-0.0018	-0.0005	-0.0007	-0.0002
ρ_2	0.1	-0.0099	-0.0034	-0.0018	-0.0012	-0.0004	ρ_2	0.1	-0.0091	-0.0031	-0.0014	-0.0007	-0.0004
σ^2	1	-0.0408	-0.0129	-0.0067	-0.0044	-0.0020							
γ_1	1	0.0613	0.0169	0.0095	0.0053	0.0028	γ_1	1	0.0588	0.0170	0.0101	0.0045	0.0029
γ_2	1	0.0058	0.0022	0.0012	0.0009	0.0004	γ_2	1	0.0075	0.0027	0.0002	0.0006	-0.0002
ρ_1	0.5	-0.0072	-0.0023	-0.0011	-0.0009	-0.0004	ρ_1	0.5	-0.0091	-0.0015	-0.0011	-0.0007	-0.0003
ρ_2	0.1	-0.0062	-0.0015	-0.0011	-0.0004	-0.0003	ρ_2	0.1	-0.0041	-0.0024	-0.0010	-0.0003	-0.0003
σ^2	1	-0.0394	-0.0138	-0.0054	-0.0044	-0.0018							
γ_1	1	0.2140	0.0647	0.0330	0.0230	0.0087	γ_1	1	0.2206	0.0657	0.0323	0.0181	0.0094
γ_2	1	0.0076	0.0022	0.0016	0.0002	0.0005	γ_2	1	0.0065	0.0029	0.0003	0.0007	-0.0002
ρ_1	0.8	-0.0091	-0.0024	-0.0016	-0.0005	-0.0005	ρ_1	0.8	-0.0107	-0.0025	-0.0015	-0.0010	-0.0003
ρ_2	0.1	-0.0020	-0.0010	-0.0001	-0.0006	0.0001	ρ_2	0.1	-0.0007	-0.0009	-0.0001	0.0000	-0.0002
σ^2	1	-0.0411	-0.0131	-0.0076	-0.0041	-0.0020							
γ_1	1	0.0161	0.0054	0.0023	0.0025	0.0014	γ_1	1	-0.0035	0.0008	0.0008	0.0009	0.0002
γ_2	1	0.0058	-0.0001	0.0015	-0.0001	0.0001	γ_2	1	0.0195	0.0055	0.0017	0.0013	0.0008
ρ_1	0	-0.0068	-0.0006	-0.0011	-0.0002	-0.0007	ρ_1	0	-0.0063	-0.0015	-0.0009	-0.0008	-0.0001
ρ_2	-0.3	-0.0069	-0.0026	-0.0016	-0.0010	-0.0002	ρ_2	-0.3	-0.0115	-0.0039	-0.0018	-0.0011	-0.0005
σ^2	1	-0.0380	-0.0121	-0.0067	-0.0043	-0.0022							
γ_1	1	0.0167	0.0066	0.0017	0.0024	0.0012	γ_1	1	0.0066	0.0029	0.0022	0.0010	0.0008
γ_2	1	0.0061	0.0022	0.0006	0.0007	0.0003	γ_2	1	0.0126	0.0044	0.0011	0.0009	-0.0002
ρ_1	0.3	-0.0064	-0.0029	-0.0012	-0.0009	-0.0004	ρ_1	0.3	-0.0078	-0.0012	-0.0008	-0.0007	-0.0001
ρ_2	-0.3	-0.0055	-0.0016	-0.0004	-0.0007	-0.0002	ρ_2	-0.3	-0.0041	-0.0027	-0.0014	-0.0004	-0.0003
σ^2	1	-0.0369	-0.0139	-0.0061	-0.0040	-0.0017							
γ_1	1	0.0220	0.0060	0.0035	0.0016	0.0010	γ_1	1	0.0142	0.0054	0.0027	0.0015	0.0013
γ_2	1	0.0075	0.0016	0.0008	0.0010	0.0003	γ_2	1	0.0093	0.0036	0.0016	0.0008	-0.0002
ρ_1	0.6	-0.0072	-0.0020	-0.0008	-0.0007	-0.0005	ρ_1	0.6	-0.0078	-0.0014	-0.0010	-0.0008	-0.0003
ρ_2	-0.3	-0.0030	-0.0007	-0.0007	-0.0004	0.0001	ρ_2	-0.3	-0.0012	-0.0019	-0.0007	-0.0001	-0.0001
σ^2	1	-0.0399	-0.0124	-0.0059	-0.0044	-0.0020							
γ_1	1	0.0332	0.0128	0.0055	0.0030	0.0011	γ_1	1	0.0321	0.0117	0.0039	0.0030	0.0018
γ_2	1	0.0074	0.0026	0.0014	0.0003	0.0005	γ_2	1	0.0072	0.0031	0.0010	0.0007	-0.0003
ρ_1	0.9	-0.0081	-0.0031	-0.0015	-0.0007	-0.0003	ρ_1	0.9	-0.0077	-0.0023	-0.0017	-0.0009	-0.0003
ρ_2	-0.3	-0.0001	-0.0001	0.0002	0.0000	0.0000	ρ_2	-0.3	-0.0003	-0.0006	0.0005	0.0001	0.0000
σ^2	1	-0.0391	-0.0130	-0.0056	-0.0042	-0.0023							

Model: $y_t = \max(0, \gamma_1 + \gamma_2 * x_t + \rho_1 * y_{t-1} + \rho_2 * y_{t-2} + \varepsilon_t)$

$x_t = 0.5 + 0.5x_{t-1} + v_t$

Table 2
Demiralp and Jordà (2002) Tobit regression for Open Market Operations
Coefficient Estimates for Lags of the Dependent Variable

	OB			TB			PB		
	1984-94	**1994-1998**	**1998-2000**	**1984-94**	**1994-1998**	**1998-2000**	**1984-94**	**1994-1998**	**1998-2000**
Lag 1	0.628	0.513	0.063	0.265	0.129	-0.024	0.145	0.217	-0.068
	(0.141)	(0.093)	(0.078)	(0.146)	(0.105)	(0.120)	(0.044)	(0.174)	(0.282)
Lag 2	0.268	0.300	-0.017	0.594	0.080	-0.039	0.108	0.329	0.737
	(0.146)	(0.091)	(0.070)	(0.099)	(0.095)	(0.114)	(0.046)	(0.164)	(0.213)
Lag 3	0.025	0.236	0.087	0.404	0.154	0.113	0.101	0.281	0.421
	(0.148)	(0.087)	(0.065)	(0.104)	(0.098)	(0.120)	(0.045)	(0.168)	(0.227)
Smallest root									
	1.063	*0.972*	2.154	*0.894*	1.574	2.069	1.661	1.095	*0.940*

	OS			TS			PS		
	1984-94	**1994-1998**	**1998-2000**	**1984-94**	**1994-1998**	**1998-2000**	**1984-94**	**1994-1998**	**1998-2000**
Lag 1	0.938	1.897	1.104	0.759	2.623	-	-	-98.549	-31.074
	(0.240)	(0.406)	(0.685)	(0.269)	(2.871)			(12658677)	(19394016)
Lag 2	-0.064	-0.025	-0.185	1.269	0.487	-		-77.082	1.515
	(0.213)	(0.357)	(0.703)	(0.261)	(0.502)		-	(11727925)	(0.840)
Lag 3	0.197	0.522	0.084	1.069	0.322	-		-77.958	-62.098
	(0.178)	(0.336)	(0.702)	(0.239)	(0.591)		-	(11021784)	(28391370)
Smallest root									
	0.908	*0.494*	*0.858*	*0.557*	*0.353*			*1.000*	*0.812*

Note: Standard errors reported in parenthesis. Smallest root denotes the smallest root of the $\rho max(z)$ polynomial; for complex roots the modulus is reported.

Table 3
Tobit and CLAD Estimates for Open Market Temporary Purchases
1986-2000

	Tobit			CLAD		
Variable	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*Day7	-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 root of the $\rho_{max}(z)$ polynomial; for complex roots the modulus is reported.

Appendix

Define $\hat{y}_t^m = 0$ for $m \leq 0$ and $\hat{y}_t^m = \max(0, \eta_t + \sum_{i=1}^{p} \rho_i \hat{y}_{t-i}^{m-i})$. 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* $(1-\max(0,\rho_1)B-\ldots-\max(0,\rho_p)B^p)$ *has all its roots outside the unit circle and* $\sup_{t\in\mathbb{Z}} E\max(0,\eta_t) < \infty$*, then for an almost surely finite random variable* $f_t = f(\eta_t, \eta_{t-1}, \ldots) = \sum_{j=0}^{\infty} L_1^j \max(0, \eta_{t-j})$*, and* L_1^j *that are such that* $L_1^j \leq c_1 \exp(-c_2 j)$ *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,

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

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

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

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

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

$$\leq \sum_{j=0}^{\infty} L_1^j \max(0,\eta_{t-j}).$$

The L_i^j satisfy, for $j \geq 2$,

$$L_1^j = L_2^{j-1} + \max(0,\rho_1)L_1^{j-1},$$

$$L_2^j = L_3^{j-1} + \max(0, \rho_2) L_1^{j-1},$$

$$\vdots$$

$$L_{p-1}^j = L_p^{j-1} + \max(0, \rho_{p-1}) L_1^{j-1},$$

$$L_p^j = \max(0, \rho_p) L_1^{j-1}.$$

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

$$(1 - \sum_{j=1}^{p} \max(0, \rho_j) B^j) L_1^j = 0.$$

From the fact that the above lag polynomial has all its roots outside the unit circle by assumption, it follows that $L_1^j \leq c_1 \exp(-c_2 j)$ for positive constants c_1 and c_2. Also, if $\sup_{t \in \mathbb{Z}} E \max(0, \eta_t) < \infty$, then $\sum_{j=0}^{\infty} L_1^j \max(0, \eta_{t-j})$ 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 η_t is strictly stationary, $\rho_{max}(B)$ has all its roots outside the unit circle, and $\| \max(0, \eta_t) \|_r < \infty$ for some $r \geq 1$, then $\sup_{t \in \mathbb{Z}} \| f_t \|_r < \infty$.*

Proof of Lemma 2:

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

$$\| f_t \|_r \leq \sum_{j=0}^{\infty} L_1^j \| \max(0, \eta_{t-j}) \|_r < \infty.$$

□

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

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

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

and

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

Lemma 3 *Assume that* η_t *is strictly stationary and strong mixing and satisfies* $\| \max(0,\eta_t) \|_2 < \infty$*. Then for all* $t \in \mathbb{Z}$ *and* $\delta > 0$*, as* $m \to \infty$*,*

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

Proof of Lemma 3:

Note that we can write

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

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

Note that

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

and for all t and j,

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

$$= \eta_{t-(m-p+1-l)-j} + \sum_{i=1}^{p} \sum_{k=0}^{\infty} \rho_i L_1^k \max(0, \eta_{t-(m-p+1-l)-j-i-k})$$

$$= \eta_{t-(m-p+1-l)-j} + \sum_{k=0}^{\infty} \max(0, \eta_{t-(m-p+1-l)-j-i-k}) \sum_{i=1}^{p} \rho_i L_1^{k-i} I(i \leq k) = w_{t-(m-p+1-l)-j}$$

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

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

$$\leq \| \max(0, \eta_t) \|_2 \sum_{k=M+1}^{\infty} \sum_{i=1}^{p} \rho_i L_1^{k-i} I(i \leq k),$$

and the last expression converges to 0 as $M \to \infty$ at exponential rate because L_1^k converges to zero at an exponential rate. Therefore, because $H^{\zeta}(\cdot)$ is Lipschitz-continuous,

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

is also L_2-near epoch dependent on η_t with an exponentially decreasing $\nu(\cdot)$ sequence, and so is

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

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

Lemma 4 *Under the assumptions of Theorem 1, for all $\zeta > 0$,*

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

Proof of Lemma 4:

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

$$\eta_t + \sum_{i=1}^{p} \rho_i f_{t-i} \leq u_t + c \sum_{j=0}^{\infty} |a|^j |u_{t-j}|.$$

Noting that

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

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

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

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

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

(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 $\sum_{j=0}^{\infty}|a|^j|w_{t-j}|$, it now follows that for all $\zeta > 0$,

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

□

Lemma 5 *For some almost surely finite random variable y_t such that (y_t, η_t) is strictly stationary,*

$$\hat{y}_t^m \xrightarrow{as} y_t \quad as \quad m\to\infty.$$

Proof of Lemma 5:

We will use the Cauchy criterion to show that $\hat{y}_t^m$ converges almost surely, and we will define y_t to be this limit. By the Cauchy criterion, $\hat{y}_t^m$ converges a.s. if $\max_{k \geq m} |\hat{y}_t^k - \hat{y}_t^m|$ converges to zero in probability as $m \to \infty$. 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 $\max_{k \geq m} |\hat{y}_t^k - \hat{y}_t^m| = 0$ for all $m > p$ if there can be found p consecutive "small" η_{t-l} that are negative and large in absolute value in the range $l = 1, \ldots, m-1$; i.e. if

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

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

$$P[\max_{k \geq m} |\hat{y}_t^k - \hat{y}_t^m| > 0]$$

$$\leq P[\text{there are no } p \text{ consecutive "small" } \eta_t]$$

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

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

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

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

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

$$\leq \exp(-(m-p)c)) + P[(m-p)^{-1}\sum_{l=1}^{m-p}(I_{tl}^{\zeta}\log(\delta) + \log(1+\delta)(1-I_{tl}^{\zeta})) > -c], \tag{16}$$

where

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

and

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

for

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

Note that $I_{tl} \geq I_{tl}^{\zeta}$ because $I(x \leq 0) \geq H^{\zeta}(x)$. Both terms in Equation (16) now converge to zero as $m \to \infty$ for a suitable choice of ζ, c and δ if

$$E(m-p)^{-1}\sum_{l=1}^{m-p}(I_{tl}^{\zeta}\log(\delta) + \log(1+\delta)(1-I_{tl}^{\zeta}))$$

$$= E(I_{tl}^{\zeta}\log(\delta) + \log(1+\delta)(1-I_{tl}^{\zeta})) < 0 \tag{17}$$

and

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

satisfies a weak law of large numbers as $m \to \infty$. This weak law of large numbers is proven in Lemma 3. Now if $EI_{tl}^{\zeta} > 0$, we can pick $\delta > 0$ small enough to satisfy the requirement of Equation (17). Now,

$$EI_{tl}^{\zeta} = E\prod_{j=1}^{p} H^{\zeta}(\eta_{t-l-j} - \sum_{i=1}^{p}\rho_j f_{t-l-j-i})$$

$$\geq E\prod_{j=1}^{p} I(\eta_{t-l-j} + \sum_{i=1}^{p}\rho_j f_{t-l-j-i} \leq -\zeta),$$

and the last term is positive by Lemma 4.
Since $\hat{y}_t^m = f_m(\eta_t, \ldots, \eta_{t-m})$ is strictly stationary because it depends on a finite numbers of η_t, $\lim_{m\to\infty}(\hat{y}_t^m, \eta_t) = (y_t, \eta_t)$ is also strictly stationary. □

Proof of Theorem 1:

Noting that y_t as constructed in Lemma 5 is a solution to the dynamic censored regression model, part (i) of Theorem 1 follows. Also, by the reasoning of Lemma 5 it follows that any $z_t = f(\eta_t, \eta_{t-1}, \ldots)$ that is a solution to the model also satisfies $\max_{k\geq m}|\hat{y}_t^k - z_t| = 0$, 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 (y_t, η_t) and by noting that the conditional expectation is the best L_2-approximation,

$$\sup_{t\in\mathbb{Z}} E|y_t - E(y_t|\eta_{t-m}, \eta_{t-m+1}, \ldots, \eta_t)|^2 = E|y_t - E(y_t|\eta_{t-m}, \eta_{t-m+1}, \ldots, \eta_t)|^2 \leq E|y_t - \hat{y}_t^m|^2,$$

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

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

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

for $p \geq 1$ and $q \geq 1$ such that $p^{-1} + q^{-1} = 1$. Now by choosing p small enough, $E|f_t|^{2p} < \infty$ by assumption and by Lemma 1. Therefore, it suffices to show that $P(|\hat{y}_t^m - y_t| > 0)$ decays as $c_1\exp(-c_2 m^{1/3})$ with m. By the earlier reasoning,

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

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