Dynamic censored regression and the Open Market Desk reaction function

Robert de Jong* Ana María Herrera†

July 13, 2004

Abstract

The censored regression model and the Tobit model are standard tools in econometrics. This paper provides a formal asymptotic theory for dynamic time series censored regression when lags of the dependent variable have been included among the regressors. We derive fading memory properties of the model under the assumption that the regression error is strong mixing. This paper shows 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. We conclude 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. The asymptotic theory for the Tobit model in cross-section situations has long been understood;

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

†Department of Economics, Michigan State University, 215 Marshall Hall, East Lansing, MI 48824, email herrer20@mail.msu.edu. We thank Selva Demiralp and Oscar Jorda for making their data available, and Bruce Hansen for the use of his code for kernel density estimation.

see for example the treatment in Amemiya (1973). However, there appears to be no asymptotic theory for the time series dynamic censored regression model where lagged values of the dependent variable are included among the regressors, and this paper seeks to provide such a treatment. The dynamic censored regression model under consideration is

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

where x_t denotes the regressor and ε_t is a regression error, and we assume that $\gamma \in \mathbb{R}^q$, and we define $\sigma^2 = E\varepsilon_t^2$. One feature of the treatment of the censored regression model in this paper is that ε_t is itself allowed to be strong mixing, and therefore potentially correlated, process. While stationarity results for general nonlinear models have been derived in e.g. Meyn and Tweedie (1994), there appear to be no results for the case where innovations are not i.i.d. (i.e. weakly dependent or heterogeneously distributed). The reason for this appears to be that the derivation of results such as those of Meyn and Tweedie (1994) depends on a Markov chain argument, and this line of reasoning appears to break down when the i.i.d. assumption is dropped. This means that in the current setting, Markov chain techniques cannot be used for the derivation of stationarity properties, which complicates the analysis substantially, but also puts the 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 $\rho_{max}(z) = 1 - \sum_{j=1}^{p} \max(0, \rho_j) z^j$ 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 attempt to analyze the dynamic Tobit model on a level of generality that is comparable to the level of generality under which results for the linear model AR(p) model can be derived. Note that intuitively, negative values for ρ_j can never be problematic when considering the stationarity properties of y_t. This intuition is formalized by the fact that only $\max(0, \rho_j)$ shows up in our stationarity requirement.

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

Papers that estimate censored regression models in a time series framework 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 (2001) who use a dynamic Tobit model to analyze the determinants of U.S. butter prices with particular attention to the effects of market liberalization via reductions in floor prices. Zangari and Tsurumi (1996), and Wei (1999) use a Bayesian approach to analyze the demand for Japanese exports of passenger cars to the U.S., which were subject to quotas negotiated between the U.S. and Japan after the oil crisis of the 1970's.

We are aware of only a few applications in time series macroeconomics. Peristiani (1994) estimates a censored regression model to study the decision of individual banks to borrow from the Federal Reserve's discount window. Demiralp and Jorda (2002) use a dynamic Tobit model to study the determinants of the daily transactions conducted by the Open Market Desk.

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

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

2 Main results

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

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

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

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. In the situation considered in this paper, y_t and η_t will be assumed to be strictly stationary, and in that situation, the "sup" can be deleted from the definition of near epoch dependence.
For studying the weak dependence properties of the dynamic censored regression model, assume that y_t is generated as

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

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.
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. In order to do this however, we first need to derive a moment bound for y_t. Let B denote the backward operator, and define the lag polynomial $\rho_{max}(B) = 1 - \sum_{j=1}^{p} \max(0, \rho_j) B^j$.

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

The formal weak dependence result for the dynamic censored regression model is now the following:

Theorem 2 *If η_t is strictly stationary and strong mixing, $\rho_{max}(B)$ has all its roots outside the unit circle, $\| \max(0, \eta_t) \|_2 < \infty$, and*

$$P[\eta_{t-1} \leq y_1, \ldots, \eta_{t-p} \leq y_p | \eta_{t-p}, \eta_{t-p-1}, \ldots] \geq F(y_1, \ldots, y_p) > 0 \tag{5}$$

for all $(y_1, \ldots, y_p) \in \mathbb{R}^p$, then (i) (y_t, η_t) is strictly stationary; and (ii) y_t is L_2-near epoch dependent on η_t.

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

$$l_t(b) = I(y_t > 0) \log(s^{-1} \phi((y_t - \sum_{j=1}^{p} r_j y_{t-j} - c' x_t)/s))$$

$$+ I(y_t = 0) \log(\Phi((- \sum_{j=1}^{p} r_j y_{t-j} - c' x_t)/s)). \tag{7}$$

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 mixing assumptions, by the use of Theorem 2.

Let $\hat{\beta}_T$ denote a maximizer of $L_T(b)$ over $b \in B$. Define $w_t = (y_{t-1}, \ldots, y_{t-p}, x_t', 1)'$. The "1" at the end of the definition of w_t allows us to write "$b'w_t$". For showing consistency, we need the following two assumptions.

Assumption 1 *$(x_t', \varepsilon_t)'$ is a sequence of strictly stationary strong mixing random variables with α-mixing numbers $\alpha(m)$, where $x_t \in \mathbb{R}^q$ and*

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

Assumption 2

1. *$(x_t', \varepsilon_t)'$ is a sequence of strictly stationary strong mixing random variables with α-mixing numbers $\alpha(m)$, and $E|x_t|^2 < \infty$.*
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 $\mathbb{R}^{p+q+1}$.*
4. *$Ew_t w_t' I(\sum_{j=1}^{p} r_j y_{t-j} + c' x_t > \delta)$ is positive definite for some positive δ.*

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

For asymptotic normality, we need the following additional assumption.

Assumption 3

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

4 Powell's LAD for dynamic censored regression

For this section, define $\beta = (\rho', \gamma')'$, where $\rho = (\rho_1, \ldots, \rho_p)$, and define $b = (r', c')'$ where r is a $(p \times 1)$ vector and c is a $(q \times 1)$ vector and $w_t = (y_{t-1}, \ldots, y_{t-p}, x_t')'$. This redefines the b and 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

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

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

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

Assumption 4

1. *$\beta \in B$, where B is a compact subset of $\mathbb{R}^{p+q}$.*
2. *$\varepsilon_t | w_t$ has a density $f(.|.)$ that is bounded and positive at 0, and $E|\varepsilon_t|^3 < \infty$.*
3. *$E|x_t|^3 < \infty$, and $E w_t w_t' I(\sum_{j=1}^{p} r_j y_{t-j} + c' x_t > \delta)$ is nonsingular for some positive δ.*

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

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

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

$\psi(.,.)$ 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 $r = 0, 1, 2$,*

$$\sup_{|b-\beta|<\zeta_0} |G(z, b, r)| \leq K_1 z. \tag{10}$$

3. $M = Ef(0|w_t)I(w_t'\beta > 0)w_t w_t'$ *is a well-defined matrix, and*

$$\Omega = \lim_{T\to\infty} E(T^{-1/2}\sum_{t=1}^{T}\psi(w_t,\beta))(T^{-1/2}\sum_{t=1}^{T}\psi(w_t,\beta))' \qquad (11)$$

is invertible.

4. For some $r \geq 2$, $E|x_t|^{2r} < \infty$, $E|\varepsilon_t|^{2r} < \infty$, *and the strong mixing numbers* $\alpha(\cdot)$ *for* $(x_t', \varepsilon_t)'$ *satisfy* $\sum_{m=0}^{\infty}\alpha(m)^{1/(4r)} < \infty$, *and* $r > 2(p+q)$.

5. The conditional distribution $F(\varepsilon_t|w_t)$ *satisfies* $F(0|w_t) = 1/2$, *and* $f(\varepsilon|w_t) = (\partial/\partial\varepsilon)F(\varepsilon|w)$ *satisfies* $c_2 \geq f(0|w_t) \geq c_1 > 0$ *for constants* $c_1, c_2 > 0$.

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

$$|f(\varepsilon|w_t) - f(\tilde{\varepsilon}|w_t)| \leq L_0|\varepsilon - \tilde{\varepsilon}|. \qquad (12)$$

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

Assumption 5.1 is identical to Powell's Assumption P.2, and Assumption 5.2 is the same as Powell's Assumption R.2. Theorem 6 imposes rth order moment conditions, for some $r \geq 4$. The conditions imposed by Theorem 6 restrict the rate of polynomial decay of the $\alpha(m)$ sequence, where the rate of polynomial decay depends on the r sequence. These conditions originate from the stochastic equicontinuity proof of Hansen (1996), which is used in the proof. One would expect that some progress in establishing stochastic equicontinuity results for dependent variables could aid in relaxing condition 4 imposed in Theorem 6.

5 Empirical Application

In this section we discuss an application of the dynamic censored regression model. Although there is a significant number of papers that model and estimate the Federal Open Market Committee's (FOMC's) reaction function, we are only aware of one recent study where lags of the dependent variable (i.e. open market operations) are included among the regressors. Demiralp and Jorda (2002) use a dynamic Tobit model to analyze whether the February 4, 1994, Fed decision to publicly announce changes in the federal funds rate target affected the manner in which the Open Market Desk conducts operations. We re-evaluate some findings of Demiralp and Jorda (2002).

5.1 Data

The data used in the analysis can be found in Demiralp and Jorda (2002). The data are daily and span the period between April 25, 1984 and August 14, 2000. We classify open market operations in four groups: (a) temporary purchases comprise overnight reversible repurchase agreements (RP) and term RP; (b) permanent purchases include T-bill purchases, coupon purchases; (c) temporary sales are overnight and term matched sale-purchases; and (d) permanent sales comprise T-bill sales and coupon sales. We restrict our analysis to the change in the maintenance-period-average level of reserves brought about by temporary purchases of the Open Market Desk. Because the computation of reserves is based on a 14-day maintenance period that starts on Thursday and finishes on the "Settlement Wednesday" two weeks later, the maintenance-period average is the object of attention of the Open Desk. Thus, all operations are adjusted according to the number of days spanned by the transaction, and standardized by the aggregate level of reserves held by depository institutions in the maintenance period previous to the execution of the transaction. Daily values for temporary purchases are plotted in Figure 1. Note that, although not clearly apparent in the figure, the Open Market Desk engaged in temporary purchases only 37% of the time.
Changes in the federal funds rate are separated into an expected component and a surprise component. The expected component corresponds to the expectation of a target change in day t, conditional on the information available at the beginning of the 14-day maintenance period. The variable surprise records the unexpected component of the target change for the 115 days in the sample when there was a change in the target, and zero otherwise. As a proxy for the projected reserve need, we use the difference between the federal funds rate in day t, and the value of the target at the start of the maintenance period, plus the expected change in the target.

5.2 Model and estimation procedure

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

$$TB_t = \max(0, \gamma_t^{\alpha} + \sum_{j=1}^{3} \rho_j TB_{t-j} + \sum_{j=1}^{3} \gamma_j^{TS} TS_{t-j} + \sum_{j=1}^{3} \gamma_j^{PB} PB_{t-j} + \sum_{j=1}^{3} \gamma_j^{PS} PS_{t-j}$$

$$+ \sum_{j=1}^{10} \gamma_j^{N} NEED_{t-j} \times DAY_{tj} + \sum_{j=1}^{10} \gamma_j^{E} EXPECT_{t-j} \times DAY_{tj}$$

$$+\sum_{j=0}^{3}\gamma_j^S SURPRISE_{t-j}+\varepsilon_t) \tag{13}$$

where γ_t^α denotes a vector of maintenance-day dummies, TB_t denotes temporary purchases, TS_t denotes temporary sales, PB_t denotes permanent purchases, PS_t denotes permanent sales, $NEED_t$ denotes reserve needs, $EXPECT_t$ denotes the expected change in the federal funds target, $SURPRISE_t$ denotes the unexpected component of the target change, and ε_t is a stochastic disturbance.

This model is a simplified version of the one estimated by Demiralp and Jorda (2002) in that γ_t^α does not include dummies for all days in the maintenance period. Instead, to control for differences in the reserve levels that the Federal Reserve might want to leave in the system at the end of the day, we include only dummies for the first Thursday, the first Friday, the second Friday, and Settlement Wednesday. However, we do allow the response of temporary purchases to reserve needs and expected changes in the Fed funds rate to vary across all days of the maintenance period. Due to computational issues related to the CLAD estimation procedure, we aggregate overnight and term open market purchases, and estimate the model using all the sample.

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

$$\min_b\{(T-p)^{-1}\sum_{t=p+1}^{T}[\frac{1}{2}\text{sgn}(y_t-b'w_t)(y_t-b'w_t)]\} \tag{14}$$

to obtain the $\hat{b}^{(1)}_{CLAD}$ estimates. Then, solve the LP problem for $\hat{b}^{(2)}_{CLAD}$ using the observations for which $\hat{b}^{(1)\prime}_{CLAD}w_t>0$. This procedure is repeated until the set of observation used in two consecutive iterations is the same. Standard errors for $\hat{b}_{CLAD}$ are obtained according to equation (15). We compute $\hat{\Omega}$ as the long-run variance of

$$\psi(w_t,\hat{b}_{CLAD})=I(\hat{b}'_{CLAD}w_t>0)[\frac{1}{2}-I(y_t<\hat{b}'_{CLAD}w_t)]w_t,$$

following the suggestions of Andrews (1991) to select the bandwidth for the Bartlett kernel. To compute $\hat{M}$, we estimate $f(0|w_t)$ using a higher-order Gaussian kernel with the order and bandwidth selected according to Hansen (2003, 2004). The reported standard errors for the Tobit estimates are the quasi-maximum likelihood standard errors.

5.3 Estimation Results

Maximum likelihood estimates of the Tobit model and corresponding standard errors are presented in the first two columns of Table 1. Of interest is the presence of statistically significant coefficients on the lags of the dependent variable, TB_{t-j}. This persistence suggests that in order to attain the desired target, the Open Market Desk had to exercise pressure on the fed funds market in a gradual manner, on consecutive days. The negative and statistically significant coefficients on lagged temporary sales, TS_{t-j}, confirms Demiralp and Jorda's (2002) finding that temporary sales have constituted substitutes for temporary purchases. In other words, in the face of a reserve shortage the Open Market Desk could react by conducting temporary purchases and/or delaying temporary sales. The positive and statistically significant coefficients on the $NEED_{t-1} \times Day$ 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 where accommodated in the first day of the maintenance period, and did not significantly affect temporary purchases on the remaining days. As for the effect of surprise changes in the target, the estimated coefficients are statistically insignificant. According to Demiralp and Jorda (2002), statistically insignificant coefficients on $SURPRISE_{t-j}$ $(j = 0, ..3)$ can be interpreted as evidence of the announcement effect.[1] This suggest that the Fed did not require temporary purchases to signal the change in the target, once it had been announced (or inferred by the markets).

The Tobit estimate will be inconsistent if the error terms are heteroskedastic or nonnormal. In fact, a Lagrange multiplier test for heteroskedasticity obtained by assuming $Var(\varepsilon_t|w_t) = \sigma^2 \exp(\delta' z_t)$, where z_t is a vector that contains all elements in w_t but the constant, rejects the null $H_0 : \delta = 0$ at the 1% level. In addition, the Jarque-Bera statistics leads us to reject the null that the residuals are normally distributed at a 1% level. Thus, we consider the Powell's (1994) CLAD estimator, which allows for both heteroskedasticity and nonnormality.

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

First, the CLAD estimates indicate that expected changes in the target were generally ac-

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

commodated on days other than the first Monday of the maintenance period, not only on the first Thursday as the Tobit estimates suggest. But, more importantly, a negative and statistically significant coefficient on $SURPRISE_{t-3}$ is consistent with the conventional liquidity effect. Furthermore, this liquidity effect is still present if we restrict the sample to the post August 18, 1998 period when the lagged reserve accounting was operational, and the federal funds target was being announced. Thus, is there an announcement effect? Our CLAD estimates, which are robust to nonnormality and very general forms of heteroskedasticity, suggest there is not.

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 and 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.

Bierens, H.J. (1981) *Robust methods and asymptotic theory in nonlinear econometrics.* New York: Springer-Verlag.

Bierens (1994) *Topics in Advanced Econometrics.* Cambridge: Cambridge University Press.

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`.

Buchinsky, M. (1994), Changes ins the U..S. Wage Structure 1963-1987: Application of

Table 1: Estimation results for Temporary Open Market Purchases

	Tobit		CLAD	
Variable	estimate	std.err.	estimate	std.err.
Constant	-18.742***	1.753	-0.504*	0.269
First Thursday	16.962***	2.833	1.896***	0.468
First Friday	-17.046***	2.917	-1.229*	0.709
Second Friday	-12.051***	2.683	-3.167***	0.996
Settlement Wednesday	4.456***	1.553	2.828***	0.431
TB(-1)	0.264***	0.035	0.047***	0.007
TB(-2)	0.292***	0.042	0.079***	0.007
TB(-3)	0.305***	0.049	0.069***	0.007
TS(-1)	-1.726***	0.611	-6.934***	1.381
TS(-2)	-0.865*	0.447	-1.058*	0.568
TS(-3)	-1.895***	0.408	-0.570*	0.273
PB(-1)	-0.018	0.085	-0.031	0.035
PB(-2)	-0.065	0.074	0.079***	0.013
PB(-3)	-0.073	0.072	-0.167***	0.049
PS(-1)	0.146	0.253	0.178***	0.045
PS(-2)	-0.151	0.245	0.032	0.045
PS(-3)	-0.198	0.223	0.039	0.043
NEED(-1)×Day1	-0.501	2.974	1.582***	0.503
NEED(-1)×Day2	11.645**	4.691	3.112***	0.727
NEED(-1)×Day3	24.030***	8.595	3.612*	1.976
NEED(-1)×Day4	-7.489	7.775	-1.509	1.745
NEED(-1)×Day5	21.671***	7.612	5.751***	1.848
NEED(-1)×Day6	5.312	10.941	-88.851***	3.165
NEED(-1)×Day7	33.429***	10.471	11.853***	2.347
NEED(-1)×Day8	6.842	7.789	8.894***	2.145
NEED(-1)×Day9	13.402***	4.953	4.557***	1.171
NEED(-1)×Day10	3.972*	2.083	1.132**	0.508
EXPECT(-1)×Day1	121.451**	53.231	38.392***	5.334
EXPECT(-1)×Day2	48.599	34.018	11.161*	6.105
EXPECT(-1)×Day3	-14.656	38.828	-27.103***	9.720
EXPECT(-1)×Day4	-25.528	29.346	4.957	5.303
EXPECT(-1)×Day5	-54.997*	29.859	4.929	5.134
EXPECT(-1)×Day6	51.573*	30.636	48.795***	5.201
EXPECT(-1)×Day7	-37.603	36.721	-16.383	16.266
EXPECT(-1)×Day8	39.313*	20.515	10.372**	5.052
EXPECT(-1)×Day9	-48.249*	24.666	-7.071	7.439
EXPECT(-1)×Day10	16.269	16.363	8.230	5.019
SURPRISE	-14.046	15.387	-0.647	3.815
SURPRISE(-1)	11.516	13.044	1.255	3.594
SURPRISE(-2)	-14.619	17.648	19.541***	5.149
SURPRISE(-3)	-7.259	14.832	-7.466**	3.311
SCALE	1094.099	104.478		

*,** and *** denote significance at the 10%, 5% and 1% levels respectively.

quantile regression, *Econometrica* 62, 405-458.

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

Cook, T. and T. Hahn (1988), The Information Content of Discount Rate Announcements and Their Effect on Market Interest Rates, *Journal of Money, Credit, and Banking* 20(2), 167-180.

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 (2003), Dynamic time series binary choice, mimeo, Ohio State University.

Demiralp, S. and O. Jorda (2002), The announcement effect: evidence from open market desk data, *Economic Policy Review*, 29-48.

Demiralp, S. and O. Jorda, The Response of Term Rates to Fed Announcements, *Journal of Money, Credit, and Banking*, forthcoming.

Dong, D., C. Chanjin, T. Schmit, and H.M. Kaiser (2000), Modeling the effects of advertising using a dynamic Tobit model, manuscript, 2000.

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.

Hansen, B. (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 nonparmetric kernel estimation, Working Paper, University of Wisconsin.

Kodres, L. E. (1988), Test of unbiasedness in the foreign exchange futures makets: 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.

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: Springer-Verlag, 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 of Finance* 18, 183-197.

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*, 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: North-Holland.

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, pages 111-143.

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 a well-defined 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 a well-defined 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*,*

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

Proof of Lemma 1:

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

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

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

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

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

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

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

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

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

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

$$\vdots$$

$$L^j_{p-1} = L^{j-1}_p + \max(0, \rho_{p-1}) L^{j-1}_1,$$

$$L^j_p = \max(0, \rho_p) L^{j-1}_1.$$

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

$$(1 - \sum_{j=1}^{p} \max(0, \rho_j) B^j) L^j_1 = 0.$$

From the fact that the above lag polynomial has all its roots outside the unit circle by assumption, it follows that $L^j_1 \leq c_1 \exp(-c_2 j)$ for positive constants c_1 and c_2. Also, if $\sup_{t\in\mathbb{Z}} E \max(0, \eta_t) < \infty$, then $\sum_{j=0}^{\infty} L^j_1 \max(0, \eta_{t-j})$ is a well-defined random variable. The above reasoning also applies to y_t (instead of $\hat{y}^m_t$). □

We are now in a position to prove Theorem 1.

Proof of Theorem 1:

The result of Theorem 1 follows by noting that, by Lemma 1,

$$\| y_t \|_r \leq \sum_{j=0}^{\infty} L^j_1 \| \max(0, \eta_{t-j}) \|_r < \infty.$$

□

The following lemma is needed for the stationarity proof of Theorem 2. 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

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

Lemma 2 *Assume that η_t is strictly stationary and strong mixing and satisfies $\| \max(0,\eta_t) \|_2 < \infty$. Then for all t, $\delta > 0$ and $c > 0$, 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 2:

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 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 3 *If*

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

for all $(y_1, \ldots, y_p) \in \mathbb{R}^p$*, then*

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

Proof of Lemma 3:

Note that for all $K \geq 0$,

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

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

Now defining

$$g_t^K = \eta_t + \sum_{k=1}^{K} \max(0, \eta_{t-k}) \sum_{i=1}^{p} \rho_i L_1^{k-i} I(i \leq k),$$

it follows from our assumption that $(g_{t-1}^{p-1}, \ldots, g_{t-p}^{0})$ satisfies, for some function $H(., \ldots, .)$,

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

Therefore,

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

$$= E \prod_{j=1}^{p} I(g_{t-j}^{p-j} \leq - \sum_{k=p-j+1}^{\infty} \max(0, \eta_{t-j-k}) \sum_{i=1}^{p} \rho_i L_1^{k-i} I(i \leq k) - \eta)$$

$$\geq EG(h_{t-1}, \ldots, h_{t-p})$$

where

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

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

Lemma 4 $\hat{y}_t^m \xrightarrow{as} y_t \quad as \quad m \to \infty.$

Proof of Lemma 4:

The result follows if we show that $\max_{k \geq m} |\hat{y}_t^k - y_t|$ converges to zero in probability as $m \to \infty$. Now, note that

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

so certainly,

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

and therefore $y_t = \hat{y}_t^k$ for all k such that $k \geq m > p$ if there can be found p consecutive "small" η_{t-l} that are negative and large in absolute value in the range $l = 1, \ldots, m$; 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 $\delta > 0$, $\eta > 0$, and $c > 0$,

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

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

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

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

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

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

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

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

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

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

for

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

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

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

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

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. This weak law of large numbers is proven in Lemma 2. Now if $EI_{tl}^{\zeta} > 0$, we can pick $\delta > 0$ small enough to satisfy the first requirement. 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 3. □

Proof of Theorem 2:

This result easily follows by noting 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. □

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

Lemma 5 *Assume that z_t is strictly stationary and L_2-near epoch dependent on a strictly stationary strong mixing process η_t, and assume that $q(z, b)$ is continuous on $\mathbb{R}^a \times \mathbb{R}^c$, where B is a compact subset of $\mathbb{R}^c$. Then if $E\sup_{b\in B} |q(z_t, b)| < \infty$,*

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

Proof of Lemma 5:

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

Lemma 6 *Under the conditions of Theorem 4,*

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

Proof of Lemma 6:

Note that by assumption, $E((\partial l_t(b)/\partial b)|_{b=\beta}|y_{t-1}, \ldots, x_t) = 0$ so that $E(\partial l_t(b)/\partial b)|_{b=\beta} = 0$, implying that $(\partial l_t(b)/\partial b)|_{b=\beta}$ is a martingale difference sequence. In particular, by noting that (y_t, x_t) has a "strong mixing base" in Bierens' (2004) terminology, asymptotic normality now follows from the version of Bierens (2004, Theorem 7.11) of a central limit theorem of McLeish (1974). Applying the information matrix equality yields the result. □

Proof of Theorem 4:

We prove Theorem 4 by checking the conditions of Newey and McFadden (1994, Theorem 3.1). Consistency was shown in Theorem 3. Condition (i) was assumed. Condition (ii), twice differentiability of the log likelihood, follows from the Tobit specification. Condition (iii) was shown in Lemma 6. Note that stationarity and the strong mixing base imply ergodicity. Condition (iv) follows from the result of Lemma 5, the Tobit specification, and the assumption of finite second moments for $|x_t|$ and ε_t. Condition (v) is assumed. □

Proof of Theorem 5:

Under Assumption 4, it follows from the discussion in Powell (1984, p. 318) that $S_T(b)$ is uniquely minimized at $\beta = (\rho', \gamma')'$. From the assumptions of Assumption 4 is follows that $E\sup_{b\in B}|s(y_{t-1},\ldots,y_{t-p},x_t,\varepsilon_t,b)| < \infty$, and therefore the uniform law of large numbers of Lemma 5 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 z_t, let

$$\lambda(b) = Eq(z_t, b).$$

Lemma 7 *Assume that $q(z_t, b)$ is continuous in both arguments and assume that $q(z, b)$ is continuous on $\mathbb{R}^a \times \mathbb{R}^c$, where B is a compact subset of $\mathbb{R}^c$. Assume that z_t is strictly stationary and that $|\hat{\beta}_T - \beta| = o_p(1)$. In addition assume that*

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

and assume that

$$T^{-1/2}\sum_{t=1}^{T}(q(z_t, b) - Eq(z_t, b))$$

is stochastically equicontinuous on B. Then

$$T^{1/2}\lambda(b)|_{b=\hat{\beta}_T} = -T^{-1/2}\sum_{t=1}^{T} q(z_t, \beta) + o_p(1).$$

Proof of Lemma 7:

This follows by writing

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

$$= T^{-1/2} \sum_{t=1}^{T} (q(z_t, \hat{\beta}_T) - Eq(z_t, \hat{\beta}_T) - q(z_t, \beta) + Eq(z_t, \beta))$$

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

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

Define that $z_t = (y_{t-1}, \ldots, y_{t-p}, x_1', \ldots, x_T')'$ and remember that $w_t = (y_{t-1}, \ldots, y_{t-p}, x_t')'$ and $b = (c', r', s)'$. To show the stochastic equicontinuity of $T^{-1/2} \sum_{t=1}^{T} (\psi(w_t, b) - E\psi(w_t, b))$ and thereby obtain our analogue of Powell's lemma A3, we first need the following results.

Lemma 8 *Assume $u_{1t} \in \mathbb{R}$ and $u_{2t} \in \mathbb{R}^q$, and assume that $(u_{1t}, u_{2t}')'$ is strictly stationary and L_r-near epoch dependent, $r \geq 2$, on η_t with an exponentially decreasing $\nu(\cdot)$ sequence. Then if $|u_{1t}| \leq 1$ and $\parallel u_{2t} \parallel_{r+\delta} < \infty$ for $\delta < r/(r-1)$, $u_{1t}u_{2t}$ is L_r-near epoch dependent on η_t with an exponentially decreasing $\nu(\cdot)$ sequence.*

Proof of Lemma 8:

This follows by noting that

$$\parallel u_{1t}u_{2t} - E(u_{1t}u_{2t} | \eta_t, \ldots, \eta_{t-m}) \parallel_r$$

$$\leq \parallel u_{1t}u_{2t} - E(u_{1t} | \eta_t, \ldots, \eta_{t-m}) E(u_{2t} | \eta_t, \ldots, \eta_{t-m}) \parallel_r$$

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

$$\leq \parallel u_{2t} - E(u_{2t} | \eta_t, \ldots, \eta_{t-m}) \parallel_r$$

$$+ \parallel E(u_{2t} | \eta_t, \ldots, \eta_{t-m}) \parallel_{r+\delta} \parallel u_{1t} - E(u_{1t} | \eta_t, \ldots, \eta_{t-m}) \parallel_{r(1+r/\delta)}$$

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

$$+ \| E(u_{2t}|\eta_t, \ldots, \eta_{t-m}) \|_{r+\delta} \| u_{1t} - E(u_{1t}|\eta_t, \ldots, \eta_{t-m}) \|_r^{r/(1+r/\delta)},$$

and all terms in the last expression decay exponentially with m. □

We also need the following result.

Lemma 9 *Under Assumption 5, for all $\eta > 0$, $I(b'w_t > 0)$, $I(b'w_t > -\eta|w_t|)$, $I(b'w_t \leq \eta|w_t|)$, and $I(\varepsilon_t + (\beta - b)'w_t > 0))$, $I(\varepsilon_t + (\beta - b)'w_t > -\eta|w_t|)$, $I(\varepsilon_t + (\beta - b)'w_t \leq \eta|w_t|)$ are L_r-near epoch dependent on η_t with an exponentially decreasing $\nu(\cdot)$ sequence that does not depend on b.*

Proof of Lemma 9:

We can limit attention to the case where $\eta \neq 0$. Note that w_t is near epoch dependent on $\eta_t = \gamma' x_t + \varepsilon_t$ with exponentially decreasing $\nu(\cdot)$ sequence by Theorem 2. In addition, for any $\delta > 0$, let $T_\delta(\cdot)$ be a continuously differentiable function such that $T_\delta(x) = I(x > 0)$ for $|x| > \delta$ and $\sup_{|x|\leq\delta} |(\partial/\partial x)T(x)| = K/\delta < \infty$. Then

$$\| I(b'w_t + \eta|w_t| > 0) - E(I(b'w_t + \eta|w_t| > 0)|\eta_t, \ldots, \eta_{t-m}) \|_r$$

$$\leq 2 \| I(b'w_t + \eta|w_t| > 0) - T_\delta(b'w_t + \eta|w_t|) \|_r$$

$$+ \| T_\delta(b'w_t + \eta|w_t| > 0) - E(T_\delta(b'w_t + \eta|w_t| > 0)|\eta_t, \ldots, \eta_{t-m}) \|_r$$

$$\leq 2 \| I(|b'w_t + \eta|w_t|| \leq \delta) \|_q + C_2\delta^{-1} \| b'w_t - E(b'w_t|\eta_t, \ldots, \eta_{t-m}) \|_r$$

$$+ C_2\delta^{-1} \| w_t - E(w_t|\eta_t, \ldots, \eta_{t-m}) \|_r$$

$$\leq C_1\delta + C_3\nu(m)\delta^{-1}$$

because for all b such that $r \neq 0$ and all $\eta > 0$, $b'w_t + \eta|w_t|$ has a uniformly bounded density because the density of $\varepsilon_t|w_t$ is uniformly bounded by assumption. Therefore by setting $\delta = \nu(m)^{1/2}$, it follows that $I(b'w_t + \eta|w_t| > 0)$ is near epoch dependent on η_t with exponentially decaying $\nu(\cdot)$ sequence as well.
For $I(\varepsilon_t + (\beta - b)'w_t + \eta|w_t| > 0))$, note that

$$E|I(\varepsilon_t + (\beta - b)'w_t + \eta|w_t| > 0) - E(I(\varepsilon_t + (\beta - b)'w_t + \eta|w_t| > 0)|\eta_t, \ldots, \eta_{t-m})|^r$$

$$\leq E|I(\varepsilon_t + (\beta - b)'w_t + \eta|w_t| > 0)$$

$$-I(\varepsilon_t + (\beta - b)'E(w_t|\eta_t, \ldots, \eta_{t-m}) + \eta|E(w_t|\eta_t, \ldots, \eta_{t-m})| > 0)|^r$$

$$= E|F((\beta - b)'w_t + \eta|w_t|) - F((\beta - b)'E(w_t|\eta_t, \ldots, \eta_{t-m}) + \eta|E(w_t|\eta_t, \ldots, \eta_{t-m})|)|^r$$

$$\leq L|(\beta - b)'(w_t - E(w_t|\eta_t, \ldots, \eta_{t-m})) + \eta(|w_t| - |E(w_t|\eta_t, \ldots, \eta_{t-m})|)|^r$$

$$\leq C(\sup_{b \in B}(\beta - b)'(\beta - b)^{r/2} + \eta^r)E|w_t - E(w_t|\eta_t, \ldots, \eta_{t-m})|^r,$$

and the last term decays exponentially with m. □

Lemma 10 *Under Assumption 5,* $T^{-1/2}\sum_{t=1}^{T}(\psi(w_t, b) - E\psi(w_t, b))$ *is stochastically equicontinuous on* B.

Proof of Lemma 10:

Note that for $\psi(w_t, b)$ we have for b such that $|b - \tilde{b}| < \eta$, because $\psi(.,.)$ is increasing in $w_t'b$,

$$|\psi(w_t, b) - \psi(w_t, \tilde{b})| \leq I(b'w_t > -\eta|w_t|)|1/2 - I(\varepsilon_t + (\beta - b)'w_t > -\eta|w_t|)||w_t|$$

and

$$E|I(b'w_t > -\eta|w_t|)(1/2 - I(\varepsilon_t + (\beta - b)'w_t > -\eta|w_t|))|w_t||$$

$$\leq EI(b'w_t > -\eta|w_t|)|w_t| \leq K_1\eta,$$

where the last result follows by Assumption 5.2. Therefore, we can cover B by $O(\eta^{\cdot\cdot})$ balls with center b_j and radius η and we can define the bracketing functions as

$$f_j^L(w_t) = I(b_j'w_t > 0)(1/2 - I(\varepsilon_t + (\beta - b_j)'w_t > 0))w_t$$

$$-I(b_j'w_t > -\eta|w_t|)|1/2 - I(\varepsilon_t + (\beta - b_j)'w_t > -\eta|w_t|)||w_t|$$

and

$$f_j^U(w_t) = I(b_j'w_t > 0)(1/2 - I(\varepsilon_t + (\beta - b_j)'w_t > 0))w_t$$

$$+I(b_j'w_t > -\eta|w_t|)|1/2 - I(\varepsilon_t + (\beta - b_j)'w_t > -\eta|w_t|)||w_t|.$$

By the result of Lemma 9, the bracketing functions $f_j^L(\cdot)$ and $f_j^U(\cdot)$ as well as the $\psi(w_t,\beta)$ are near epoch dependent on η_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 $\psi(m)=\nu(m)+\alpha(m)^{1/r-1/(2r)}$ and uniformly bounded mixingale numbers. We will now apply Theorem 3 of Hansen (1996); note that while Hansen's smoothness condition with respect to the parameter on the function class under consideration does not hold in our situation, his argument will still go through, because his cover number and weak dependence conditions hold in exactly the same way as for his proof. For Hansen's proof to work, we set Hansen's constants γ, q and s equal to 1/2, and our r and $2r$ respectively, and we note that the bracketing functions $f_j^L(\cdot)$ and $f_j^U(\cdot)$ as well as the $\psi(w_t,\beta)$ are also L_r-mixingales with mixingale numbers $\psi(m)=\nu(m)^{1/2}+\alpha(m)^{1/(2r)-1/(4r)}$ and mixingale numbers $c_t=\| f_j^L(w_t) \|_{2r}^{1/2}$, $c_t=\| f_j^L(w_t) \|_{2r}^{1/2}$, or $c_t=\| f_j^L(w_t) \|_{2r}^{1/2}$ respectively. The condition

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

now corresponds to Hansen's (1996) condition 2 of his Assumption 1, and Hansen's condition $q>a/(\lambda\gamma)$ now corresponds to, in our notation, $r>(p+q)/(1/2)=2(p+q)$. □

Proof of Theorem 6:

We follow the asymptotic normality proof of Powell (1984). The strategy of our proof is to replace Powell's Lemma A3 by the result of Lemma 7, and we note that under the conditions of Theorem 6, the stochastic equicontinuity condition of Lemma 7 follows from the result of Lemma 10. 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 4.2, 5.2, and 5.5, (i.e. the equivalents of Powell's E.1, R.1, and R.2),

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

By Lemma 7 and Lemma 10, for $\lambda(b)=E\psi(w_t,b)$,

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

It now follows from Lemma 8 and Lemma 9 that $\psi(z_t, b)$ is near epoch dependent on η_t with an exponentially decreasing $\nu(\cdot)$ sequence. Therefore by the central limit theorem of Theorem 2 of de Jong (1997), it follows that

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

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

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

and identically to the discussion in Powell (1984, p. 320-321, equations A.16-A.19), it now follows that, under Assumptions 4.2, 5.2, and 5.6, (i.e. the analogues of Powell's E.2, R.1, and R.2),

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

Therefore, it now follows that

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

as asserted by the theorem. □