TIME SERIES APPROACH TO THE EVOLUTION OF NETWORKS: PREDICTION AND ESTIMATION. SUPPLEMENTARY MATERIAL.

ANNA BYKHOVSKAYA

1. Truncated ordinary least squares estimator

As Example 2 suggests, the ordinary least squares (OLS) does not produce a consistent estimator. We show below the same inconsistency for generic values of parameters such that y_t is strongly mixing and converges to a stationary distribution, as in Theorems 2 and 3.

Lemma S1. *Suppose that $\mathbb{E}u_t = 0$. Then OLS is inconsistent.*

Proof. Define $\theta = (\alpha, \beta, \gamma)'$. If X is the matrix with rows $(1,\, y_{t-1},\, z_{t-1})$ and $Y = (y_1, \dots, y_T)'$, then the OLS estimator is

$$\begin{aligned}\hat{\theta}_{OLS} &= (X'X)^{-1}X'Y \\ &= \theta + (X'X)^{-1}X'U - (X'X)^{-1}\begin{pmatrix} \sum\limits_{t\colon y_t=0}(\alpha+\beta y_{t-1}+\gamma z_{t-1}+u_t) \\ \sum\limits_{t\colon y_t=0}(\alpha+\beta y_{t-1}+\gamma z_{t-1}+u_t)y_{t-1} \\ \sum\limits_{t\colon y_t=0}(\alpha+\beta y_{t-1}+\gamma z_{t-1}+u_t)z_{t-1}\end{pmatrix} \qquad (1)\end{aligned}$$

The term $(X'X)^{-1}X'U$ converges to zero as T goes to infinity by the law of large numbers, because $(1,\, y_{t-1},\, z_{t-1})$ is independent of u_t. However, the last term does not converge to zero:

$$\begin{aligned}&\frac{1}{T}\begin{pmatrix} \sum\limits_{t\colon y_t=0}(\alpha+\beta y_{t-1}+\gamma z_{t-1}+u_t) \\ \sum\limits_{t\colon y_t=0}(\alpha+\beta y_{t-1}+\gamma z_{t-1}+u_t)y_{t-1} \\ \sum\limits_{t\colon y_t=0}(\alpha+\beta y_{t-1}+\gamma z_{t-1}+u_t)z_{t-1}\end{pmatrix} \\ &\xrightarrow[T\to\infty]{\mathbb{P}} \begin{pmatrix} \mathbb{E}(\alpha+\beta y_{t-1}+\gamma z_{t-1}+u_t)\mathbf{1}(u_t < -\alpha-\beta y_{t-1}-\gamma z_{t-1}) \\ \mathbb{E}(\alpha+\beta y_{t-1}+\gamma z_{t-1}+u_t)y_{t-1}\mathbf{1}(u_t < -\alpha-\beta y_{t-1}-\gamma z_{t-1}) \\ \mathbb{E}(\alpha+\beta y_{t-1}+\gamma z_{t-1}+u_t)z_{t-1}\mathbf{1}(u_t < -\alpha-\beta y_{t-1}-\gamma z_{t-1})\end{pmatrix},\end{aligned}$$

where the expectations do not equal to zero. Thus, OLS is inconsistent. Moreover, each term in the expectation is negative when the indicator equals to 1. So OLS overestimates the coefficients as T goes to infinity. □

Date: December 27, 2019.

The advantage of the OLS procedure is the closed form for the estimator. We also recall that in the linear models the OLS estimator is more efficient than the LAD. These two properties motivate us to attempt to adjust the OLS procedure to restore consistency.

The idea of the modified procedure is the following: when y_{t-1} (z_{t-1}) is large, while z_{t-1} (y_{t-1}) is small, we can treat the constant and the second regressor as part of an error. Thus, we are left effectively with the classical autoregression model and can use standard theory. Mathematically, to estimate β, we need to condition on $T_{1M} = \{t | y_t > 0,\, y_{t-1} > M,\, z_{t-1} < M/h(M)\}$ for some number $M > 0$ and function $h(\cdot)$ such that $h(M) \xrightarrow[M\to\infty]{} \infty$. When M is large, $-\alpha - \beta y_{t-1} - \gamma z_{t-1}$ is very negative, so the indicator $\mathbf{1}(u_t < -\alpha - \beta y_{t-1} - \gamma z_{t-1})$ almost always equals zero, and the last term in Eq. (1) disappears as $T \to \infty$. Similarly, we can condition on $T_{2M} = \{t | y_t > 0,\, y_{t-1} < M/h(M),\, z_{t-1} > M\}$ to recover γ. The next theorem, which is proved in the Appendix, summarizes the above heuristics.

Theorem S2. *Separate OLS estimates of β and γ based on T_{1M} and T_{2M} are consistent, respectively, as $(M,T)_{seq} \to \infty$:*

$$\frac{\sum\limits_{T_{1M}} y_{t-1}y_t}{\sum\limits_{T_{1M}} y_{t-1}^2} \xrightarrow[(M,T)_{seq}\to\infty]{\mathbb{P}} \beta, \qquad \frac{\sum\limits_{T_{2M}} z_{t-1}y_t}{\sum\limits_{T_{2M}} z_{t-1}^2} \xrightarrow[(M,T)_{seq}\to\infty]{\mathbb{P}} \gamma.$$

After β and γ are estimated, one can estimate α using

$$\frac{1}{|T_{1M}| + |T_{2M}|} \sum_{T_{1M}\cup T_{2M}} (y_t - \beta y_{t-1} - \gamma z_{t-1}) \xrightarrow[(M,T)_{seq}\to\infty]{\mathbb{P}} \alpha. \tag{2}$$

Proof. Define $\hat{\beta} = \frac{\sum_{T_{1M}} y_{t-1}y_t}{\sum_{T_{1M}} y_{t-1}^2}$ and $\hat{\gamma} = \frac{\sum_{T_{1M}} z_{t-1}y_t}{\sum_{T_{1M}} z_{t-1}^2}$. Let us show that $\hat{\beta} \xrightarrow[(M,T)_{seq}\to\infty]{\mathbb{P}} \beta$. The proof for $\hat{\gamma}$ is the same.

$$\hat{\beta} = \frac{\sum\limits_{T_{1M}} y_{t-1}y_t}{\sum\limits_{T_{1M}} y_{t-1}^2} = \beta + \frac{\sum\limits_{T_{1M}} (\alpha + \gamma z_{t-1} + u_t) y_{t-1}}{\sum\limits_{T_{1M}} y_{t-1}^2}$$
$$\xrightarrow[T\to\infty]{\mathbb{P}} \beta + \frac{\alpha \mathbb{E}(y_{t-1}|T_{1M}) + \gamma \mathbb{E}(y_{t-1}z_{t-1}|T_{1M}) + \mathbb{E}(u_t y_{t-1}|T_{1M})}{\mathbb{E}(y_{t-1}^2|T_{1M})}.$$

First note that u_t and y_{t-1} are independent and $\mathbb{E}(u_t|T_{1M}) \xrightarrow[M\to\infty]{} 0$. Then $\frac{\mathbb{E}(y_{t-1}|T_{1M})}{\mathbb{E}(y_{t-1}^2|T_{1M})} \leq \frac{\mathbb{E}(y_{t-1}|T_{1M})}{\mathbb{E}(My_{t-1}|T_{1M})} = \frac{1}{M} \xrightarrow[M\to\infty]{} 0$. Finally, by Cauchy–Schwarz inequality, $\mathbb{E}(y_{t-1}z_{t-1}|T_{1M}) \leq \sqrt{\mathbb{E}(y_{t-1}^2|T_{1M})\mathbb{E}(z_{t-1}^2|T_{1M})}$ so that $\frac{\mathbb{E}(y_{t-1}z_{t-1}|T_{1M})}{\mathbb{E}(y_{t-1}^2|T_{1M})} \leq \sqrt{\frac{\mathbb{E}(z_{t-1}^2|T_{1M})}{\mathbb{E}(y_{t-1}^2|T_{1M})}} \leq \sqrt{\frac{M^2/h^2(M)}{M^2}} = \frac{1}{h(M)} \xrightarrow[M\to\infty]{} 0$.

Thus, $\hat{\beta} \xrightarrow[(M,T)_{seq}\to\infty]{\mathbb{P}} \beta$.

Finally, notice that both under T_{1M} and T_{2M},

$$\alpha + u_t = y_t - \beta y_{t-1} - \gamma z_{t-1},$$

so that

$$\frac{1}{|T_{1M}| + |T_{2M}|} \sum_{T_{1M} \cup T_{2M}} (y_t - \beta y_{t-1} - \gamma z_{t-1}) = \alpha + \frac{1}{|T_{1M}| + |T_{2M}|} \sum_{T_{1M} \cup T_{2M}} u_t \xrightarrow[T\to\infty]{\mathbb{P}} \alpha + \mathbb{E}(u|T_{1M} \cup T_{2M})$$

and

$$\alpha + \mathbb{E}(u|T_{1M} \cup T_{2M}) \xrightarrow[M\to\infty]{} \alpha + \mathbb{E}u = \alpha. \quad \square$$

In practice, to estimate α we need to use a different, smaller threshold. That is, we first estimate β and γ based on some M_1 and then we plug the estimates into Eq. (2), evaluated at $M_2 < M_1$, to estimate α.

Let us note, that in practice we can use a simpler procedure. We denote it as OLS_M. One can condition on $T_M := \{t | y_t > 0,\, y_{t-1} > M\}$ for some $M > 0$ and run OLS with three regressors. The problem here is that the limit behavior of the inverse of conditional matrix $\begin{pmatrix} 1 & \mathbb{E}(y|T_M) & \mathbb{E}(z|T_M) \\ \mathbb{E}(y|T_M) & \mathbb{E}(y^2|T_M) & \mathbb{E}(yz|T_M) \\ \mathbb{E}(z|T_M) & \mathbb{E}(yz|T_M) & \mathbb{E}(z^2|T_M) \end{pmatrix}^{-1}$ is unclear. It may crucially depend on the properties of the error distribution. As long as post-multiplication by the vector of cross product covariances $(0, Cov(y_{t-1}, u_t|T_M), Cov(z_{t-1}, u_t|T_M))'$ results in the zero vector in the limit, the sequential limit of the corresponding OLS estimate equals (α, β, γ). That is, the inverse matrix must not explode faster than the conditional covariance vector goes to zero. This is summarized in the next theorem.

Theorem S3. *The sequential limit $(M, T)_{seq} \to \infty$ of the OLS estimator based on $t \in T_M$ equals the true value (α, β, γ) when the product*

$$\begin{pmatrix} 1 & \mathbb{E}(y_{t-1}|T_M) & \mathbb{E}(z_{t-1}|T_M) \\ \mathbb{E}(y_{t-1}|T_M) & \mathbb{E}(y_{t-1}^2|T_M) & \mathbb{E}(y_{t-1}z_{t-1}|T_M) \\ \mathbb{E}(z_{t-1}|T_M) & \mathbb{E}(y_{t-1}z_{t-1}|T_M) & \mathbb{E}(z_{t-1}^2|T_M) \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ Cov(y_{t-1}, u_t|T_M) \\ Cov(z_{t-1}, u_t|T_M) \end{pmatrix}$$

converges to zero as $M \to \infty$.

Proof. Conditional on T_M, the OLS estimate equals to

$$\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} + \begin{pmatrix} \sum_{T_M} 1 & \sum_{T_M} y_{t-1} & \sum_{T_M} z_{t-1} \\ \sum_{T_M} y_{t-1} & \sum_{T_M} y_{t-1}^2 & \sum_{T_M} y_{t-1}z_{t-1} \\ \sum_{T_M} z_{t-1} & \sum_{T_M} y_{t-1}z_{t-1} & \sum_{T_M} z_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{T_M} u_t \\ \sum_{T_M} y_{t-1}u_t \\ \sum_{T_M} z_{t-1}u_t \end{pmatrix}$$

$$\xrightarrow[T\to\infty]{\mathbb{P}} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} + \begin{pmatrix} 1 & \mathbb{E}(y_{t-1}|T_M) & \mathbb{E}(z_{t-1}|T_M) \\ \mathbb{E}(y_{t-1}|T_M) & \mathbb{E}(y_{t-1}^2|T_M) & \mathbb{E}(y_{t-1}z_{t-1}|T_M) \\ \mathbb{E}(z_{t-1}|T_M) & \mathbb{E}(y_{t-1}z_{t-1}|T_M) & \mathbb{E}(z_{t-1}^2|T_M) \end{pmatrix}^{-1} \begin{pmatrix} \mathbb{E}u_t \\ \mathbb{E}(y_{t-1}u_t|T_M) \\ \mathbb{E}(z_{t-1}u_t|T_M) \end{pmatrix}.$$

Let us rewrite the second term. The goal is to show that it converges to zero as $M \to \infty$.

$$\begin{pmatrix} 1 & \mathbb{E}(y_{t-1}|T_M) & \mathbb{E}(z_{t-1}|T_M) \\ \mathbb{E}(y_{t-1}|T_M) & \mathbb{E}(y_{t-1}^2|T_M) & \mathbb{E}(y_{t-1}z_{t-1}|T_M) \\ \mathbb{E}(z_{t-1}|T_M) & \mathbb{E}(y_{t-1}z_{t-1}|T_M) & \mathbb{E}(z_{t-1}^2|T_M) \end{pmatrix}^{-1} \begin{pmatrix} \mathbb{E}u_t \\ \mathbb{E}(y_{t-1}u_t|T_M) \\ \mathbb{E}(z_{t-1}u_t|T_M) \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \mathbb{E}(y|T_M) & \mathbb{E}(z|T_M) \\ \mathbb{E}(y|T_M) & \mathbb{E}(y^2|T_M) & \mathbb{E}(yz|T_M) \\ \mathbb{E}(z|T_M) & \mathbb{E}(yz|T_M) & \mathbb{E}(z^2|T_M) \end{pmatrix}^{-1}$$

$$\cdot \left(\mathbb{E}(u|T_M) \begin{pmatrix} 1 \\ \mathbb{E}(y|T_M) \\ \mathbb{E}(z|T_M) \end{pmatrix} + \begin{pmatrix} 0 \\ \mathbb{E}(yu|T_M) - \mathbb{E}(y|T_M)\mathbb{E}(u|T_M) \\ \mathbb{E}(zu|T_M) - \mathbb{E}(z|T_M)\mathbb{E}(u|T_M) \end{pmatrix} \right)$$

$$= \mathbb{E}(u|T_M) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & \mathbb{E}(y|T_M) & \mathbb{E}(z|T_M) \\ \mathbb{E}(y|T_M) & \mathbb{E}(y^2|T_M) & \mathbb{E}(yz|T_M) \\ \mathbb{E}(z|T_M) & \mathbb{E}(yz|T_M) & \mathbb{E}(z^2|T_M) \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ Cov(y,u|T_M) \\ Cov(z,u|T_M) \end{pmatrix}.$$

Because $\mathbb{E}(u|T_M) \to \mathbb{E}u = 0$ as $M \to \infty$, the first term converges to zero. By assumption, the second term also converges to zero. (Note that as $M \to \infty$ the correlation between u_t and y_{t-1} drops to zero, so that $Cov(y_{t-1}, u_t|T_M) \to 0$. Similarly, $Cov(z_{t-1}, u_t|T_M) \to 0$. However, the behavior of the inverse matrix per se is unclear.)

Thus, sequential limit of the OLS estimate based on T_M equals the true values of the parameters. □

In simulations, the product of the inverse conditional matrix of second moments and the conditional covariance vector goes to zero. Thus, in the empirical example studied below we use the above procedure to calculate adjusted OLS estimates. Moreover, as the next theorem suggests, when there are no peer effects ($\gamma \equiv 0$) and both u_t and y_t have exponential tails, the product goes to zero. When there is no γ, the product of the inverse conditional matrix of second moments and the conditional covariance vector reduces to

$$\begin{pmatrix} 1 & \mathbb{E}(y_{t-1}|T_M) \\ \mathbb{E}(y_{t-1}|T_M) & \mathbb{E}(y_{t-1}^2|T_M) \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ Cov(y_{t-1}, u_t|T_M) \end{pmatrix}$$
$$= \frac{1}{\mathbb{V}(y_{t-1}|T_M)} \begin{pmatrix} -\mathbb{E}(y_{t-1}|T_M)Cov(y_{t-1}, u_t|T_M) \\ Cov(y_{t-1}, u_t|T_M) \end{pmatrix}$$

Theorem S4. *Assume that the stationary distribution of y_t has density $f_y(x)$ for large positive x and that the noise u_t has density $f_u(x)$ for large negative x. Further, assume that there exist six positive constants $c_1, c_2, c_3, d_1, d_2, d_3 > 0$, such that for all large enough positive x:*

$$f_y(x) = \exp(-g_y(x)), \text{ where } c_1 x^{d_1} \le g_y'(x) \le c_2 x^{d_2} \tag{3}$$

and for all large enough negative x:

$$f_u(x) = \exp(-g_u(x)), \text{ where } g_u(x) \ge c_3|x|^{d_3}. \tag{4}$$

Then the vector $\frac{1}{\mathbb{V}(y_{t-1}|T_M)} \begin{pmatrix} -\mathbb{E}(y_{t-1}|T_M)Cov(y_{t-1}, u_t|T_M) \\ Cov(y_{t-1}, u_t|T_M) \end{pmatrix}$ goes to 0 as $M \to \infty$, i.e. the OLS_M estimate is consistent as $M \to \infty$.

Proof. Because $g_y'(x) \ge c_1 x^{d_1}$, for any $x \ge M$ $g_y'(x) \ge c_1 M^{d_1}$ and

$$\begin{aligned} \mathbb{P}(y_{t-1} > M) &= \int_M^\infty f_y(x)dx = \int_M^\infty e^{-g_y(x)}dx = e^{-g_y(M)} \int_M^\infty e^{-\int_M^x g_y'(w)dw} dx \\ &\le e^{-g_y(M)} \int_M^\infty e^{-c_1 M^{d_1}(x-M)}dx \le e^{-g_y(M)} \frac{1}{c_1 M^{d_1}} \le e^{-g_y(M)}, \end{aligned} \tag{5}$$

where the last inequality holds for M large enough.

We want to show that conditional variance of y_{t-1} is polynomial in M^{-1}. To do this, let us show that if a variance of a random variable X is bounded from below by $C > 0$ on some interval $[a, b]$, then $\mathbb{V}X \ge \frac{C}{8}(b-a)^3$:

$$\mathbb{V}X = \int_{\mathbb{R}} (x - \mathbb{E}X)^2 f_X(x)dx \ge C \int_a^b (x - \mathbb{E}X)^2 dx \ge C \int_0^{\frac{b-a}{2}} x^2 dx = \frac{C}{24}(b-a)^3, \tag{6}$$

where the last inequality holds because if $a + \frac{b-a}{2} \ge \mathbb{E}X$ then $(x - \mathbb{E}X)^2 \ge \left(x - a - \frac{b-a}{2}\right)^2$ for $x \in [a + (b-a)/2, b]$ and if $a + \frac{b-a}{2} \le \mathbb{E}X$ then $(x - \mathbb{E}X)^2 \ge (x-a)^2$ for $x \in [a, a + (b-a)/2]$.

Consider the interval $\Delta = [M, M + (c_2 M^{d_2})^{-1}]$. Because $f_y(x) = e^{-g_y(x)}$ and $g_y' > 0$, the density of y is decreasing for $x \in \Delta$ for M large enough. Thus, for $x \in \Delta$,

$$\begin{aligned} f_y(x) &\geq f_y(M + (c_2 M^{d_2})^{-1}) = \exp(-g_y(M + (c_2 M^{d_2})^{-1})) \\ &\geq \exp(-g_y(M) - g_y'(M + (c_2 M^{d_2})^{-1})(c_2 M^{d_2})^{-1}) \\ &\geq \exp(-g_y(M)) \exp(-c_2 (M + (c_2 M^{d_2})^{-1})^{d_2} (c_2 M^{d_2})^{-1}) \\ &= \exp(-g_y(M)) \exp(-(1 + (c_2 M^{d_2+1})^{-1})^{d_2}) \geq \exp(-g_y(M)) e^{-2^{d_2}}. \end{aligned} \tag{7}$$

Therefore, combining Eq. (5) and Eq. (7), for $x \in \Delta$,

$$f_{y_{t-1}|T_M}(x) = f_y(x)/\mathbb{P}(y_{t-1} > M) \geq e^{-g_y(M)} e^{-2^{d_2}} / e^{-g_y(M)} = e^{-2^{d_2}}.$$

Using the bound from Eq. (6), we get

$$\mathbb{V}(y_{t-1}|T_M) = \int (x - \mathbb{E}(y_{t-1}|T_M))^2 f_y(x) dx \geq \frac{1}{24 e^{2^{d_2}}} (c_2 M^{d_2})^{-3}. \tag{8}$$

Let us show that the conditional expectation of y_{t-1} does not grow faster than linearly in M.

(9)

$$\mathbb{E}(y_{t-1}|T_M) = \int_M^\infty x \frac{f_y(x)}{\mathbb{P}(y_{t-1} > M)} dx = \frac{\int_M^\infty x e^{-g_y(x)} dx}{\int_M^\infty e^{-g_y(x)} dx} \leq \frac{2 \int_M^{2M} x e^{-g_y(x)} dx}{\int_M^\infty e^{-g_y(x)} dx} \leq 4M \frac{\int_M^{2M} e^{-g_y(x)} dx}{\int_M^\infty e^{-g_y(x)} dx} \leq 4M,$$

where the first inequality comes from the fact that $x e^{-g_y(x)}$ is decreasing exponentially, so that for M large enough $\int_M^{2M} x e^{-g_y(x)} dx > \int_{2M}^\infty x e^{-g_y(x)} dx$.

We are left with analyzing conditional covariance between y_{t-1} and u_t.

$$\begin{aligned} &Cov(y_{t-1}, u_t|T_M) = \mathbb{E}\left(y_{t-1} \mathbb{E}\left(u_t - \mathbb{E}(u_t|T_M)|y_{t-1}\right)|T_M\right) \\ &= \int_M^\infty x \int_{-\alpha-\beta x}^\infty v \frac{f_u(v)}{\mathbb{P}(u_t > -\alpha - \beta x)} dv \frac{f_y(x)}{\mathbb{P}(y_{t-1} > M)} dx - \mathbb{E}(y_{t-1}|T_M)\mathbb{E}(u_t|T_M). \end{aligned} \tag{10}$$

First, note that, for $x \geq M$,

(11)

$$\mathbb{P}(u_t > -\alpha - \beta x) \geq \mathbb{P}(u_t > -\alpha - \beta M) = 1 - \int_{-\infty}^{-\alpha-\beta M} f_u(v) dv \geq 1 - \int_{-\infty}^{-\alpha-\beta M} e^{-c_3 |v|^{d_3}} dv \xrightarrow[M \to \infty]{} 1,$$

so that $\mathbb{P}(u_t > -\alpha - \beta x) \geq 0.5$ for M large enough.

Second, because $\mathbb{E}u = 0$,

$$\int_{-\alpha-\beta x}^{\infty} v f_u(v)dv = -\int_{-\infty}^{-\alpha-\beta x} v f_u(v)dv = \int_{-\infty}^{-\alpha-\beta x} (-v)e^{-g_u(v)}dv \leq \int_{-\infty}^{-\alpha-\beta x} (-v)e^{-c_3|v|^{d_3}}dv$$
$$\leq 2(\alpha+\beta x)e^{-c_3(\alpha+\beta x)^{d_3}}, \tag{12}$$

where the last inequality holds for M large enough as the integrand decreases exponentially.

Third, using Eq. (12),

$$\begin{aligned}\mathbb{E}(u_t|T_M) &= \int_M^{\infty}\int_{-\alpha-\beta x}^{\infty} u\frac{f_u(v)}{\mathbb{P}(u_t > -\alpha-\beta x)}dv\frac{f_y(x)}{\mathbb{P}(y_{t-1} > M)}dx \\ &\leq \int_M^{\infty} 2(\alpha+\beta x)e^{-c_3(\alpha+\beta x)^{d_3}}\frac{f_y(x)}{\mathbb{P}(u_t > -\alpha-\beta x)\mathbb{P}(y_{t-1} > M)}dx \\ &\leq \int_M^{\infty} 4(\alpha+\beta x)e^{-c_3(\alpha+\beta x)^{d_3}}\frac{f_y(x)}{\mathbb{P}(y_{t-1} > M)}dx \\ &= 4\mathbb{E}\left((\alpha+\beta y_{t-1})e^{-c_3(\alpha+\beta y_{t-1})^{d_3}}|T_M\right) \leq 4(\alpha+\beta M)e^{-c_3(\alpha+\beta M)^{d_3}},\end{aligned} \tag{13}$$

because the function under expectation is decreasing in y for M large enough.

Plugging Eq. (9), (12), and (13) into Eq. (10), we get

$$\begin{aligned}&|Cov(y_{t-1}, u_t|T_M)| \\ &\leq 4\int_M^{\infty} x(\alpha+\beta x)e^{-c_3(\alpha+\beta x)^{d_3}}\frac{f_y(x)dx}{\mathbb{P}(y_{t-1} > M)} + 16M(\alpha+\beta M)e^{-c_3(\alpha+\beta M)^{d_3}} \\ &\leq 4M(\alpha+\beta M)e^{-c_3(\alpha+\beta M)^{d_3}} + 16M(\alpha+\beta M)e^{-c_3(\alpha+\beta M)^{d_3}}.\end{aligned} \tag{14}$$

Combining Eq. (8) and (14), we get

$$\begin{aligned}&\frac{|Cov(y_{t-1}, u_t|T_M)|}{\mathbb{V}(y_{t-1}|T_M)} \\ &\leq 24e^{2^{d_2}}\frac{4M(\alpha+\beta M)e^{-c_3(\alpha+\beta M)^{d_3}} + 16M(\alpha+\beta M)e^{-c_3(\alpha+\beta M)^{d_3}}}{(c_2M^{d_2})^{-3}} \xrightarrow[M\to\infty]{} 0.\end{aligned}$$

Combining Eq. (8), (9) and (14), we get

$$\begin{aligned}&\frac{\mathbb{E}(y_{t-1}|T_M)|Cov(y_{t-1}, u_t|T_M)|}{\mathbb{V}(y_{t-1}|T_M)} \\ &\leq 96e^{2^{d_2}}M\frac{4M(\alpha+\beta M)e^{-c_3(\alpha+\beta M)^{d_3}} + 16M(\alpha+\beta M)e^{-c_3(\alpha+\beta M)^{d_3}}}{(c_2M^{d_2})^{-3}} \xrightarrow[M\to\infty]{} 0.\end{aligned}$$

So that $\frac{1}{\mathbb{V}(y_{t-1}|T_M)}\begin{pmatrix} -\mathbb{E}(y_{t-1}|T_M)Cov(y_{t-1},u_t|T_M) \\ Cov(y_{t-1},u_t|T_M) \end{pmatrix} \xrightarrow[M\to\infty]{} 0.$ □

Remark S1. The conditions (3), (4) mean that both the noise and the stationary distribution have light tails. The condition (3) additionally requires that the tail probability of the stationary distribution of y_t does not decay too fast. These conditions are not intended to be optimal and can likely be considerably weakened. Instead, they are intended to illustrate the type of conditions where Theorem 3 holds.

The disadvantage of the adjusted OLS procedures is that we have to discard a lot of observations. Moreover, it is unclear how to choose M and $h(M)$. The tradeoff is that the larger is M, the more observations we have to discard, yet the closer to the consistent limit we are. Thus, we see that as we restore the consistency by increasing M, we lose the efficiency of the estimator.

2. Maximum likelihood estimator

Suppose that we know the density, f_u, of the error u_t. Then we can calculate the likelihood. It will consist of two types of terms. The first type corresponds to the cases when y_t is non-zero, the positive part is non-binding, so we can write $y_t = \alpha + \beta y_{t-1} + \gamma z_{t-1} + u_t$ or $u_t = y_t - \alpha - \beta y_{t-1} - \gamma z_{t-1}$. The second type corresponds to time periods with $y_t = 0$. If y_t is zero, then it is equivalent to $\alpha + \beta y_{t-1} + \gamma z_{t-1} + u_t$ being non-positive. That is, $y_t = 0$ is equivalent to $u_t \le -\alpha - \beta y_{t-1} - \gamma z_{t-1}$. Thus, the likelihood and its logarithm are

$$\begin{aligned} \mathcal{L} &= \prod_{t:\, y_t>0} f_u(y_t - \alpha - \beta y_{t-1} - \gamma z_{t-1}) \times \prod_{t:\, y_t=0} F_u(-\alpha - \beta y_{t-1} - \gamma z_{t-1}), \\ \log \mathcal{L} &= \sum_{t:\, y_t>0} \log f_u(y_t - \alpha - \beta y_{t-1} - \gamma z_{t-1}) + \sum_{t:\, y_t=0} \log F_u(-\alpha - \beta y_{t-1} - \gamma z_{t-1}). \end{aligned} \tag{15}$$

Following common practice, we assume a normal distribution for u_t. It turns our, as the Theorem 5 shows, that when the true distribution is normal, the MLE produces consistent estimators. However, as the simulations suggest, and in agreement with the well-known results in the i.i.d. censored regression model, when the true distribution is far from normal, the estimates are poor. Moreover, numerical optimization is very sensitive to the choice of the initial point and the calculations for the MLE sometimes explode.

The proof of Theorem 5, which is shown below, uses extremum estimation techniques. In a similar spirit it is possible to show $\sqrt{T}$- asymptotic normality of the MLE estimator under Gaussian errors.

Theorem S5. *If $u_t \sim i.i.d.\,\mathcal{N}(0,\sigma^2)$, then MLE is consistent.*

Proof. Define $\theta = (\alpha, \beta, \gamma, \sigma)$ and assume that θ_0 is the true value of θ. MLE estimate $\hat{\theta}$ maximizes sample log-likelihood, Q_n. The sample and population log-likelihoods are

$$\begin{aligned} Q_n(\theta) &= \frac{1}{T}\sum_{t=1}^{T}\Big[\log f_y(y_t|y_{t-1}, z_{t-1}, \theta)\mathbf{1}(y_t > 0) + \log F_y(0|y_{t-1}, z_{t-1}, \theta)\mathbf{1}(y_t = 0)\Big] \\ &= \frac{1}{T}\sum_{t=1}^{T}\Big[\log f_u(y_t - \alpha - \beta y_{t-1} - \gamma z_{t-1}|y_{t-1}, z_{t-1}, \theta)\mathbf{1}(y_t > 0) \\ &+ \log F_u(-\alpha - \beta y_{t-1} - \gamma z_{t-1}|y_{t-1}, z_{t-1}, \theta)\mathbf{1}(y_t = 0)\Big] \end{aligned} \tag{16}$$

and

$$\begin{aligned} Q(\theta) &= \mathbb{E}\Big[\log f_y(y_t|y_{t-1}, z_{t-1}, \theta)\mathbf{1}(y_t > 0) + \log F_y(0|y_{t-1}, z_{t-1}, \theta)\mathbf{1}(y_t = 0)\Big] \\ &= \mathbb{E}\Big[\log f_u(y_t - \alpha - \beta y_{t-1} - \gamma z_{t-1}|y_{t-1}, z_{t-1}, \theta)\mathbf{1}(y_t > 0) \\ &+ \log F_u(-\alpha - \beta y_{t-1} - \gamma z_{t-1}|y_{t-1}, z_{t-1}, \theta)\mathbf{1}(y_t = 0)\Big], \end{aligned} \tag{17}$$

where expectation is taken with respect to y_t, y_{t-1}, z_{t-1}.

Let us first show that θ_0 uniquely minimizes Q.

$$\begin{aligned} Q(\theta) - Q(\theta_0) &= \mathbb{E}(\log f_y(y_t|y_{t-1}, z_{t-1}, \theta) - \log f_y(y_t|y_{t-1}, z_{t-1}, \theta_0))\mathbf{1}(y_t > 0) \\ &+ \mathbb{E}(\log F_y(0|y_{t-1}, z_{t-1}, \theta) - \log F_y(0|y_{t-1}, z_{t-1}, \theta_0))\mathbf{1}(y_t = 0). \end{aligned} \tag{18}$$

First note, that

$$\begin{aligned} &\mathbb{E}\mathbf{1}(y_t = 0)\log F_y(0|y_{t-1}, z_{t-1}, \theta) = \mathbb{E}\log F_y(0|y_{t-1}, z_{t-1}, \theta)\left(\mathbb{E}\left(\mathbf{1}(y_t = 0)|y_{t-1}, z_{t-1}\right)\right) \\ &= \mathbb{E}\log F_y(0|y_{t-1}, z_{t-1}, \theta)\mathbb{P}_y(0|y_{t-1}, z_{t-1}) = \mathbb{E}\log F_y(0|y_{t-1}, z_{t-1}, \theta)F_y(0|y_{t-1}, z_{t-1}, \theta_0). \end{aligned}$$

Then, because $\log x \le x - 1$,

(19)

$$\begin{aligned} &\mathbb{E}(\log F_y(0|y_{t-1}, z_{t-1}, \theta) - \log F_y(0|y_{t-1}, z_{t-1}, \theta_0))\mathbf{1}(y_t = 0) \\ &= \mathbb{E}\log\left(\frac{F_y(0|y_{t-1}, z_{t-1}, \theta)}{F_y(0|y_{t-1}, z_{t-1}, \theta_0)}\right)F_y(0|y_{t-1}, z_{t-1}) \le \mathbb{E}\left(F_y(0|y_{t-1}, z_{t-1}, \theta) - F_y(0|y_{t-1}, z_{t-1}, \theta_0)\right). \end{aligned}$$

Similarly,

$$
\begin{aligned}
&\mathbb{E}(\log f_y(y_t|y_{t-1},z_{t-1},\theta)-\log f_y(y_t|y_{t-1},z_{t-1},\theta_0))\mathbf{1}(y_t>0)\\
&=\mathbb{E}\log\left(\frac{f_y(y_t|y_{t-1},z_{t-1},\theta)}{f_y(y_t|y_{t-1},z_{t-1},\theta_0)}\right)\mathbf{1}(y_t>0)\le\mathbb{E}\left(\frac{f_y(y_t|y_{t-1},z_{t-1},\theta)}{f_y(y_t|y_{t-1},z_{t-1},\theta_0)}-1\right)\mathbf{1}(y_t>0)\\
&=\mathbb{E}\left(\mathbb{E}\left(\left(\frac{f_y(y_t|y_{t-1},z_{t-1},\theta)}{f_y(y_t|y_{t-1},z_{t-1},\theta_0)}-1\right)\mathbf{1}(y_t>0)|y_{t-1},z_{t-1}\right)\right)\\
&=\mathbb{E}\int\left(f_y(y_t|y_{t-1},z_{t-1},\theta)-f_y(y_t|y_{t-1},z_{t-1},\theta_0)\right)\mathbf{1}(y_t>0)dy_t\\
&=(1-\mathbb{E}F_y(0|y_{t-1},z_{t-1},\theta))-(1-\mathbb{E}F_y(0|y_{t-1},z_{t-1},\theta_0))\\
&=\mathbb{E}\left(F_y(0|y_{t-1},z_{t-1},\theta_0)-F_y(0|y_{t-1},z_{t-1},\theta)\right)
\end{aligned}
\tag{20}
$$

Plugging Eq. (19) and (20) into Eq. (18), we get

$$
\begin{aligned}
Q(\theta)-Q(\theta_0)\le\,&\mathbb{E}\left(F_y(0|y_{t-1},z_{t-1},\theta)-F_y(0|y_{t-1},z_{t-1},\theta_0)\right)\\
&+\mathbb{E}\left(F_y(0|y_{t-1},z_{t-1},\theta_0)-F_y(0|y_{t-1},z_{t-1},\theta)\right)=0.
\end{aligned}
$$

Thus, θ_0 minimizes Q. Moreover, equality holds only when $\mathbb{P}\left(f_y(y_t|y_{t-1},z_{t-1},\theta)=f_y(y_t|y_{t-1},z_{t-1},\theta_0)\right)=1$, which can not happen for gaussian errors with density $f_y(y_t|y_{t-1},z_{t-1},\theta)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(y_t-\alpha-\beta y_{t-1}-\gamma z_{t-1})^2\right)$.

To apply the theorem for extremum estimators, we need to reduce the domain of θ to a compact space. That is, we need to show that when some of the parameters go to infinity, Q_n goes to minus infinity and, thus, such values can not be solutions to $\max Q_n$. Here we are going to use the fact that $f_u(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{x^2/2\sigma^2}$. Let us plug the density into Eq. (16):

$$
\begin{aligned}
Q_n=\,&\frac{1}{T}\sum_{t=1}^{T}\left(-0.5\log(2\pi\sigma^2)-\frac{1}{2\sigma^2}(y_t-\alpha-\beta y_{t-1}-\gamma z_{t-1})^2\right)\mathbf{1}(y_t>0)\\
&+\frac{1}{T}\sum_{t=1}^{T}\left(-0.5\log(2\pi\sigma^2)+\log\int\limits_{-\infty}^{-\alpha-\beta y_{t-1}-\gamma z_{t-1}}e^{-u^2/2\sigma^2}du\right)\mathbf{1}(y_t=0)
\end{aligned}
\tag{21}
$$

If σ goes to infinity, then $-0.5\log(2\pi\sigma^2)\to-\infty$, while other terms remain non-positive: $\int\limits_{-\infty}^{A}e^{-u^2/2\sigma^2}du\le\sqrt{2\pi}\sigma$. Thus, $Q_n\to-\infty$ when $\sigma\to\infty$ independently of the values of other parameters, and we can restrict σ to a bounded interval.

After we know that σ is bounded, we can guarantee that the second summation is bounded by zero from above for any values of α,β,γ: $\left(-0.5\log(2\pi\sigma^2)+\log\int\limits_{-\infty}^{-\alpha-\beta y_{t-1}-\gamma z_{t-1}}e^{-u^2/2\sigma^2}du\right)\le-0.5\log(2\pi\sigma^2)+0.5\log(2\pi\sigma^2)=0$. When $|\alpha|$ goes to infinity or $|\beta|\to\infty$ or $|\gamma|\to\infty$, we have $(y_t-\alpha-\beta y_{t-1}-\gamma z_{t-1})^2\to\infty$.

Note that as y and z are random with correlation below one, parameters can not compensate each other. Thus, $Q_n \to -\infty$ as $|\alpha| \to \infty$ or $\beta \to \infty$ or $|\gamma| \to \infty$. Therefore, those parameters can also be restricted to bounded intervals. Thus, we are left with compact set.

Plugging the density of u into Eq. (17), we get

$$\begin{aligned} Q(\theta) = \mathbb{E}\left(-0.5\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(y_t - \alpha - \beta y_{t-1} - \gamma z_{t-1})^2\right)\mathbf{1}(y_t > 0) \\ + \mathbb{E}\left(-0.5\log(2\pi\sigma^2) + \log\int_{-\infty}^{-\alpha-\beta y_{t-1}-\gamma z_{t-1}} e^{-u^2/2\sigma^2}du\right)\mathbf{1}(y_t = 0). \end{aligned} \tag{22}$$

Function under expectation in Eq. (22),

$$\begin{aligned} g(y_t, y_{t-1}, z_{t-1}, \theta) := \left(-0.5\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(y_t - \alpha - \beta y_{t-1} - \gamma z_{t-1})^2\right)\mathbf{1}(y_t > 0) \\ + \left(-0.5\log(2\pi\sigma^2) + \log\int_{-\infty}^{-\alpha-\beta y_{t-1}-\gamma z_{t-1}} e^{-u^2/2\sigma^2}du\right)\mathbf{1}(y_t = 0), \end{aligned}$$

is continuous at every θ with probability 1 and, because parameters are restricted to a compact set, $\mathbb{E}\sup_\theta |g(y_t, y_{t-1}, z_{t-1}, \theta)|$ is finite.

Finally, we can apply Proposition 7.3 (Consistency of M-estimators with compact parameter space) from **?**. Our model satisfies all the conditions of the proposition. Thus, the MLE estimate $\hat{\theta}$ is consistent. □

3. Omitted proofs

3.1. Finite time until the network is empty.

Lemma S6. *If Assumptions* 1, 2, *and* 5 *are satisfied,* $\mathbb{E}u_{ijt}^4 < \infty$ *for all* i, j, t, *and for all* i, j $\max(0, \beta_{ij}) + |\gamma_{ij}| < C < 1$, *then* $\mathbb{E}($*time until graph is empty for* H *periods*$)$ *is finite. That is, the expected time until* $y_{ijt} = \ldots = y_{ij,t+H-1} = 0$ *for all* i, j *is finite.*

Proof. Denote by $\bar{u}_t = \{u_{ijt}\}_{i,j}$ the vector of all errors at time t. Fix some number $M > 0$ (it will be specified later) and define three independent random variables

$$\bar{u}_t^- = \{\{u_{ijt}\}_{i,j} | u_{ijt} < -M \;\forall i, j\},$$

$$\bar{u}_t^+ = \{\{u_{ijt}\}_{i,j} | u_{i'j't} \geq -M \text{ for some } i', j'\},$$

$$\xi_t = \begin{cases} 0, \text{with probability } \mathbb{P}(\forall i, j \;\; u_{ijt} < -M), \\ 1, \text{with probability } \mathbb{P}(\exists i, j \text{ s.t. } u_{ijt} \geq -M). \end{cases}$$

Then

$$\bar{u}_t \overset{d}{=} \xi_t \bar{u}_t^+ + (1 - \xi_t)\bar{u}_t^-. \tag{23}$$

Fix realizations of $(\bar{u}_t^+, \bar{u}_t^-, \xi_t)$ for $t = 1, \dots, T$ and calculate the corresponding $\bar{u}_t$ from Eq. (23). Define

$$v_t = |\mathcal{A}| + \max_{\substack{i,j \\ s=t,\dots,t-H+1}} \left[u_{ijs}^+ + \alpha_{ij}\right]_+ \geq 0.$$

Now construct a new time series

$$y_t' = C \max_{s=t-1,\dots,t-H} y_s' + v_t,$$

$$y_p' = \max_{i,j} y_{ijp} \text{ for } p = 0, \dots, H-1.$$

One can easily show by induction that $y_t' \geq y_{ijt}$ for all i, j, t.

By the same argument as in proof of Theorem 1, we can divide time periods into blocks of length H and get a bound

$$y_{t+p}' \leq C \max_{s=t-1,\dots,t-H} y_s' + \sum_{s=0}^{H-1} v_{t+s} \text{ for all } p = 0, \dots, H-1. \tag{24}$$

Now define another random process and error, $x_\tau = \max_{s=(\tau-1)H,\dots,\tau H-1} y_s'$, $w_\tau = \sum_{s=0}^{H-1} v_{(\tau-1)H+s}$. Then by Eq. (24),

$$x_{\tau+1} \leq C x_\tau + w_{\tau+1}.$$

We need to find M such that for some $\varepsilon > 0$, $\mathbb{P}(\sharp\{\tau \in [1, \dots, \lfloor T/H \rfloor] | x_\tau < M\} \geq \varepsilon T) \geq \frac{1}{T^2}$ and $\mathbb{P}(\forall i, j \;\; u_{ijt} < -M) > 0$. This is a condition on u_{ijt} which generally may fail to be true. For example, if u_{ijt} are almost surely larger than some positive constant. Let us show that such M exists under assumptions 2 and $\mathbb{E} u_{ijt} < C_4 < \infty$. Assumption 2 implies that for all M large enough $\mathbb{P}(\forall i, j \;\; u_{ijt} < -M) > 0$.

Let us show that if $\mathbb{E} u_{ijt} < C_4 < \infty$, then $\mathbb{E} w_\tau < \tilde{C}_4 < \infty$, and constant $\tilde{C}_4$ does not depend on M. Note that the fourth moment of u_{ijt}^+ is bounded as

$$\mathbb{E}|u_{ijt}^+|^4 \leq \mathbb{E}|u_{ijt}^4| \frac{1}{\mathbb{P}(\exists i, j \text{ s.t. } u_{ijt} \geq -M)}.$$

Further, as $|w_\tau| = \sum_{s=0}^{H-1} v_{(\tau-1)H+s}$, it is less than the sum of absolute values of several instances of u_{ijt}^+ and constants. Thus, the fourth moment of the sum can be bounded by a combination of the individual fourth moments, which are bounded. As $M \to \infty$, $\mathbb{P}(\exists i, j \text{ s.t. } u_{ijt} \geq -M) \to 1$, so that setting a lower bound for M to be such that $\mathbb{P}(\exists i, j \text{ s.t. } u_{ijt} \geq -M') = 0.5$, we get a bound which does not depend on $M > M'$ and then $\mathbb{E}|u_{ijt}^+|^4 \leq 2\mathbb{E}|u_{ijt}^4|$.

Define one more process, $\tilde{x}_\tau$, by

$$\tilde{x}_{\tau+1} = C \tilde{x}_\tau + w_{\tau+1}, \; \tilde{x}_1 = x_1.$$

It can be shown by induction, that for all τ, $\tilde{x}_\tau \geq x_\tau$. Thus, it is enough to show that $\mathbb{P}(\sharp\{\tau \in [1, \dots, T/H] | \tilde{x}_\tau < M\} \geq \varepsilon T) \geq \frac{1}{T^2}$. Let us show that $\exists Q$ such that $\tilde{x}_1 + \tilde{x}_2 + \dots +$

$\tilde{x}_{\lfloor T/H \rfloor} < Q\lfloor T/H \rfloor$ with probability greater that $1 - \frac{const}{T^2}$.

$$\tilde{x}_1 + \tilde{x}_2 + \ldots + \tilde{x}_{\lfloor T/H \rfloor} = \tilde{x}_1 + \sum_{\tau=2}^{\lfloor T/H \rfloor} (w_\tau + C w_{\tau-1} + \ldots + C^{\tau-2} w_2 + C^{\tau-1} \tilde{x}_1)$$
$$\leq \frac{1}{1-C} \sum_{\tau=1}^{\lfloor T/H \rfloor} w_\tau + \frac{\tilde{x}_1}{1-C}.$$

The expectation of the right hand side of the last expression is $\frac{1}{1-C} \left(\lfloor T/H \rfloor \mathbb{E} w_\tau + \mathbb{E} x_1 \right)$

$$\mathbb{P}\left(\left| \frac{1}{1-C} \sum_{\tau=1}^{\lfloor T/H \rfloor} (w_\tau - \mathbb{E} w_\tau) \right| > \lfloor T/H \rfloor \right) \leq \frac{\mathbb{E} \left| \sum_{\tau=1}^{\lfloor T/H \rfloor} (w_\tau - \mathbb{E} w_\tau) \right|^4}{(1-C)^4 \lfloor T/H \rfloor^4} \leq \frac{const \cdot T^2}{T^4} \leq \frac{const}{T^2},$$

where we used the fact that $w_\tau - \mathbb{E} w_\tau$ are i.i.d. with zero mean and with bounded fourth and second moments. Thus,

$$\mathbb{P}\left(\frac{1}{1-C} \sum_{\tau=1}^{\lfloor T/H \rfloor} w_\tau > Q\lfloor T/H \rfloor \right) \leq \mathbb{P}\left(\left| \frac{1}{1-C} \sum_{\tau=1}^{\lfloor T/H \rfloor} (w_\tau - \mathbb{E} w_\tau) \right| > \lfloor T/H \rfloor \right) \leq \frac{const}{T^2},$$

where $Q = 1 + \frac{\mathbb{E} w_\tau}{1-C}$. Thus,

(25)
$$\mathbb{P}\left(\tilde{x}_1 + \tilde{x}_2 + \ldots + \tilde{x}_{\lfloor T/H \rfloor} < Q\lfloor T/H \rfloor \right) \geq \mathbb{P}\left(\frac{1}{1-C} \sum_{\tau=1}^{\lfloor T/H \rfloor} w_\tau + \frac{\tilde{x}_1}{1-C} < Q\lfloor T/H \rfloor \right) \geq 1 - \frac{const}{T^2}.$$

Finally, if $\tilde{x}_1 + \tilde{x}_2 + \ldots + \tilde{x}_{\lfloor T/H \rfloor} < Q\lfloor T/H \rfloor$, then $x_1 + \ldots + x_{\lfloor T/H \rfloor} < Q\lfloor T/H \rfloor$. The latter implies that $\sharp\{\tau | x_\tau > 2Q\} < 0.5\lfloor T/H \rfloor$ and $\sharp\{\tau | x_\tau \leq 2Q\} > 0.5\lfloor T/H \rfloor$. That is, we have shown that $\exists M$ (any number larger than $2Q$ and M') such that $\sharp\{\tau | x_\tau \leq M\} > \varepsilon T$ has probability greater than $1 - \frac{const}{T^2}$. For each such τ we flip a coin to determine ξ_t. If it zero, then the whole process y_{ijt} jumps to zero. Thus, with probability of at most $(\mathbb{P}(\exists i, j \text{ s.t. } u_{ijt} \geq -M))^{\varepsilon T}$ the process does not jump to zero. Thus,

$$\mathbb{E}(\text{length until } H \text{ zero periods}) = \sum_{T=1}^{\infty} \mathbb{P}(length \geq T)$$
$$\leq \sum_{T=1}^{\infty} \left(\frac{const}{T^2} + (\mathbb{P}(\exists i, j \text{ s.t. } u_{ijt} \geq -M))^{\varepsilon T} \right) < \infty. \quad \square$$

Corollary S7. *If Assumptions* 1 *and* 2 *are satisfied for the model without* γ *and if* $\beta < 1$, *then* $\mathbb{E}$(*length until zero*) *is finite.*

Proof. If $\beta < 1$ and there is no γ, $\max(0, \beta) + |\gamma| = \max(0, \beta) < 1$. Thus, Lemma 6 applies. $\square$

Lemma S8. *If Assumptions* 1 *and* 2 *are satisfied,* $\alpha < 0$, $\beta = 1$ *and if* $\mathbb{E}u_t^4 < \infty$, *then* $\mathbb{E}$(*length until zero*) *is finite.*

Proof. We can write the expected length until $y_t = 0$ as

$$\mathbb{E}(\text{length until zero}) = \sum_{T=1}^{\infty} \mathbb{P}(\textit{length} \geq T). \tag{26}$$

Define $S_t = y_0 + t\alpha + u_1 + \ldots + u_t$ for all $t \in \mathbb{N}$. If length until zero is greater than T, then $S_1 > 0, \ldots, S_{T-1} > 0$. (Otherwise the process S_t becomes negative, so that non-negative process y_t becomes zero before T). Thus, $\mathbb{P}(\textit{length} \geq T) \leq \mathbb{P}(S_1 > 0, \ldots, S_{T-1} > 0)$. Note that

$$\mathbb{P}(S_1 > 0, \ldots, S_{T-1} > 0) = \mathbb{P}\left(y_0 + \alpha + u_1 > 0, \ldots, y_0 + (T-1)\alpha + \sum_{t=1}^{T-1} u_t > 0\right)$$
$$\leq \mathbb{P}\left(y_0 + (T-1)\alpha + \sum_{t=1}^{T-1} u_t > 0\right) = \mathbb{P}\left(\sum_{t=1}^{T-1} u_t > -y_0 - (T-1)\alpha\right).$$

Because $\alpha < 0$, there exists T' such that $\forall T > T'$ $-y_0 - (T-1)\alpha > 0$. Let us look at any $T > T'$.

$$\mathbb{P}\left(\sum_{t=1}^{T-1} u_t > -y_0 - (T-1)\alpha\right) \leq \mathbb{P}\left(\left|\sum_{t=1}^{T-1} u_t\right| > -y_0 - (T-1)\alpha\right) \leq \frac{\mathbb{E}\left|\sum_{t=1}^{T-1} u_t\right|^4}{(y_0 + (T-1)\alpha)^4} \tag{27}$$
$$= \frac{(T-1)\mathbb{E}u^4 + 3(T-1)(T-2)\mathbb{E}u^2}{(y_0 + (T-1)\alpha)^4} \leq \frac{\textit{const}}{T^2},$$

where we used Markov inequality to bound probability by expectation.

Plugging Eq. (27) into Eq. (26), we get

$$\mathbb{E}(\text{length until zero}) \leq \sum_{T=1}^{\infty} \frac{\textit{const}}{T^2} < \textit{const}_1 < \infty. \qquad \square$$

3.2. Explosive LAD.

Lemma S9. *Consider the process* $y_{t+1} = [\alpha + \beta y_t + u_t]_+$, $\beta > 1$. *For any* $\beta' \in (1, \beta)$, *almost surely exists* T *such that* $y_{t+1} > \beta' y_t$ *for all* $t > T$.

Proof. Fix $\varepsilon > 0$. Denote $v_t = |\alpha| + |u_t|$. Write $y_{t+1} \geq \beta y_t - v_t$. Iterating, we get

$$y_{t+k} \geq \beta^k \left(y_t - \sum_{i=1}^{k} v_{t+i-1}\beta^{-i}\right) \tag{28}$$

Further, note $\sum_{i=1}^{\infty} v_{t+i-1}\beta^{-i}$ is a positive finite random variable, whose distribution does not depend on the choice of t. Choose large $M > 2$ such that this random variable is less than M with probability greater than $1 - \varepsilon$. By Classification Theorem (Theorem 3) we already know that almost surely $y_t \xrightarrow[t\to\infty]{} \infty$. Thus, we can choose T such that $y_T > 2M$ with probability greater than $1-\varepsilon$. Then with probability greater than $1-2\varepsilon$, we have by (28):

$$y_{T+k} \geq \beta^k (y_T/2), \quad \text{for all } k = 1, 2, \ldots. \tag{29}$$

Let us call the event where (29) holds $\mathcal{A}_T$. We thus know that $\mathbb{P}(\mathcal{A}_T) \geq 1 - 2\varepsilon$ for large enough T.

Next, consider the events

$$\mathcal{B}_k = \{|v_{T+k}| > \beta^k\}.$$

Note that $\sum_k \mathbb{P}(\mathcal{B}_k) < \infty$, since v_t is a random variable, whose distribution does not depend on t and whose expectation exists. Therefore, there exists K such that for the event $\mathcal{C}_K = \{v_{T+k} \leq \beta^k \text{ for all } k > K\}$, $\mathbb{P}(\mathcal{C}_K) \geq 1 - \sum_{k=K}^{\infty} \mathbb{P}(\mathcal{B}_k) > 1 - \varepsilon$.

Now consider the event $\mathcal{D} = \mathcal{A}_T \cap \mathcal{C}_K$. We have $\mathbb{P}(\mathcal{D}) \geq 1 - (\mathbb{P}(\neg \mathcal{A}_T) + \mathbb{P}(\neg \mathcal{C}_K)) > 1 - 3\varepsilon$. On the other hand, on this event, for each $t > T + K$, we have

$$y_{t+1} \geq \beta y_t - v_t = \beta' y_t + (\beta - \beta') y_t \left(1 - \frac{v_t}{y_t}\right) \tag{30}$$

Since $y_t \geq \frac{y_T}{2} \beta^{t-T} > M\beta^{t-T} > 2\beta^{t-T}$ and $v_t \leq \beta^{t-T}$, the last term in (30) is positive and we conclude that $y_{t+1} \geq \beta' y_t$, as desired.

Since $\varepsilon > 0$ was arbitrary, we conclude that with probability 1 for all large enough t, $y_{t+1} \geq \beta' y_t$, as desired. □

(Anna Bykhovskaya) University of Wisconsin-Madison
E-mail address: anna.bykhovskaya@wisc.edu