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

ANNA BYKHOVSKAYA

Contents

Appendix A. Game Theoretical Model

In this section we present a stylized game theoretical model that leads to our equation of interest, Eq. (3). It is another justification of Eq. (3) along with the primary one, which is to capture a number of essential properties of networks (non-negativity of edges, positive probability of vanishing of each edge, and interactions between edges, which affect the whole network).

Consider a world with n myopic agents (people/firms/countries/etc.) with quadratic adjustment costs. Agents can interact with each other over time. Time is discrete and goes from 1 to T. Every period, each agent i chooses how much time to spend with or how much to trade with each other agent j. The decision is based on two components: costs and benefits.

Benefits are characterized by a per unit gain of $\alpha_{ij} + u_{ijt}$. Here α_{ij} is a constant, while u_{ijt} is random component that is independent across time. Thus, y units of communication/trade leads to a benefit of $y(\alpha_{ij} + u_{ijt})$.

The second component is a quadratic adjustment cost function. Agents get disutility whenever there are deviations from some target expected level of communication/trade.

Date: August 2, 2021.

The target is composed from an own past and a peer-effect or interactions component. The interaction term aggregates the whole structure of the network for up to H periods. That is, we assume that agent i by choosing to devote y units to agent j has to pay

$$\frac{1}{2}\left(y-\beta_{ij}y_{ijt-1}-\gamma_{ij}p_{ij}\left(\left\{y_{kls}\right\}_{\substack{k,l=1,\dots,n\\ s=t-H,\dots,t-1}}\right)\right)^2,$$

where

$$p_{ij}\left(\left\{y_{kls}\right\}_{\substack{k,l=1,\dots,n\\ s=t-H,\dots,t-1}}\right):\mathbb{R}_+^{NH}\to\mathbb{R}$$

represents peer effects/interactions function, which depends on H previous periods.

The interpretation is that β_{ij} represents a rate at which stock/relationship depreciates/appreciates. If $\beta_{ij}<1$, then the agent is introverted and tends to decrease communication; while if $\beta_{ij}>1$, the agent is an extrovert, who tends to expand communication. The interpretation of β_{ij} for firms corresponds to production depreciation (i.g. technology wears out) or production appreciation (better technology management over time increases production). The coefficient γ_{ij} indexes the sensitivity of the reference level with respect to peer effects/interactions, which are, in turn, represented by the function p_{ij}. The peer effects function depends on H past periods of the whole network, and captures interactions between different edges y_{kls} across time.

Agents have separate maximization problems for each time period t and with each peer j. Agent i solves the following maximization problem at day t with respect to agent j:

$$\max_{y\geq 0}\left[y(\alpha_{ij}+u_{ijt})-\frac{1}{2}\left(y-\beta_{ij}y_{ijt-1}-\gamma_{ij}p_{ij}\left(\left\{y_{kls}\right\}_{\substack{k,l=1,\dots,n\\ s=t-H,\dots,t-1}}\right)\right)^2\right]. \tag{A.1}$$

The solution to the maximization problem (A.1) is

$$y^*_{ijt}=\left[\alpha_{ij}+\beta_{ij}y_{ijt-1}+\gamma_{ij}p_{ij}\left(\left\{y_{kls}\right\}_{\substack{k,l=1,\dots,n\\ s=t-H,\dots,t-1}}\right)+u_{ijt}\right]_+,$$

which leads to the network evolution process described by Eq. (3).

Appendix B. 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 B.1. *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

$$\hat{\theta}_{OLS} = (X'X)^{-1}X'Y$$

$$= \theta + (X'X)^{-1}X'U - (X'X)^{-1} \begin{pmatrix} \sum_{t:\, y_t=0} (\alpha + \beta y_{t-1} + \gamma z_{t-1} + u_t) \\ \sum_{t:\, y_t=0} (\alpha + \beta y_{t-1} + \gamma z_{t-1} + u_t) y_{t-1} \\ \sum_{t:\, y_t=0} (\alpha + \beta y_{t-1} + \gamma z_{t-1} + u_t) z_{t-1} \end{pmatrix} \tag{B.1}$$

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:

$$\frac{1}{T} \begin{pmatrix} \sum_{t:\, y_t=0} (\alpha + \beta y_{t-1} + \gamma z_{t-1} + u_t) \\ \sum_{t:\, y_t=0} (\alpha + \beta y_{t-1} + \gamma z_{t-1} + u_t) y_{t-1} \\ \sum_{t:\, y_t=0} (\alpha + \beta y_{t-1} + \gamma z_{t-1} + u_t) z_{t-1} \end{pmatrix}$$

$$\xrightarrow[T\to\infty]{\mathbb{P}} \begin{pmatrix} \mathbb{E}(\alpha + \beta y_{t-1} + \gamma z_{t-1} + u_t)\mathbf{1}(u_t < -\alpha - \beta y_{t-1} - \gamma z_{t-1}) \\ \mathbb{E}(\alpha + \beta y_{t-1} + \gamma z_{t-1} + u_t) y_{t-1}\mathbf{1}(u_t < -\alpha - \beta y_{t-1} - \gamma z_{t-1}) \\ \mathbb{E}(\alpha + \beta y_{t-1} + \gamma z_{t-1} + u_t) z_{t-1}\mathbf{1}(u_t < -\alpha - \beta y_{t-1} - \gamma z_{t-1}) \end{pmatrix},$$

where the expectations do not equal to zero. Thus, OLS is inconsistent. □

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. We do this in the "non-negative" setting: $\beta > 0, \gamma > 0, z_t \geq 0$.

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 $\beta > 0$, 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. (B.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 $\gamma > 0$. The next theorem summarizes the above heuristics.

Theorem B.2. *Suppose that $\beta > 0, \gamma > 0$, $z_t \geq 0$ for all t, and $\mathbb{E}u_t = 0$. Then 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

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

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

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

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. (B.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 $\mathrm{OLS_M}$. 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 go to infinity faster than the conditional covariance vector goes to zero. This is summarized in the next theorem.

Theorem B.3. *Suppose that $\beta > 0, \gamma > 0$, $z_t \geq 0$ for all t, and $\mathbb{E}u_t = 0$. Then 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. Moreover, as the next theorem suggests, when $\beta > 0$ and 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 B.4. *Assume that $\beta > 0$ and 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:*

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

and for all large enough negative x:

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

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 $g_y'(x) \geq c_1 x^{d_1}$, for any $x \geq M$ $g_y'(x) \geq c_1 M^{d_1}$ and

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

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 $\mathbb{V}X \geq \frac{C}{8}(b-a)^3$:

$$\mathbb{V}X = \int\limits_{\mathbb{R}} (x - \mathbb{E}X)^2 f_X(x)dx \geq C \int\limits_a^b (x - \mathbb{E}X)^2 dx \geq C \int\limits_0^{\frac{b-a}{2}} x^2 dx = \frac{C}{24}(b-a)^3, \tag{B.6}$$

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

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

Therefore, combining Eq. (B.5) and Eq. (B.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. (B.6), we get

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

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

(B.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 $xe^{-g_y(x)}$ is decreasing exponentially, so that for M large enough $\int_M^{2M} xe^{-g_y(x)}dx > \int_{2M}^\infty xe^{-g_y(x)}dx$.

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

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

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

(B.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 $\mathbb{E}u = 0$,

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

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

Third, using Eq. (B.12),

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

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

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

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

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

$$\begin{aligned}&\frac{|Cov(y_{t-1}, u_t|T_M)|}{\mathbb{V}(y_{t-1}|T_M)} \\ &\le 24e^{2^{d_2}} \frac{4M(\alpha+\beta M) e^{-c_3(\alpha+\beta M)^{d_3}} + 16M(\alpha+\beta M) e^{-c_3(\alpha+\beta M)^{d_3}}}{(c_2 M^{d_2})^{-3}} \xrightarrow[M\to\infty]{} 0.\end{aligned}$$

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

$$\begin{aligned}&\frac{\mathbb{E}(y_{t-1}|T_M)|Cov(y_{t-1}, u_t|T_M)|}{\mathbb{V}(y_{t-1}|T_M)} \\ &\le 96e^{2^{d_2}} M \frac{4M(\alpha+\beta M) e^{-c_3(\alpha+\beta M)^{d_3}} + 16M(\alpha+\beta M) e^{-c_3(\alpha+\beta M)^{d_3}}}{(c_2 M^{d_2})^{-3}} \xrightarrow[M\to\infty]{} 0.\end{aligned}$$

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 B.1. The conditions (B.3), (B.4) mean that both the noise and the stationary distribution have light tails. The condition (B.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 B.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.

Appendix C. 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 $u_t \leq -\alpha - \beta y_{t-1} - \gamma z_{t-1}$. Thus, the likelihood and its logarithm are

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

Following common practice, we assume a normal distribution for u_t. It turns our, as the Theorem C.1 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 C.1, 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 C.1. *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

$$\begin{aligned}Q_n(\theta) &= \frac{1}{T}\sum_{t=1}^{T}\Big[\log f_y(y_t|y_{t-1}, z_{t-1}, \theta)\mathbf{1}(y_t > 0) + \log F_y(0|y_{t-1}, z_{t-1}, \theta)\mathbf{1}(y_t = 0)\Big]\\ &= \frac{1}{T}\sum_{t=1}^{T}\Big[\log f_u(y_t - \alpha - \beta y_{t-1} - \gamma z_{t-1}|y_{t-1}, z_{t-1}, \theta)\mathbf{1}(y_t > 0)\\ &\quad + \log F_u(-\alpha - \beta y_{t-1} - \gamma z_{t-1}|y_{t-1}, z_{t-1}, \theta)\mathbf{1}(y_t = 0)\Big]\end{aligned} \tag{C.2}$$

and

$$
\begin{aligned}
Q(\theta) &= \mathbb{E}\Big[\log f_y(y_t|y_{t-1}, z_{t-1}, \theta)\mathbf{1}(y_t > 0) + \log F_y(0|y_{t-1}, z_{t-1}, \theta)\mathbf{1}(y_t = 0)\Big] \\
&= \mathbb{E}\Big[\log f_u(y_t - \alpha - \beta y_{t-1} - \gamma z_{t-1}|y_{t-1}, z_{t-1}, \theta)\mathbf{1}(y_t > 0) \\
&\quad + \log F_u(-\alpha - \beta y_{t-1} - \gamma z_{t-1}|y_{t-1}, z_{t-1}, \theta)\mathbf{1}(y_t = 0)\Big],
\end{aligned} \tag{C.3}
$$

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.

$$
\begin{aligned}
Q(\theta) - Q(\theta_0) = {}& \mathbb{E}(\log f_y(y_t|y_{t-1}, z_{t-1}, \theta) - \log f_y(y_t|y_{t-1}, z_{t-1}, \theta_0))\mathbf{1}(y_t > 0) \\
& + \mathbb{E}(\log F_y(0|y_{t-1}, z_{t-1}, \theta) - \log F_y(0|y_{t-1}, z_{t-1}, \theta_0))\mathbf{1}(y_t = 0).
\end{aligned} \tag{C.4}
$$

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 $\log x \leq x - 1$,

$$
\begin{aligned}
&\mathbb{E}(\log F_y(0|y_{t-1}, z_{t-1}, \theta) - \log F_y(0|y_{t-1}, z_{t-1}, \theta_0))\mathbf{1}(y_t = 0) \\
&= \mathbb{E}\log\left(\frac{F_y(0|y_{t-1}, z_{t-1}, \theta)}{F_y(0|y_{t-1}, z_{t-1}, \theta_0)}\right)F_y(0|y_{t-1}, z_{t-1}) \leq \mathbb{E}\left(F_y(0|y_{t-1}, z_{t-1}, \theta) - F_y(0|y_{t-1}, z_{t-1}, \theta_0)\right).
\end{aligned} \tag{C.5}
$$

Similarly,

$$
\begin{aligned}
&\mathbb{E}(\log f_y(y_t|y_{t-1}, z_{t-1}, \theta) - \log f_y(y_t|y_{t-1}, z_{t-1}, \theta_0))\mathbf{1}(y_t > 0) \\
&= \mathbb{E}\log\left(\frac{f_y(y_t|y_{t-1}, z_{t-1}, \theta)}{f_y(y_t|y_{t-1}, z_{t-1}, \theta_0)}\right)\mathbf{1}(y_t > 0) \leq \mathbb{E}\left(\frac{f_y(y_t|y_{t-1}, z_{t-1}, \theta)}{f_y(y_t|y_{t-1}, z_{t-1}, \theta_0)} - 1\right)\mathbf{1}(y_t > 0) \\
&= \mathbb{E}\left(\mathbb{E}\left(\left(\frac{f_y(y_t|y_{t-1}, z_{t-1}, \theta)}{f_y(y_t|y_{t-1}, z_{t-1}, \theta_0)} - 1\right)\mathbf{1}(y_t > 0)|y_{t-1}, z_{t-1}\right)\right) \\
&= \mathbb{E}\int\left(f_y(y_t|y_{t-1}, z_{t-1}, \theta) - f_y(y_t|y_{t-1}, z_{t-1}, \theta_0)\right)\mathbf{1}(y_t > 0)dy_t \\
&= (1 - \mathbb{E}F_y(0|y_{t-1}, z_{t-1}, \theta)) - (1 - \mathbb{E}F_y(0|y_{t-1}, z_{t-1}, \theta_0)) \\
&= \mathbb{E}\left(F_y(0|y_{t-1}, z_{t-1}, \theta_0) - F_y(0|y_{t-1}, z_{t-1}, \theta)\right)
\end{aligned} \tag{C.6}
$$

Plugging Eq. (C.5) and (C.6) into Eq. (C.4), we get

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

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. (C.2):

$$\begin{aligned} Q_n &= \frac{1}{T}\sum_{t=1}^{T}\left(-0.5\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(y_t - \alpha - \beta y_{t-1} - \gamma z_{t-1})^2\right)\mathbf{1}(y_t > 0) \\ &+ \frac{1}{T}\sum_{t=1}^{T}\left(-0.5\log(2\pi\sigma^2) + \log\int_{-\infty}^{-\alpha-\beta y_{t-1}-\gamma z_{t-1}} e^{-u^2/2\sigma^2}du\right)\mathbf{1}(y_t = 0) \end{aligned} \tag{C.7}$$

If σ goes to infinity, then $-0.5\log(2\pi\sigma^2) \to -\infty$, while other terms remain non-positive: $\int_{-\infty}^{A} e^{-u^2/2\sigma^2}du \leq \sqrt{2\pi\sigma}$. 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: $\left(-0.5\log(2\pi\sigma^2) + \log\int_{-\infty}^{-\alpha-\beta y_{t-1}-\gamma z_{t-1}} e^{-u^2/2\sigma^2}du\right) \leq -0.5\log(2\pi\sigma^2) + 0.5\log(2\pi\sigma^2) = 0$. 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. (C.3), we get

$$\begin{aligned} Q(\theta) &= \mathbb{E}\left(-0.5\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(y_t - \alpha - \beta y_{t-1} - \gamma z_{t-1})^2\right)\mathbf{1}(y_t > 0) \\ &+ \mathbb{E}\left(-0.5\log(2\pi\sigma^2) + \log\int_{-\infty}^{-\alpha-\beta y_{t-1}-\gamma z_{t-1}} e^{-u^2/2\sigma^2}du\right)\mathbf{1}(y_t = 0). \end{aligned} \tag{C.8}$$

Function under expectation in Eq. (C.8),

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

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 Hayashi (2000). Our model satisfies all the conditions of the proposition. Thus, the MLE estimate $\hat{\theta}$ is consistent. □

Appendix D. Proofs: Stationarity and Explosiveness

This subsection presents proofs on stationary/explosive behavior of the process y_{ijt} as $t \to \infty$.

Proof of Theorem 1. To simplify notation, let us denote $z_{ijt-1} = p\left(\mathbf{y}_{t-1}, \ldots, \mathbf{y}_{t-H}\right)$. Then

$$\begin{aligned} y_{ijt} &= [\alpha_{ij} + \beta_{ij} y_{ijt-1} + \gamma_{ij} z_{ijt-1} + u_{ijt}]_+ \\ &\leq \max(0, \beta_{ij}) y_{ijt-1} + |\gamma_{ij} z_{ijt-1}| + \max(0, \alpha_{ij} + u_{ijt}) \\ &\leq (\max(0, \beta_{ij}) + |\gamma_{ij}|) \max_{k,l;s=t-H,\ldots,t-1} y_{kls} + |\mathcal{A}| + \max(0, \alpha_{ij} + u_{ijt}) \\ &\leq C \max_{k,l;s=t-H,\ldots,t-1} y_{kls} + |\mathcal{A}| + \max(0, \alpha_{ij} + u_{ijt}). \end{aligned} \tag{D.1}$$

Denote $v_t = |\mathcal{A}| + \max_{k,l;s=t-H,\ldots,t-1} \max(0, \alpha_{kl} + u_{kls})$. Then from Eq. (D.1) we get

$$\max_{i,j} y_{ijt} \leq C \max_{k,l;s=t-H,\ldots,t-1} y_{kls} + v_t. \tag{D.2}$$

We are going to show that for any $m \in \mathbb{N}$,

$$\max_{\substack{i,j;\\ s=mH+1,\ldots,mH+H}} y_{ijs} \leq C \max_{\substack{i,j;\\ s=(m-1)H+1,\ldots,mH}} y_{ijs} + w_m,$$

where $w_m = v_{mH+1} + \ldots + v_{mH+H}$.

Fix some $m \in \mathbb{N}$. By Eq. (D.2) for $t = mH + 1$,

$$\max_{i,j} y_{i,j,mH+1} \leq C \max_{i,j;s=(m-1)H+1,\ldots,mH} y_{ijs} + v_{mH+1}.$$

Applying Eq. (D.2) twice (for $t = mH + 2$ and $t = mH + 1$) we get

$$\begin{aligned} \max_{i,j} y_{i,j,mH+2} &\leq C \max_{i,j;s=(m-1)H+2,\ldots,mH+1} y_{ijs} + v_{mH+2} \\ &\leq C \max\left(C \max_{i,j;s=(m-1)H+1,\ldots,mH} y_{ijs} + v_{mH+1}, \max_{i,j;s=(m-1)H+2,\ldots,mH} y_{ijs}\right) + v_{mH+2} \\ &\leq C \max_{i,j;s=(m-1)H+1,\ldots,mH} y_{ijs} + v_{mH+1} + v_{mH+2}, \end{aligned}$$

because $C < 1$ and $v_t \geq 0$.

We can redo the same for $t = mH + 3, \ldots, mH + H$, so that

$$\max_{i,j} y_{i,j,mH+r} \leq C \max_{i,j;s=(m-1)H+1,\ldots,mH} y_{ijs} + v_{mH+1} + \ldots + v_{mH+r}.$$

Thus,

$$\max_{\substack{i,j;\\ s=mH+1,\ldots,mH+H}} y_{ijs} \leq C \max_{\substack{i,j;\\ s=(m-1)H+1,\ldots,mH}} y_{ijs} + w_m,$$

where $w_m = v_{mH+1} + \ldots + v_{mH+H} \geq 0$.

Iterative back-substitution leads to

$$\begin{aligned}\max_{\substack{i,j;\\ s=mH+1,\ldots,mH+H}} y_{ijs} &\leq C \max_{\substack{i,j;\\ s=(m-1)H+1,\ldots,mH}} y_{ijs} + w_m \\ &\leq w_m + Cw_{m-1} + C^2 \max_{\substack{i,j;\\ s=(m-2)H+1,\ldots,(m-1)H}} y_{ijs} \leq \sum_{s=0}^{m-1} C^s w_{m-s} + C^m \max_{\substack{i,j;\\ s=1,\ldots,H}} y_{ijs}.\end{aligned} \tag{D.3}$$

Note that $\mathbb{E}w_t = H\mathbb{E}v_t < C_1 < \infty$, as H and number of agents n are finite. Then from Eq. (D.3)

$$\mathbb{E}y_{ijt} \leq \mathbb{E} \max_{\substack{k,l;\\ s=H\lfloor\frac{t-1}{H}\rfloor+1,\ldots,H\lfloor\frac{t-1}{H}\rfloor+H}} y_{kls} \leq \frac{\mathbb{E}w}{1-C} + const < C_2 < \infty. \quad \square$$

In the following set of theorems we show sufficient condition for stationarity for the general model (Theorem 2) and the full characterization of stationary/explosive limiting behavior of y_t depending on α, β (Theorem 3) for the model without peer effects.

Lemma D.1. *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

$$\bar{u}_t \overset{d}{=} \xi_t \bar{u}_t^+ + (1 - \xi_t)\bar{u}_t^-. \tag{D.4}$$

Fix realizations of $(\bar{u}_t^+, \bar{u}_t^-, \xi_t)$ for $t = 1, \ldots, T$ and calculate the corresponding $\bar{u}_t$ from Eq. (D.4). Define

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

Now construct a new time series

$$y'_t = C \max_{s=t-1,\ldots,t-H} y'_s + v_t,$$

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

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

$$y'_{t+p} \leq C \max_{s=t-1,\ldots,t-H} y'_s + \sum_{s=0}^{H-1} v_{t+s} \text{ for all } p = 0, \ldots, H-1. \tag{D.5}$$

Now define another random process and error, $x_\tau = \max_{s=(\tau-1)H,\ldots,\tau H-1} y'_s$, $w_\tau = \sum_{s=0}^{H-1} v_{(\tau-1)H+s}$. Then by Eq. (D.5),

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

We need to find M such that for some $\varepsilon > 0$, $\mathbb{P}(\sharp\{\tau \in [1, \ldots, \lfloor T/H \rfloor] | x_\tau < M\} \geq \varepsilon T) \geq \frac{1}{T^2}$ 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

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

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 $\mathbb{E}|u^+_{ijt}|^4 \leq 2\mathbb{E}|u^4_{ijt}|$.

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 $\mathbb{P}(\sharp\{\tau \in [1, \ldots, T/H] | \tilde{x}_\tau < M\} \geq \varepsilon T) \geq \frac{1}{T^2}$. Let us show that $\exists Q$ such that $\tilde{x}_1 + \tilde{x}_2 + \ldots + \tilde{x}_{\lfloor T/H \rfloor} < Q\lfloor T/H \rfloor$ with probability greater that $1 - \frac{const}{T^2}$.

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

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

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

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,

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

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

(D.6)

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

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 $\sharp\{\tau | x_\tau \le 2Q\} > 0.5\lfloor T/H\rfloor$. That is, we have shown that $\exists M$ (any number larger than $2Q$ and M') such that $\sharp\{\tau | x_\tau \le M\} > \varepsilon T$ 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 $(\mathbb{P}(\exists i, j \text{ s.t. } u_{ijt} \ge -M))^{\varepsilon T}$ the process does not jump to zero. Thus,

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

Corollary D.2. *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 D.1 applies. $\square$

Lemma D.3. *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

$$\mathbb{E}(\text{length until zero}) = \sum_{T=1}^{\infty}\mathbb{P}(length \ge T). \tag{D.7}$$

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, $\mathbb{P}(length \geq T) \leq \mathbb{P}(S_1 > 0, \ldots, S_{T-1} > 0)$. 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 $T > T'$.

(D.8)

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

where we used Markov inequality to bound probability by expectation.

Plugging Eq. (D.8) into Eq. (D.7), we get

$$\mathbb{E}(\text{length until zero}) \leq \sum_{T=1}^{\infty} \frac{const}{T^2} < const_1 < \infty. \qquad \square$$

Theorem D.4. *Suppose that Assumptions 1, 2, and 3 are satisfied and* $\mathbb{E}u_{ijt}^4 < \infty$ *for all* i, j, t. *For the model without* γ, *if* $\beta < 1$ *or* $\beta = 1$, $\alpha < 0$, *then* y_t *converges to a stationary distribution. For the model with* γ, *if for some* C $\max(0, \beta_{ij}) + |\gamma_{ij}| < C < 1$ *for all* i, j, *then* $\mathbf{y}_t$ *converges to a stationary distribution*[1].

Proof. Let us first show convergence to a stationary distribution. The proof follows the lines of section $XI.8$ in Feller (2008). Let us show the proof for the model without peer effects.

From Lemmas D.1 and D.3, we know that $\mathbb{E}$(length until zero) is finite. Thus, with probability one the process reaches zero. The continuation of a process after it reaches zero is a probabilistic replica of the whole process started from the previous zero.

For any Borel set Δ denote by $P_\Delta(t)$ the probability that $y_{t+s} \in \Delta$ given that s is the (finite) time before the process first hits zero. The process y_t is by definition strongly Markov, so that such probability does not depend on s.

[1]Formally this means that the finite-dimensional distributions of the process $\{y_{t+\tau}\}_{\tau\in\mathbb{Z}}$, converge to those of a stationary in τ process as $t \to \infty$.

We are going to show that for any Borel set Δ there exists $\lim_{t\to\infty} P_\Delta(t) = P_\Delta$ such that $P_\Delta \geq 0$, $P_{\mathbb{R}_+} = 1$, and P_Δ is countably additive. This would imply that the one-point distribution of y_t converges to a limit at $t \to \infty$; by the Markov property the latter further implies the desired convergence of all finite-dimensional distributions of $\{y_{t+\tau}\}_{\tau\in\mathbb{Z}}$.

Define by S_1 the first time of hitting zero, by S_2 the second time of hitting zero, etc. Also define $q_\Delta(t) = \mathbb{P}(S_1 > t, y_t \in \Delta)$. Then

$$q_\Delta(t) + q_{\mathbb{R}_+\setminus\Delta}(t) = 1 - F(t),$$

where F is a distribution of time between two consequent moments of hitting zero $(S_{n+1}-S_n)$. By the strong Markov property the distribution of $S_{n+1} - S_n$ does not depend on n and we denote the expected time it takes the process to return to zero after starting at zero by μ, i.e.,

$$\mu = \mathbb{E}(S_{n+1} - S_n) = \int t dF(t).$$

Because the probability that $y_{t+s} \in \Delta$ given $S_1 = s$ does not depend on s, we can write

$$P_\Delta(t) = \mathbb{P}(S_1 > t,\, y_t \in \Delta) + \mathbb{P}(S_1 \leq t,\, y_t \in \Delta) = q_\Delta(t) + \int_0^t P_\Delta(t-y)F(dy).$$

The function $q_\Delta(t)$ is directly integrable since it is dominated by the monotone integrable function $1 - F$. Therefore by the renewal theorem $\lim_{t\to\infty} P_\Delta(t) = \frac{1}{\mu}\sum_{t=0}^{\infty} q_\Delta(t) \geq 0$.

By definition, $q_{\mathbb{R}_+}(t) = 1 - F(t)$ so that $\lim_{t\to\infty} P_{\mathbb{R}_+}(t) = \frac{1}{\mu}\sum_{t=0}^{\infty}(1 - F(t)) = \frac{\mu}{\mu} = 1$. Similarly, for any Δ, $q_\Delta(t) \leq 1 - F(t)$, and, thus, $\lim_{t\to\infty} P_\Delta(t) \leq 1$. Finally, we need to check that P_Δ is countably additive. This follows from the fact that for a countable number of pairwise disjoint sets Δ_i,

$$q_{\cup_i \Delta_i}(t) = \mathbb{P}(S_1 > t, y_t \in \cup_i \Delta_i) = \sum_i q_{\Delta_i}(t).$$

Thus, $\lim_{t\to\infty} P_{\cup\Delta_i}(t) = \frac{1}{\mu}\sum_{t=0}^{\infty} q_{\cup\Delta_i}(t) = \frac{1}{\mu}\sum_{t=0}^{\infty}\sum_i q_{\Delta_i}(t) = \sum_i \lim_{t\to\infty} P_{\Delta_i}(t)$.

The proof for the model with peer effects is the same, with the only difference that now y_t is a vector $\{y_{ijt}\}_{i,j}$. Such process is strongly Markov in extended space, where the element is a vector $\{y_{ij,t-1}, \ldots, y_{ij,t-H}\}_{i,j}$. (That is, we divide the time scale into blocks of length H.) Moreover, by Lemma D.1, $\mathbb{E}$(length until zero) is finite. □

Theorem D.5. *Suppose that Assumptions 1, 2, and 3 are satisfied and* $\mathbb{E}u_{ijt}^4 < \infty$ *for all* i, j, t. *For the model without* γ, *if* $\beta < 1$ *or* $\beta = 1, \alpha < 0$, *then* y_t *is strongly mixing. For the model with* γ, *if for some* C $\max(0, \beta_{ij}) + |\gamma_{ij}| < C < 1$ *for all* i, j, *then* $\mathbf{y}_t$ *is strongly mixing.*

Proof. The proof of strong mixing adapts some ideas from Michel and de Jong (2018). Let us show the proof for the model without peer effects. The proof for the model with peer effects is the same with the only difference that now y_t is a vector $\{y_{ijt}\}_{i,j}$.

For any $m, \tau \in \mathbb{N}$, $\tau > m$, define the process $\tilde{y}_t^{\tau,m}$ as follows

$$\tilde{y}_t^{\tau,m} = \begin{cases} 0 & \text{if } t < \tau - m, \\ [\alpha + \beta \tilde{y}_{t-1}^{\tau,m} + u_t]_+ & \text{otherwise,} \end{cases} \tag{D.9}$$

where u_t is the same error term as is used in the definition of y_t. That is, $\tilde{y}_t^{\tau,m}$ equals zero until time $\tau - m$, and then starts to evolve with the same error and coefficients as y_t.

Define the probability that two processes merge at time τ as

$$P^{\tau,m} := \mathbb{P}(\tilde{y}_\tau^{\tau,m} = y_\tau).$$

Notice that because both processes are generated by the same errors and coefficients, after they coincide at some time t, they are also exactly the same for any time $> t$. We claim that

$$\sup_{\tau:\, \tau > m} P^{\tau,m} \xrightarrow[m\to\infty]{} 1. \tag{D.10}$$

That is, after evolving for m periods ($t = \tau - m, ..., \tau - 1$) the processes eventually merge at time τ with probability 1 as $m \to \infty$. This follows from the convergence to stationary distribution (Theorem D.4), which guarantees that $y_\tau = O_P(1)$, so that y_τ takes large values with very small probability, and from slight modification of Lemmas D.1 and D.3. The latter shows that with probability tending to 1 process $\tilde{y}_t^{\tau,m}$ and y_t simultaneously reach zero between times $t = \tau - m$ and $t = \tau$, thus, they coincide in all future periods.

Now we are ready to show that

$$\lim_{t\to\infty} \sup_s \sup_{\substack{\Delta_1 \in \mathcal{Y}_{0,s}, \\ \Delta_2 \in \mathcal{Y}_{t+s,\infty}}} |\mathbb{P}(\Delta_1 \cap \Delta_2) - \mathbb{P}(\Delta_1)\mathbb{P}(\Delta_2)| = 0. \tag{D.11}$$

Each $\Delta_2 \in \mathcal{Y}_{t+s,\infty}$ can be written as $\Delta_2 = \big(\{y_{t+s}, y_{t+s+1}, \dots\} \in A\big)$ for some Borel set A. For each Δ_2, define the event $\tilde{\Delta}_2^m = \big(\{\tilde{y}_{t+s}^{t+s,m}, \tilde{y}_{t+s+1}^{t+s,m}, \dots\} \in A\big)$. That is, instead of the process $y_\cdot$ being in the set A, we require the modified process $\tilde{y}_\cdot^{t+s,m}$ to be in that set. The symmetric difference $\ominus$ between Δ_2 and $\tilde{\Delta}_2^m$ (union of $\Delta_2 \setminus \tilde{\Delta}_2^m$ and $\tilde{\Delta}_2^m \setminus \Delta_2$) has probability

$$\mathbb{P}(\Delta_2 \ominus \tilde{\Delta}_2^m) \le 1 - P^{t+s,m} \xrightarrow[m\to\infty]{} 0 \quad \text{uniformly over } t+s. \tag{D.12}$$

Moreover, if $t > m$, then $t + s - m > s$ and $\tilde{\Delta}_2^m$ depends only on the errors u_τ, $\tau > s$, while Δ_1 depends only on the errors up until time s. Thus, the random components in two sets are independent and

$$\mathbb{P}(\Delta_1 \cap \tilde{\Delta}_2^m) = \mathbb{P}(\Delta_1)\mathbb{P}(\tilde{\Delta}_2^m). \tag{D.13}$$

Combining (D.12) and (D.13), we get the desired limit (D.11), i.e. strong mixing for y_t. □

Theorem D.6. *Let Assumption 1 hold. Consider the model without γ and assume that $\mathbb{P}(u_t > -\alpha) > 0$, $\mathbb{E}u_t = 0$, and $\mathbb{E}u_t^4 < \infty$. If either (a) $\beta = 1$ and $\alpha > 0$ or (b) $\beta > 1$, then y_t converges to ∞ almost surely ($y_t \xrightarrow[t\to\infty]{a.s.} \infty$).*

Proof. We need to show that $\forall M \geq 0$ $\mathbb{P}(\lim_{t\to\infty} y_t < M) = 0$.

Assume first that $\alpha > 0$ and note, that because taking positive part of a random variable can only increase it, we get

$$y_t = [\alpha + \beta y_{t-1} + u_t]_+ \geq \alpha + \beta y_{t-1} + u_t \geq \alpha + y_{t-1} + u_t.$$

Thus, if $\alpha > 0$, then $y_t \geq \alpha t + y_0 + \sum_{s=1}^{t} u_s$, and

$$\text{If } y_t < M \text{ then } \alpha t + y_0 + \sum_{s=1}^{t} u_s < M \text{ and } \frac{1}{t}\sum_{s=1}^{t} u_s < \frac{M - y_0}{t} - \alpha.$$

By the strong law of large numbers, $\frac{1}{t}\sum_{s=1}^{t} u_s \xrightarrow[t\to\infty]{a.s.} \mathbb{E}u_t = 0$. Therefore, $\mathbb{P}\left(\lim_{t\to\infty} \frac{1}{t}\sum_{s=1}^{t} u_s = 0\right) = 1$ and $\forall \varepsilon > 0$ $\mathbb{P}\left(\lim_{t} \frac{1}{t}\sum_{s=1}^{t} u_s < -\varepsilon\right) = 0$.

Fix $\varepsilon = \frac{2\alpha}{3}$ and T' such that $\frac{M-y_0}{T'} < \frac{\alpha}{3}$. Then

$$\mathbb{P}\left(\lim_{t\to\infty} y_t < M\right) \leq \mathbb{P}\left(\lim_{t\to\infty} \frac{1}{t}\sum_{s=1}^{t} u_s - \frac{M - y_0}{t} < -\alpha\right) \leq \mathbb{P}\left(\lim_{t\to\infty} \frac{1}{t}\sum_{s=1}^{t} u_s < -2\alpha/3\right) = 0.$$

So that $\mathbb{P}\left(\lim_{t\to\infty} y_t < M\right) = 0$.

Now suppose that $\alpha \leq 0$ and $\beta > 1$. We first will show convergence in probability. We are going to show that

$$\forall M \quad \mathbb{P}(y_t \in [M, M+1]) \xrightarrow[t\to\infty]{} 0.$$

Then $\mathbb{P}(|y_t| < M) \to 0$. As $y_t \geq 0$, we only need to consider intervals in $\mathbb{R}_+$.

Consider events $S_t(M) = \begin{cases} y_t \in [M, M+1] \\ y_{t+s} > M+1 \;\; \forall s \geq 1. \end{cases}$ Such events are disjoint for $t \neq t'$ ($S_t(M) \cap S_{t'}(M) = \emptyset$ when $t \neq t'$). Thus,

$$\sum_{t=1}^{\infty} \mathbb{P}(S_t(M)) \leq 1. \tag{D.14}$$

Choose $A > 1$ such that

$$\left(\mathbb{E}u^4 + \left(\mathbb{E}u^2\right)^2\right) \sum_{s=1}^{\infty} \frac{3s^2}{(As-1)^4} < \frac{1}{2}. \tag{D.15}$$

Also choose $\varepsilon > 0$ such that $\beta - \varepsilon > 1$ and M' such that $\varepsilon M' > |\alpha| + A$. Thus, conditional on being in $S_t(M')$, $\forall s \geq 1$

$$\begin{aligned} y_{t+s} &= [\alpha + \beta y_{t+s-1} + u_{t+s}]_+ \geq \alpha + \beta y_{t+s-1} + u_{t+s} \\ &\geq A + (\beta - \varepsilon) y_{t+s-1} + u_{t+s} \geq As + y_t + \sum_{r=1}^{s} u_{t+r}. \end{aligned} \tag{D.16}$$

If $y_t \geq M$, then from Eq. (D.16), if $A + u_{t+1} > 1$, we get $y_{t+1} > M + 1$. If also $2A + u_{t+1} + u_{t+2} > 1$, then similarly $y_{t+2} > M + 1$. Thus, if $Ap + \sum_{r=1}^{p} u_{t+r} > 1$ for all $p = 1, \ldots, s$ we get $y_{t+s} > M + 1$. Let us calculate the probability that $Ap + \sum_{r=1}^{p} u_{t+r} > 1$ for all $p \geq 1$.

$$\begin{aligned} &\mathbb{P}\left(As + \sum_{r=1}^{s} u_{t+r} \leq 1\right) = \mathbb{P}\left(\sum_{r=1}^{s} u_{t+r} \leq 1 - As\right) \leq \mathbb{P}\left(\left|\sum_{r=1}^{s} u_{t+r}\right| \geq As - 1\right) \\ &\leq \frac{\mathbb{E}\left(\sum_{r=1}^{s} u_{t+r}\right)^4}{(As-1)^4} = \frac{s\mathbb{E}u^4 + 3s(s-1)\left(\mathbb{E}u^2\right)^2}{(As-1)^4} \leq \frac{3s^2\left(\mathbb{E}u^4 + \left(\mathbb{E}u^2\right)^2\right)}{(As-1)^4}, \end{aligned}$$

where we first used the fact that $1 - As < 0$ to go from $\sum_{r=1}^{s} u_{t+r}$ to $\left|\sum_{r=1}^{s} u_{t+r}\right|$, and then we used the Markov inequality for $\left|\sum_{r=1}^{s} u_{t+r}\right|^4$ to bound probability by expectation.

Therefore,

$$\begin{aligned} \mathbb{P}\left(Ap + \sum_{r=1}^{p} u_{t+r} > 1 \ \forall p \geq 1\right) &\geq 1 - \sum_{s=1}^{\infty} \mathbb{P}\left(As + \sum_{r=1}^{s} u_{t+r} \leq 1\right) \\ &\geq 1 - \sum_{s=1}^{\infty} \frac{3s^2\left(\mathbb{E}u^4 + \left(\mathbb{E}u^2\right)^2\right)}{(As-1)^4} \geq 1 - \frac{1}{2} = \frac{1}{2}. \end{aligned} \tag{D.17}$$

Thus, for $M \geq M'$

(D.18)

$$\begin{aligned} \mathbb{P}(S_t(M)) &\geq \mathbb{P}\left(y_t \in [M, M+1],\ Ap + \sum_{r=1}^{p} u_{t+r} > 1 \ \forall p \geq 1\right) \\ &= \mathbb{P}\left(y_t \in [M, M+1]\right) \mathbb{P}\left(Ap + \sum_{r=1}^{p} u_{t+r} > 1 \ \forall p \geq 1\right) \geq 0.5\mathbb{P}\left(y_t \in [M, M+1]\right). \end{aligned}$$

Plugging Eq. (D.18) into Eq. (D.14), we get $\sum_{t=1}^{\infty} \mathbb{P}\left(y_t \in [M, M+1]\right) \leq 2$. Thus, the series converges, so it must be that $\mathbb{P}\left(y_t \in [M, M+1]\right) \xrightarrow[t\to\infty]{} 0$.

We have shown that for $M \geq M'$, $\mathbb{P}\left(y_t \in [M, M+1]\right) \xrightarrow[t\to\infty]{} 0$. Let us show that this also holds for $M < M'$.

We know that for all $M \geq M'$, $\lim_t \mathbb{P}(y_t \in [M, M+1]) = 0$. Thus, also $\lim_t \mathbb{P}(y_t \in [M', \beta M' + K]) = 0$ for any $K \geq 0$. Suppose by contradiction that

$$\lim_t \mathbb{P}\left(y_t \in [\beta^{-1}M', M']\right) = p_{M'} > 0.$$

That is, $\forall \varepsilon > 0\ \exists T$ s.t. $\forall t > T$, $\mathbb{P}(y_t \in [\beta^{-1}M', M']) > p_{M'} - \varepsilon$.

Because $\mathbb{P}(u_t > -\alpha) > 0$, there exists $K > 0$ such that $\mathbb{P}(u_t \in [-\alpha, -\alpha + K]) := p_\alpha > 0$. Thus, $\forall t > T$ with probability at least $(p_{M'} - \varepsilon)p_\alpha > 0$, $y_{t+1} = [\alpha + \beta y_t + u_{t+1}]_+ \in [M', \beta M' + K]$, so that $\lim_t \mathbb{P}(y_t \in [M', \beta M' + K]) > 0$ and we get a contradiction. Thus, $\lim_t \mathbb{P}(y_t \in [\beta^{-1}M', M']) = 0$. We can repeat the argument with $\beta^{-1}M'$ instead of M', then with $\beta^{-2}M'$ instead of $\beta^{-1}M'$ and so on. Thus, we get that $\mathbb{P}(y_t \in (0, M]) \xrightarrow[t\to\infty]{} 0$ for all $M > 0$ (as $\beta^{-k} \to 0$). If $\lim_t \mathbb{P}(y_t = 0) > 0$, then by a similar argument we must have that $\lim_t \mathbb{P}(y_t \in [1, 2]) > 0$, which is a contradiction. (Take $u_{t+1} \in [-\alpha + 1, -\alpha + 2]$.)

Therefore, $\mathbb{P}(y_t \in [M, M+1]) \xrightarrow[t\to\infty]{} 0$ for all M and $y_t \xrightarrow{\mathbb{P}} \infty$.

Now let us show that $y_t \xrightarrow{a.s.} \infty$. Note that we can choose $A(k)$ in Eq. (D.15) such that

$$\left(\mathbb{E}u^4 + \left(\mathbb{E}u^2\right)^2\right) \sum_{s=1}^{\infty} \frac{3s^2}{(A(k)s - 1)^4} < \frac{1}{k}$$

and $M(k)$ such that $\varepsilon M(k) > |\alpha| + A(k)$. Then if $y_t > M(k)$, for Eq. (D.17) with A replaced by $A(k)$, we get that $y_{t+s} > M(k)$ for all $s \geq 1$ with probability at least $1 - \frac{1}{k}$.

Because $y_t \xrightarrow{\mathbb{P}} \infty$, for any M $\mathbb{P}(y_t > M) \xrightarrow[t\to\infty]{} 1$. Thus, for any M and $\delta > 0$ there exists T such that $\mathbb{P}(y_T > M) > 1 - \delta$. Therefore, if $M > M(k)$, $\lim_t y_t > M$ with probability of at least $(1-\delta)(1-1/k)$. Because δ is arbitrary, we must have $\mathbb{P}(\lim_t y_t > M) \geq 1 - 1/k$ for any $M > M(k)$. Because k can be chosen arbitrary, we must have $\mathbb{P}(\lim_t y_t > M) = 1$ for any M. (Note that if $M' > M''$, then $\mathbb{P}(\lim_t y_t > M'') \geq \mathbb{P}(\lim_t y_t > M')$). Thus, $y_t \xrightarrow{a.s.} \infty$. ☐

Theorem D.7. *Let Assumption 1 hold. Consider the model without γ and assume that u_t has unbounded support from below, $\mathbb{E}u_t = 0$, and $\mathbb{E}u_t^4 < \infty$. If $\beta = 1$, $\alpha = 0$, then y_t is mean-divergent ($\mathbb{E}y_t \xrightarrow[t\to\infty]{} \infty$).*

Proof. Because u_t has unbounded support from below, with positive probability $u_t < -y_{t-1}$. Thus,

$$\mathbb{E}y_t = \mathbb{E}[y_{t-1} + u_t]_+ > \mathbb{E}(y_{t-1} + u_t) = \mathbb{E}y_{t-1},$$

and $\mathbb{E}y_t$ is a strictly increasing sequence of t. Therefore, either $\mathbb{E}y_t \to \infty$ or $\mathbb{E}y_t \to const.$ Suppose that the latter is true. Then by Markov's inequality for any $C > 0$,

$$\mathbb{P}(y_t \geq C) \leq \frac{\mathbb{E}y_t}{C} \leq \frac{\lim_t \mathbb{E}y_t}{C}.$$

Let us choose C such that $\frac{\lim_t \mathbb{E} y_t}{C} \leq \frac{1}{2}$. Thus, for any t, $\mathbb{P}(y_t \geq C) \leq \frac{1}{2}$ and $\mathbb{P}(y_t < C) \geq \frac{1}{2}$.

$$\mathbb{E} y_t = \mathbb{E}[y_{t-1} + u_t]_+ (\mathbf{1}\{y_{t-1} \geq C\} + \mathbf{1}\{y_{t-1} < C\}), \tag{D.19}$$

$$\mathbb{E}[y_{t-1} + u_t]_+ \mathbf{1}\{y_{t-1} \geq C\} \geq \mathbb{E}(y_{t-1} + u_t) \mathbf{1}\{y_{t-1} \geq C\} = \mathbb{E} y_{t-1} \mathbf{1}\{y_{t-1} \geq C\}, \tag{D.20}$$

(D.21)

$$\begin{aligned} &\mathbb{E}[y_{t-1} + u_t]_+ \mathbf{1}\{y_{t-1} < C\} = \mathbb{E}[y_{t-1} + \max(-C, u_t)]_+ \mathbf{1}\{y_{t-1} < C\} \\ &\quad \geq \mathbb{E}(y_{t-1} + \max(-C, u_t)) \mathbf{1}\{y_{t-1} < C\} = \mathbb{E} y_{t-1} \mathbf{1}\{y_{t-1} < C\} + \mathbb{P}(y_{t-1} < C) \mathbb{E} \max(-C, u_t). \end{aligned}$$

Because $\mathbb{P}(y_t < C) \geq \frac{1}{2}$ and $\mathbb{E} \max(-C, u_t) > \mathbb{E} u_t = 0$, combining Eq. (D.19), (D.20), and (D.21), we get

$$\mathbb{E} y_t \geq \mathbb{E} y_{t-1} \mathbf{1}\{y_{t-1} \geq C\} + \mathbb{E} y_{t-1} \mathbf{1}\{y_{t-1} < C\} + \mathbb{P}(y_{t-1} < C) \mathbb{E} \max(-C, u_t) \geq \mathbb{E} y_{t-1} + C_1.$$

where $C_1 = \mathbb{P}(y_{t-1} < C) \mathbb{E} \max(-C, u_t) > 0$. Thus, $\mathbb{E} y_t \geq t C_1 + const$, and $\mathbb{E} y_t \to \infty$. □

Lemma D.8. *If $\beta = 1$, $\alpha = 0$, then we can equivalently rewrite the evolution of y_t as follows*

$$y_t = [y_{t-1} + u_t]_+ = y_0 + \sum_{s=1}^{t} u_s + \sup_{r=0,\dots,t} \left[-y_0 - \sum_{s=1}^{r} u_s \right]_+ .$$

Proof. Define $x_t = y_0 + \sum_{s=1}^{t} u_s + \sup_{r=0,\dots,t} \left[-y_0 - \sum_{s=1}^{r} u_s \right]_+$, $x_0 = y_0$. Note that by definition x_t is always non-negative, as when $y_0 + \sum_{s=1}^{t} u_s$ becomes negative, we are adding its absolute value or even a larger positive number ($\sup_{r=0,\dots,t} \left[-y_0 - \sum_{s=1}^{r} u_s \right]_+$). Let us show that $x_t = y_t$ for all t. Let us proceed by induction.

By definition $x_0 = y_0 \geq 0$. Let us look at $t = 1$. If $y_0 + u_1 \geq 0$, then $\sup_{r=0,1} \left[-y_0 - \sum_{s=1}^{r} u_s \right]_+ = 0$ and $x_1 = y_0 + u_1 = y_1$. If $y_0 + u_1 < 0$, then $y_1 = [y_0 + u_1]_+ = 0$ and $\sup_{r=0,1} \left[-y_0 - \sum_{s=1}^{r} u_s \right]_+ = -y_0 - u_1 > 0$. Thus, $x_1 = y_0 + u_1 + (-y_0 - u_1) = 0 = y_1$.

Suppose that $x_t = y_t$ for all $t \leq t'$. Let us prove that $x_{t'+1} = y_{t'+1}$. First, suppose that $\exists p \in \{0, \dots, t'\}$ such that $\left[-y_0 - \sum_{s=1}^{p} u_s \right]_+ \geq \left[-y_0 - \sum_{s=1}^{t'+1} u_s \right]_+$. Thus, either $t+1$ is not an argmaximum or it is not a unique argmaximum over $\{0, \dots, t'+1\}$. In that case,

$$\sup_{r=0,\dots,t'} \left[-y_0 - \sum_{s=1}^{r} u_s \right]_+ = \sup_{r=0,\dots,t'+1} \left[-y_0 - \sum_{s=1}^{r} u_s \right]_+$$

and

$$x_{t'+1} = y_0 + \sum_{s=1}^{t'+1} u_s + \sup_{r=0,\dots,t'+1} \left[-y_0 - \sum_{s=1}^{r} u_s \right]_+$$
$$= y_0 + \sum_{s=1}^{t'} u_s + u_{t'+1} + \sup_{r=0,\dots,t'} \left[-y_0 - \sum_{s=1}^{r} u_s \right]_+ = x_{t'} + u_{t'+1} = y_{t'} + u_{t'+1} \geq 0,$$

where we used the induction hypothesis and the observation that $x_t \geq 0$ for all t.

Thus,

$$y_{t'+1} = [y_{t'} + u_{t'+1}]_+ = y_{t'} + u_{t'+1} = x_{t'+1}.$$

Now suppose that $\forall p \in \{0, \dots, t'\}$, $\left[-y_0 - \sum_{s=1}^{p} u_s \right]_+ < \left[-y_0 - \sum_{s=1}^{t'+1} u_s \right]_+$. Thus, $\left[-y_0 - \sum_{s=1}^{t'+1} u_s \right]_+ > 0$ and

$$x_{t'+1} = y_0 + \sum_{s=1}^{t'+1} u_s + \sup_{r=0,\dots,t'+1} \left[-y_0 - \sum_{s=1}^{r} u_s \right]_+ = y_0 + \sum_{s=1}^{t'+1} u_s + \left(-y_0 - \sum_{s=1}^{t'+1} u_s \right)_+ = 0.$$

In turn,

$$y_{t'+1} = [y_{t'} + u_{t'+1}]_+ = \left[y_0 + \sum_{s=1}^{t'} u_s + \sup_{r=0,\dots,t'} \left[-y_0 - \sum_{s=1}^{r} u_s \right]_+ u_{t'+1} \right]_+$$
$$= \left[y_0 + \sum_{s=1}^{t'+1} u_s + \sup_{r=0,\dots,t'} \left[-y_0 - \sum_{s=1}^{r} u_s \right]_+ \right]_+ = 0 = x_{t'+1},$$

as $y_0 + \sum_{s=1}^{t'+1} u_s < 0$ and $\left[-y_0 - \sum_{s=1}^{p} u_s \right]_+ < -y_0 - \sum_{s=1}^{t'+1} u_s$ for all $p \in \{0, \dots, t'\}$.

Therefore, $x_{t'+1} = y_{t'+1}$. By induction we get that $y_t = x_t$ for all t. □

Theorem D.9. *If $\beta = 1$, $\alpha = 0$, then for all $r \in (0, 1]$, $\frac{1}{\sqrt{T}} y_{\lfloor rT \rfloor} \xrightarrow[T\to\infty]{d} \sigma |W(r)|$, where $\sigma = \mathbb{E}u^2$ and W is a standard Brownian motion.*

Proof. By Lemma D.8, y_t can be alternatively written as

$$y_t = y_0 + \sum_{s=1}^{t} u_s + \sup_{p=0,\dots,t} \left[-y_0 - \sum_{s=1}^{p} u_s \right]_+ .$$

This is Skorokhod transformation for $y_0 + \sum_{s=1}^{t} u_s$, which is a continuous transformation. Thus,

$$\frac{1}{\sqrt{T}} y_{\lfloor rT \rfloor} = \frac{1}{\sqrt{T}} y_0 + \frac{1}{\sqrt{T}} \sum_{s=1}^{\lfloor rT \rfloor} u_s + \sup_{p=0,\dots,\lfloor rT \rfloor} \left[-\frac{y_0}{\sqrt{T}} - \frac{1}{\sqrt{T}} \sum_{s=1}^{p} u_s \right]_+$$

By functional central limit theorem $\frac{1}{\sqrt{T}}\sum\limits_{s=1}^{\lfloor rT\rfloor} u_s \xrightarrow[T\to\infty]{d} \sigma W(r)$. So that using the continuity of Skorokhod transformation,

$$\frac{1}{\sqrt{T}} y_{\lfloor rT\rfloor} \xrightarrow[T\to\infty]{d} \sigma W(r) + \sup_{p\in[0,r]} [-\sigma W(p)]_+ \stackrel{d}{=} \sigma |W(r)|,$$

where the last equation, which gives equivalence in distribution, was first proved in Lévy (1948). (See, for example, Section 3.6.C in Karatzas and Shreve (2012).) □

Appendix E. Proofs: Properties of the estimators

All properties of estimators are proved in this section. We follow the notations from Section 4 and estimate the equation

$$y_t = [\alpha + \beta y_{t-1} + \gamma z_{t-1} + u_t]_+.$$

E.1. Stationary LAD. We use the following lemma to justify cases, where sums are approximated by expectations.

Lemma E.1. *If the data generating process satisfies stationarity condition of either Theorem 2 or Theorem 3 and there exists* $\mathbb{E}|u_t|^{4+\varepsilon}$ *for some* $\varepsilon > 0$*, then as* $T \to \infty$

$$\frac{1}{T}\sum_{t=1}^{T} y_t \xrightarrow{\mathbb{P}} \mathbb{E}y,\ \frac{1}{T}\sum_{t=1}^{T} y_t^2 \xrightarrow{\mathbb{P}} \mathbb{E}y^2,\ \frac{1}{T}\sum_{t=1}^{T} y_t z_t \xrightarrow{\mathbb{P}} \mathbb{E}yz,\ \frac{1}{T}\sum_{t=1}^{T} z_t \xrightarrow{\mathbb{P}} \mathbb{E}z,\ \frac{1}{T}\sum_{t=1}^{T} z_t^2 \xrightarrow{\mathbb{P}} \mathbb{E}z^2,$$

where expectations are taken with respect to the stationary distribution.

Proof. Note that if the error u_t has a moment of order k, then y_t also has a moment of order k.

Let us show that $Cov(y_s, y_{s+t}) \xrightarrow[t\to\infty]{} 0$.

$$\begin{aligned} Cov(y_s, y_{s+t}) = \mathbb{E}y_s y_{t+s} - \mathbb{E}y_s \mathbb{E}y_{t+s} &= \mathbb{E}y_s y_{t+s}\mathbf{1}(y_s < M, y_{t+s} < M) - \mathbb{E}y_s\mathbb{E}y_{t+s} \\ &+ \mathbb{E}y_s y_{t+s}\mathbf{1}(y_s \geq M, y_{t+s} < M) + \mathbb{E}y_s y_{t+s}\mathbf{1}(y_s < M, y_{t+s} \geq M) \\ &+ \mathbb{E}y_s y_{t+s}\mathbf{1}(y_s \geq M, y_{t+s} \geq M) \end{aligned}$$

As y_t is mixing and converges in distribution to a random variable, its limit is independent of y_s,

$$\mathbb{E}y_s y_{t+s}\mathbf{1}(y_s < M, y_{t+s} < M) - \mathbb{E}y_s\mathbf{1}(y_s < M)\mathbb{E}y_{s+t}\mathbf{1}(y_{s+t} < M) \xrightarrow[t\to\infty]{} 0.$$

As M goes to infinity, $\mathbb{E}y_s\mathbf{1}(y_s < M) \to \mathbb{E}y_s$ for all s. Thus, by choice of M large enough we can make $\lim\limits_{t\to\infty} (\mathbb{E}y_s y_{t+s}\mathbf{1}(y_s < M, y_{t+s} < M) - \mathbb{E}y_s\mathbb{E}y_{t+s})$ arbitrarily close to zero. Moreover,

$$\mathbb{E}y_s y_{t+s}\mathbf{1}(y_s \geq M, y_{t+s} < M) \leq \mathbb{E}\frac{y_s^{1+\varepsilon/2}}{M^{\varepsilon/2}} y_{t+s}\mathbf{1}(y_s \geq M, y_{t+s} < M)$$
$$\leq \frac{1}{M^{\varepsilon/2}}\mathbb{E}(y_s^{2+\varepsilon} + y_{t+s}^2)\mathbf{1}(y_s \geq M, y_{t+s} < M) \leq \frac{1}{M^{\varepsilon/2}}\mathbb{E}(y_s^{2+\varepsilon} + y_{t+s}^2) \xrightarrow[M\to\infty]{} 0.$$

Similarly,

$$\mathbb{E}y_s y_{t+s}\mathbf{1}(y_s < M, y_{t+s} \geq M) \leq \frac{1}{M^{\varepsilon/2}}\mathbb{E}(y_s^2 + y_{t+s}^{2+\varepsilon}) \xrightarrow[M\to\infty]{} 0.$$

Finally,

$$\mathbb{E}y_s y_{t+s}\mathbf{1}(y_s \geq M, y_{t+s} \geq M) \leq \mathbb{E}(y_s^2 + y_{t+s}^2)\mathbf{1}(y_s \geq M, y_{t+s} \geq M)$$
$$\leq \frac{1}{M^{\varepsilon}}\mathbb{E}(y_s^{2+\varepsilon} + y_{t+s}^{2+\varepsilon})\mathbf{1}(y_s \geq M, y_{t+s} \geq M) \leq \frac{1}{M^{\varepsilon}}\mathbb{E}(y_s^{2+\varepsilon} + y_{t+s}^{2+\varepsilon}) \xrightarrow[M\to\infty]{} 0.$$

Thus, by choice of M large enough, $\mathbb{E}y_s y_{t+s}\mathbf{1}(y_s \geq M, y_{t+s} < M) + \mathbb{E}y_s y_{t+s}\mathbf{1}(y_s < M, y_{t+s} \geq M) + \mathbb{E}y_s y_{t+s}\mathbf{1}(y_s \geq M, y_{t+s} \geq M)$ can be made arbitrarily close to zero. That is, $Cov(y_s, y_{s+t}) \to 0$ as $t \to \infty$. This means that the variance of $\frac{1}{T}\sum\limits_{t=1}^{T} y_t$ goes to zero as $T \to \infty$, so that law of large numbers holds for y_t.

To see that the variance of $\frac{1}{T}\sum\limits_{t=1}^{T} y_t$ goes to zero as $T \to \infty$, note that

$$\mathbb{V}\sum_{t=1}^{T} y_t = \mathbb{E}\left(\sum_{t=1}^{T} y_t - T\mathbb{E}y\right)^2 = \sum_{t=1}^{T}\sum_{s=1}^{T}\mathbb{E}(y_t - \mathbb{E}y)(y_s - \mathbb{E}y) \leq \sum_{s=1}^{T}\left(\sum_{t=0}^{T} Cov(y_s, y_{s+t})\right)$$

Because $Cov(y_s, y_{s+t})$ goes to zero as $t \to \infty$, $\sum\limits_{t=0}^{T} Cov(y_s, y_{s+t}) = o(T)$ (otherwise terms would not disappear as $t \to \infty$). Thus, $\sum\limits_{s=1}^{T}\left(\sum\limits_{t=0}^{T} Cov(y_s, y_{s+t})\right) = o(T^2)$ (there are T terms each of order $o(T)$). That is, $\mathbb{V}\sum\limits_{t=1}^{T}\frac{1}{T}y_t = \frac{o(T^2)}{T^2} = o(1)$ and the law of large numbers holds for y_t.

The same logic applies to the four other limits. Corresponding moments of z_t are finite, as z_t is bounded by the maximum of finite number of variables $\{y_{ijs}\}_{s=t,\ldots,t-H+1}$ (Assumption 3). □

Proof of Theorem 5. This proof follows the lines of Powell (1984).

Define $\theta = (\alpha, \beta, \gamma)'$, $x_t = (1,\, y_{t-1},\, z_{t-1})'$, $\mathcal{F}_t = \{y_t, z_t, y_{t-1}, z_{t-1}, \ldots\}$, and

$$S_T(\theta) = \frac{1}{T}\sum_t |y_t - [\alpha + \beta y_{t-1} + \gamma z_{t-1}]_+| = \frac{1}{T}\sum_t |[x_t'\theta_0 + u_t]_+ - [x_t'\theta]_+|,$$

where θ_0 corresponds to the true value of θ.

We want to show that $S_T(\theta) - S_T(\theta_0)$ is uniformly bounded away from zero for large T and $||\theta - \theta_0|| > \varepsilon$ for any $\varepsilon > 0$. Then $\hat{\theta}_{LAD} \underset{\mathbb{P}}{\to} \theta_0$.

Let us write

$$
\begin{aligned}
Q_T(\theta) := S_T(\theta) - S_T(\theta_0) = \frac{1}{T}\sum_t \Bigg[&||[x_t'\theta_0 + u_t]_+ - [x_t'\theta]_+| - |[x_t'\theta_0 + u_t]_+ - [x_t'\theta_0]_+| \\
&- \mathbb{E}\left(|[x_t'\theta_0 + u_t]_+ - [x_t'\theta]_+| - |[x_t'\theta_0 + u_t]_+ - [x_t'\theta_0]_+|\,|\mathcal{F}_{t-1}\right)\Bigg] \\
&+ \frac{1}{T}\sum_t \mathbb{E}\left(|[x_t'\theta_0 + u_t]_+ - [x_t'\theta]_+| - |[x_t'\theta_0 + u_t]_+ - [x_t'\theta_0]_+|\,|\mathcal{F}_{t-1}\right)
\end{aligned}
\tag{E.1}
$$

Let us analyze the first summation in Eq. (E.1). Define

$$
s_t(\theta, x_t) := |[x_t'\theta_0 + u_t]_+ - [x_t'\theta]_+| - |[x_t'\theta_0 + u_t]_+ - [x_t'\theta_0]_+|
$$

and note that $s_t(\theta, x_t) - \mathbb{E}(s_t(\theta, x_t)|\mathcal{F}_{t-1})$ is a martingale difference sequence. Thus,

$$
\mathbb{E}\left(\sum_t \left(s_t(\theta, x_t) - \mathbb{E}(s_t(\theta, x_t)|\mathcal{F}_{t-1})\right)\right)^2 = \sum_t \mathbb{E}\left(s_t(\theta, x_t) - \mathbb{E}(s_t(\theta, x_t)|\mathcal{F}_{t-1})\right)^2.
$$

$$
\begin{aligned}
\mathbb{E}\left(s_t(\theta, x_t) - \mathbb{E}(s_t(\theta, x_t)|\mathcal{F}_{t-1})\right)^2 = \mathbb{E}\Bigg(&|[x_t'\theta_0 + u_t]_+ - [x_t'\theta]_+| - |[x_t'\theta_0 + u_t]_+ - [x_t'\theta_0]_+| \\
- \int \left(|[x_t'\theta_0 + u]_+ - [x_t'\theta]_+| - |[x_t'\theta_0 + u]_+ - [x_t'\theta_0]_+|\right) f(u)du\Bigg)^2,
\end{aligned}
$$

which is a function of distributions of y_t and z_t and parameter θ.

Because $\theta \in \Theta$, which is compact, and second moments of y_t and z_t exist, $\mathbb{E}\left(s_t(\theta, x_t) - \mathbb{E}(s_t(\theta, x_t)|\mathcal{F}_{t-1})\right)^2$ is bounded by some constant C_t (which depends on first two moments of y_t and z_t). Moreover, since y_t and z_t converge in distribution to y and z as $t \to \infty$, $\mathbb{E}\left(s_t(\theta, x_t) - \mathbb{E}(s_t(\theta, x_t)|\mathcal{F}_{t-1})\right)^2$ converges to a corresponding function of a stationary distribution of y and z, which is bounded by some constant C^* for all $\theta \in \Theta$. Thus, there exists t' such that $\forall t > t'$, $\mathbb{E}\left(s_t(\theta, x_t) - \mathbb{E}(s_t(\theta, x_t)|\mathcal{F}_{t-1})\right)^2 \leq C^* + 1$ for all $\theta \in \Theta$. Define $C = \max\{C_1, \ldots, C_{t'}, C^* + 1\}$. Then $\mathbb{E}\left(s_t(\theta, x_t) - \mathbb{E}(s_t(\theta, x_t)|\mathcal{F}_{t-1})\right)^2$ is bounded by C for all t, θ. Therefore, by Markov inequality, for any $a > 0$

$$
\begin{aligned}
\mathbb{P}\left(\left|\frac{1}{T}\sum_t \left(s_t(\theta, x_t) - \mathbb{E}(s_t(\theta, x_t)|\mathcal{F}_{t-1})\right)\right| > a\right) &\leq \frac{\mathbb{E}\left|\sum_t \left(s_t(\theta, x_t) - \mathbb{E}(s_t(\theta, x_t)|\mathcal{F}_{t-1})\right)\right|^2}{T^2a^2} \\
= \frac{\sum_t \mathbb{E}\left(s_t(\theta, x_t) - \mathbb{E}(s_t(\theta, x_t)|\mathcal{F}_{t-1})\right)^2}{T^2a^2} &\leq \frac{C}{Ta^2} \xrightarrow[T\to\infty]{} 0.
\end{aligned}
$$

Thus, for all $\theta \in \Theta$,

$$\frac{1}{T}\sum_t \left(s_t(\theta, x_t) - \mathbb{E}(s_t(\theta, x_t)|\mathcal{F}_{t-1})\right) \xrightarrow[T\to\infty]{\mathbb{P}} 0.$$

Let us show that the convergence is uniform over θ. Note that $s_t(x_t, \theta)$ is Lipschitz:

$$|x_t'\theta_1 - x_t'\theta_2| \le (1 + |y_t| + |z_t|)(|\alpha_1 - \alpha_2| + |\beta_1 - \beta_2| + |\gamma_1 - \gamma_2|)$$

and $|[w_1]_+ - [w_2]_+| \le |w_1 - w_2|$ for any $w_1, w_2 \in \mathbb{R}$, so that

$$||const - [x_t'\theta_1]_+| - |const - [x_t'\theta_2]_+|| \le (1 + |y_t| + |z_t|)(|\alpha_1 - \alpha_2| + |\beta_1 - \beta_2| + |\gamma_1 - \gamma_2|),$$

where we used the fact that if $h_1 : \mathbb{R}^k \to \mathbb{R}^m$ is Lipschitz with constant K_1 and $h_2 : \mathbb{R}^m \to \mathbb{R}$ is Lipschitz with constant K_2, then $h_2(h_1) : \mathbb{R}^k \to \mathbb{R}$ is Lipschitz with constant $K_1 \cdot K_2$.

Thus, $\frac{1}{T}\sum_t s_t(\theta, x_t)$ is Lipschitz with a constant $1 + \frac{1}{T}\sum_t |y_t| + \frac{1}{T}\sum_t |z_t|$. Because y_t and z_t are mixing and converge to y and z as $t \to \infty$, both $\frac{1}{T}\sum_t |y_t|$ and $\frac{1}{T}\sum_t |z_t|$ are bounded in probability, and the Lipschitz constant is finite. Similarly, $\frac{1}{T}\sum_t \mathbb{E}(s_t(\theta, x_t)|\mathcal{F}_{t-1})$ is Lipschitz, as $\frac{1}{T}\sum_t \mathbb{E}(|y_t|\,|\mathcal{F}_{t-1})$ and $\frac{1}{T}\sum_t \mathbb{E}(|z_t|\,|\mathcal{F}_{t-1})$ are bounded in probability. So $\frac{1}{T}\sum_t \left(s_t(\theta, x_t) - \mathbb{E}(s_t(\theta, x_t)|\mathcal{F}_{t-1})\right)$ is Lipschitz. Thus, it satisfies the conditions of the Prokhorov's theorem (see Kallenberg (2002, Theorem 16.5)), which implies

$$\sup_\theta \frac{1}{T}\sum_t \left(s_t(\theta, x_t) - \mathbb{E}(s_t(\theta, x_t)|\mathcal{F}_{t-1})\right) \xrightarrow[T\to\infty]{\mathbb{P}} 0.$$

Let us now analyze the second summation in Eq. (E.1). We are going to show that $\mathbb{E}\left(|y_t - [x_t'\theta]_+| - |y_t - [x_t'\theta_0]_+|\,|\mathcal{F}_{t-1}\right)$ is always non-negative.

Because $y_t = [x_t'\theta_0 + u_t]_+$,

$$\begin{aligned}\mathbb{E}(|y_t - [x_t'\theta]_+||\mathcal{F}_{t-1}) = \mathbf{1}(x_t'\theta < 0) \int_{-x_t'\theta_0}^{\infty} (x_t'\theta_0 + u) f_u(u) du \\ + \mathbf{1}(x_t'\theta \ge 0)\left(x_t'\theta F_u(-x_t'\theta_0) + \int_{-x_t'\theta_0}^{\infty} |u - x_t'(\theta - \theta_0)| f_u(u) du\right)\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}(|y_t - [x_t'\theta_0]_+||\mathcal{F}_{t-1}) = \mathbf{1}(x_t'\theta_0 < 0) \int_{-x_t'\theta_0}^{\infty} (x_t'\theta_0 + u) f_u(u) du \\ + \mathbf{1}(x_t'\theta_0 \ge 0)\left(x_t'\theta_0 F_u(-x_t'\theta_0) + \int_{-x_t'\theta_0}^{\infty} |u| f_u(u) du\right).\end{aligned}$$

Therefore, omitting the derivations, we get

$$\mathbb{E}\left(|y_t - [x_t'\theta]_+| - |y_t - [x_t'\theta_0]_+|\,|\mathcal{F}_{t-1}\right) = 2\mathbf{1}(x_t'\theta_0 \geq 0, x_t'\theta < 0) \int_{-x_t'\theta_0}^{0} (x_t'\theta_0 + u) f_u(u) du$$

$$+ 2\mathbf{1}(x_t'\theta_0 < 0, x_t'\theta \geq 0) \left(\int_{-x_t'\theta_0}^{x_t'(\theta-\theta_0)} (x_t'(\theta-\theta_0) - u) f_u(u) du + \int_{0}^{-x_t'\theta_0} x_t'\theta f_u(u) du \right) \tag{E.2}$$

$$+ 2\mathbf{1}(x_t'\theta_0 \geq 0, x_t'\theta \geq 0) \int_{0}^{x_t'(\theta-\theta_0)} (x_t'(\theta-\theta_0) - u) f_u(u) du \geq 0,$$

as every function under integral is non-negative over the domain of integration.

Because all terms in the Eq. (E.2) are non-negative,

$$\mathbb{E}\left(|y_t - [x_t'\theta]_+| - |y_t - [x_t'\theta_0]_+|\,|\mathcal{F}_{t-1}\right) \geq 2\mathbf{1}(x_t'\theta_0 \geq 0, x_t'\theta < 0) \int_{-x_t'\theta_0}^{0} (x_t'\theta_0 + u) f_u(u) du$$

$$+ 2\mathbf{1}(x_t'\theta_0 \geq 0, x_t'\theta \geq 0) \int_{0}^{x_t'(\theta-\theta_0)} (x_t'(\theta-\theta_0) - u) f_u(u) du.$$

Moreover, for $R > 0$ such that M_R is nonsingular and any $\tau \in (0, R]$,

$$\mathbb{E}\left(|y_t - [x_t'\theta]_+| - |y_t - [x_t'\theta_0]_+|\,|\mathcal{F}_{t-1}\right)$$

$$\geq 2\mathbf{1}(x_t'\theta_0 \geq R, x_t'\theta < 0)\mathbf{1}(|x_t'(\theta-\theta_0)| \geq \tau) \int_{-\tau}^{0} (\tau + u) f_u(u) du$$

$$+ 2\mathbf{1}(x_t'\theta_0 \geq R, x_t'\theta \geq 0)\mathbf{1}(|x_t'(\theta-\theta_0)| \geq \tau) \int_{0}^{\tau} (\tau - u) f_u(u) du$$

$$\geq 2 \min\left(\int_{-\tau}^{0} (\tau + u) f_u(u) du, \int_{0}^{\tau} (\tau - u) f_u(u) du \right) \mathbf{1}(x_t'\theta_0 \geq R)\mathbf{1}(|x_t'(\theta-\theta_0)| \geq \tau).$$

Thus,

$$\frac{1}{T}\sum_t \mathbb{E}\left(|y_t - [x_t'\theta]_+| - |y_t - [x_t'\theta_0]_+|\,|\mathcal{F}_{t-1}\right)$$

$$\geq \frac{2}{T} \min\left(\int_{-\tau}^{0} (\tau + u) f_u(u) du, \int_{0}^{\tau} (\tau - u) f_u(u) du \right) \sum_t \mathbf{1}(x_t'\theta_0 \geq R)\mathbf{1}(|x_t'(\theta-\theta_0)| \geq \tau).$$

As T goes to infinity, $\frac{1}{T}\sum_t \mathbf{1}(x_t'\theta_0 \geq R)\mathbf{1}(|x_t'(\theta-\theta_0)| \geq \tau)$ converges to[2]

$$\begin{aligned}&\mathbb{E}\mathbf{1}(x'\theta_0 \geq R,\, |x'(\theta-\theta_0)| \geq \tau) = \mathbb{P}(x'\theta_0 \geq R,\, |x'(\theta-\theta_0)| \geq \tau)\\ &= \mathbb{P}(x'\theta_0 \geq R)\mathbb{P}(|x'(\theta-\theta_0)| \geq \tau|\, x'\theta_0 \geq R),\end{aligned}$$

where $x = (1, y, z)'$ and (y, z) is the limit distribution of (y_t, z_t).

Because M_R is nonsingular, $\mathbb{P}(x'\theta_0 \geq R) > 0$ (otherwise indicator in M_R will always be zero, so that the matrix M_R will be identically zero).

We are going to prove that

$$\mathbb{P}(|x'(\theta-\theta_0)| \geq \tau_0|\, x'\theta_0 \geq R) \geq C_1 > 0,$$

where $\tau_0 = const \cdot ||\theta-\theta_0||^2$ and $||\cdot||$ denotes L_2 norm. This will imply that the sum of conditional expectations (Eq. (E.2)) is bounded from zero uniformly in $||\theta-\theta_0||$. Thus, initial summation (Eq. (E.1)) is also uniformly bounded from zero, so that as T goes to infinity, for any $\theta \neq \theta_0$ with probability tending to one we have

$$\frac{1}{T}\sum_t \mathbb{E}\left(|y_t - [x_t'\theta]_+| - |y_t - [x_t'\theta_0]_+|\,|\mathcal{F}_{t-1}\right) \geq const(||\theta-\theta_0||) > 0$$

and

$$\lim_{T\to\infty} \mathbb{P}(Q_T(\theta) \geq const(||\theta-\theta_0||)) = 1,$$

where constant is increasing in $||\theta-\theta_0||$.

In contrast, for any T, $Q_T(\theta_0) = 0$. Thus, we must have that LAD estimate $\hat{\theta}_T$ converges to θ_0 as $T \to \infty$, as $Q_T(\theta)$ is bounded from zero for $\theta \neq \theta_0$ and T large enough.

Let us show that $\mathbb{P}(|x'(\theta-\theta_0)| \geq \tau_0|\, x'\theta_0 \geq R) \geq C_1 > 0$, where $\tau_0 = const \cdot ||\theta-\theta_0||^2$. Define by $\lambda_{\min}$ the minimal eigenvalue of matrix $\mathbb{E}(xx'|x'\theta_0 \geq R)$. It is non-zero, because the matrix is nonsingular. Then

$$\mathbb{E}\left(|x'(\theta-\theta_0)|^2\,|x'\theta_0 \geq R\right) \geq ||\theta-\theta_0||^2\lambda_{\min},$$

as such conditional expectation corresponds to the value of the quadratic form $\mathbb{E}(xx'|x'\theta_0 > R)$ on the vector $\theta-\theta_0$.

[2]Because an indicator is a discontinuous function, LLN is not guaranteed for arbitrary R, τ. However, we can always achieve it by slightly decreasing R and τ, so that we end up in the point of continuity of the distribution function. Matrix $M_{R'}$ is still nonsingular for $R' < R$ (see the paragraph after Eq. (7)).

Choose $\varepsilon > 0$ such that $\varepsilon < ||\theta - \theta_0||^2 \lambda_{\min}$ and note that for any $A > 0$

$$\begin{aligned}
&\mathbb{E}\left(|x'(\theta-\theta_0)|^2 \, |x'\theta_0 \geq R\right) = \mathbb{E}\left(|x'(\theta-\theta_0)|^2 \mathbf{1}(|x'(\theta-\theta_0)|^2 < A) \, |x'\theta_0 \geq R\right) \\
&+ \mathbb{E}\left(|x'(\theta-\theta_0)|^2 \mathbf{1}(|x'(\theta-\theta_0)|^2 \geq A) \, |x'\theta_0 \geq R\right) \\
&\leq \mathbb{E}\left(|x'(\theta-\theta_0)|^2 \mathbf{1}(|x'(\theta-\theta_0)|^2 < A) \, |x'\theta_0 \geq R\right) \\
&+ \mathbb{E}\left(\frac{|x'(\theta-\theta_0)|^4}{A} \mathbf{1}(|x'(\theta-\theta_0)|^2 \geq A) \, |x'\theta_0 \geq R\right) \\
&\leq \mathbb{E}\left(|x'(\theta-\theta_0)|^2 \mathbf{1}(|x'(\theta-\theta_0)|^2 < A) \, |x'\theta_0 \geq R\right) + \frac{\mathbb{E}\left(|x'(\theta-\theta_0)|^4 \, |x'\theta_0 \geq R\right)}{A},
\end{aligned}$$

where the fourth moment exists, because u_t has fourth moment. Choose $A(\varepsilon)$ such that $A(\varepsilon) > \frac{||\theta-\theta_0||^4 \mathbb{E}\left(||x||^4 \, |x'\theta_0 \geq R\right)}{\varepsilon}$. Then

$$\frac{\mathbb{E}\left(|x'(\theta-\theta_0)|^4 \, |x'\theta_0 \geq R\right)}{A} < \frac{\varepsilon \mathbb{E}\left(|x'(\theta-\theta_0)|^4 \, |x'\theta_0 \geq R\right)}{||\theta-\theta_0||^4 \mathbb{E}\left(||x||^4 \, |x'\theta_0 \geq R\right)} \leq \varepsilon,$$

as $\mathbb{E}\left(|x'(\theta-\theta_0)|^4 \, |x'\theta_0 \geq R\right) \leq ||\theta-\theta_0||^4 \mathbb{E}\left(||x||^4 \, |x'\theta_0 \geq R\right)$. Thus,

$$\begin{aligned}
\mathbb{E}\left(|x'(\theta-\theta_0)|^2 \mathbf{1}(|x'(\theta-\theta_0)|^2 < A(\varepsilon)) \, |x'\theta_0 \geq R\right) &\geq \mathbb{E}\left(|x'(\theta-\theta_0)|^2 \, |x'\theta_0 \geq R\right) - \varepsilon \\
&\geq ||\theta-\theta_0||^2 \lambda_{\min} - \varepsilon > 0.
\end{aligned}$$

Finally,

$$\begin{aligned}
&\mathbb{E}\left(|x'(\theta-\theta_0)|^2 \mathbf{1}(|x'(\theta-\theta_0)|^2 < A(\varepsilon)) \, |x'\theta_0 \geq R\right) \\
&= \mathbb{E}\left(|x'(\theta-\theta_0)|^2 \mathbf{1}(|x'(\theta-\theta_0)|^2 < A(\varepsilon)) \mathbf{1}(|x'(\theta-\theta_0)| \geq \tau) \, |x'\theta_0 \geq R\right) \\
&+ \mathbb{E}\left(|x'(\theta-\theta_0)|^2 \mathbf{1}(|x'(\theta-\theta_0)|^2 < A(\varepsilon)) \mathbf{1}(|x'(\theta-\theta_0)| < \tau) \, |x'\theta_0 \geq R\right) \\
&\leq A(\varepsilon) \mathbb{P}(|x'(\theta-\theta_0)| \geq \tau | \, x'\theta_0 \geq R) + \tau,
\end{aligned}$$

so that for $\tau \in (0, ||\theta-\theta_0||^2 \lambda_{\min} - \varepsilon)$,

$$\mathbb{P}(|x'(\theta-\theta_0)| \geq \tau | \, x'\theta_0 \geq R) \geq \frac{||\theta-\theta_0||^2 \lambda_{\min} - \varepsilon - \tau}{A(\varepsilon)} = C_1 > 0.$$

Fixing $\varepsilon = \frac{1}{2}||\theta-\theta_0||^2 \lambda_{\min}$, $A(\varepsilon) = 4 \frac{||\theta-\theta_0||^2 \mathbb{E}\left(||x||^4 \, |x'\theta_0 \geq R\right)}{\lambda_{\min}}$, $\tau = \frac{1}{4}||\theta-\theta_0||^2 \lambda_{\min}$. Then

$$\mathbb{P}(|x'(\theta-\theta_0)| \geq \tau | \, x'\theta_0 \geq R) \geq \frac{\lambda_{\min}^2}{16 \mathbb{E}\left(||x||^4 | x'\theta_0 \geq R\right)} = C_1 > 0.$$

If $||\theta' - \theta_0|| > ||\theta - \theta_0||$, then

$$\begin{aligned}
&\mathbb{P}(|x'(\theta'-\theta_0)| \geq \frac{1}{4}||\theta-\theta_0||^2 \lambda_{\min} | \, x'\theta_0 \geq R) \\
&\geq \mathbb{P}(|x'(\theta'-\theta_0)| \geq \frac{1}{4}||\theta'-\theta_0||^2 \lambda_{\min} | \, x'\theta_0 \geq R) \geq \frac{\lambda_{\min}^2}{16 \mathbb{E}\left(||x_t||^4 | x'\theta_0 \geq R\right)}.
\end{aligned}$$

□

Proof of Theorem 7. The minimization problem (6) is convex, thus, if we show that if has a unique solution near the true values of (α, β, γ), then this is the only solution (i.e. global minimal).

Denote the LAD estimate as $(\hat{\alpha}, \hat{\beta}, \hat{\gamma})$. The estimate is the solution to minimization problem (6). Thus, $(\hat{\alpha}, \hat{\beta}, \hat{\gamma})$ solves $h(a, b, c) = (0, 0, 0)'$, where $h : \mathbb{R}^3 \to \mathbb{R}^3$ is the set of first order conditions:

$$h(a,b,c) = \frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix} 1 \\ y_{t-1} \\ z_{t-1} \end{pmatrix} \operatorname{sgn}\left(y_t - a - by_{t-1} - cz_{t-1}\right). \tag{E.3}$$

Fix some (a, b, c) and define $\Delta\alpha = a - \alpha$, $\Delta\beta = b - \beta$, $\Delta\gamma = c - \gamma$, $v_{t+1} = \max(u_{t+1}, -\alpha - \beta y_t - \gamma z_t)$. Then, noting that $y_t = \alpha + \beta y_{t-1} + \gamma z_{t-1} + v_t$, we can rewrite Eq. (E.3) as

$$h(a,b,c) = H(\Delta\alpha, \Delta\beta, \Delta\gamma) = \frac{1}{T}\sum_{t}\begin{pmatrix} 1 \\ y_{t-1} \\ z_{t-1} \end{pmatrix} \operatorname{sgn}\left(v_t - \Delta\alpha - \Delta\beta y_{t-1} - \Delta\gamma z_{t-1}\right). \tag{E.4}$$

Our aim is to show that equation $H(\Delta\alpha, \Delta\beta, \Delta\gamma) = (0, 0, 0)'$ has a solution which satisfies $\Delta\alpha = o(1)$, $\Delta\beta = o(1)$, $\Delta\gamma = o(1)$ as $T \to \infty$. For that we are going to study the function $H(\cdot)$, treating as $\Delta\alpha, \Delta\beta$, and $\Delta\gamma$ as numbers, not as random variables.

By the law of large numbers,[3] we get

$$H(\Delta\alpha, \Delta\beta, \Delta\gamma) = \mathbb{E}\begin{pmatrix} 1 \\ y \\ z \end{pmatrix} \operatorname{sgn}\left(v - \Delta\alpha - \Delta\beta y - \Delta\gamma z\right) + o(1), \tag{E.5}$$

where $v = \max(u, -\alpha - \beta y - \gamma z)$ and u is independent from y and z (i.e. u corresponds to the next period).

Let us linearize the expectations in Eq. (E.5).

$$\begin{aligned} &\mathbb{E}\operatorname{sgn}\left(v - \Delta\alpha - \Delta\beta y - \Delta\gamma z\right) \\ &= \int_y \int_z \int_v \operatorname{sgn}\left(v - \Delta\alpha - \Delta\beta y - \Delta\gamma z\right) f_{y,z,v}(y,z,v)\,dv\,dz\,dy \\ &= \int_y \int_z \int\limits_{\Delta\alpha+\Delta\beta y+\Delta\gamma z}^{\infty} f_{y,z,v}(y,z,v)\,dv\,dz\,dy - \int_y \int_z \int\limits_{-\infty}^{\Delta\alpha+\Delta\beta y+\Delta\gamma z} f_{y,z,v}(y,z,v)\,dv\,dz\,dy \\ &= 1 - 2\mathbb{E}_{y,z} F_{v|y,z}(\Delta\alpha + \Delta\beta y + \Delta\gamma z | y, z), \end{aligned} \tag{E.6}$$

where $\mathbb{E}_{y,z}$ stays for the expectation with respect to random variables y and z (stationary limits of y_t and z_t).

[3] LLN follows from the fact that y_t and z_t are mixing and can be formally proved in the spirit of Lemma E.1

Taylor expanding around $\Delta\alpha + \Delta\beta y + \Delta\gamma z = 0$, we can rewrite Eq. (E.6) as

$$1 - 2\mathbb{E}_{y,z}\left(F_{v|y,z}(0|y,z) + f_{v|y,z}(0|y,z)(\Delta\alpha + \Delta\beta y + \Delta\gamma z|y,z) + o(\Delta\alpha + \Delta\beta y + \Delta\gamma z)\right).$$

As $v = \max(u, -\alpha - \beta y - \gamma z)$ and $-\alpha - \beta y - \gamma z < 0$ for $\alpha > 0$, $\beta \geq 0$, $\gamma \geq 0$, the density of v has a mass point at $-\alpha - \beta y - \gamma z$ and then coincides with the density of u. Thus, $F_{v|y,z}(0|y,z) = F_u(0) = 0.5$ and $f_{v|y,z}(0|y,z) = f_u(0)$. To sum up, the first order condition with respect to α becomes

$$-2f_u(0)\mathbb{E}(\Delta\alpha + \Delta\beta y + \Delta\gamma z + o(\Delta\alpha + \Delta\beta y + \Delta\gamma z)) = o(1). \tag{E.7}$$

Let us now analyze the second line of Eq. (E.5) in a similar fashion.

(E.8)

$$\begin{aligned}
&\mathbb{E}y\mathrm{sgn}\left(v - \Delta\alpha - \Delta\beta y - \Delta\gamma z\right) = \int_y\int_z\int_v y\mathrm{sgn}\left(v - \Delta\alpha - \Delta\beta y - \Delta\gamma z\right) f_{y,z,v}(y,z,v)dvdzdy\\
&= \mathbb{E}y - 2\mathbb{E}_{y,z}yF_{v|y,z}(\Delta\alpha + \Delta\beta y + \Delta\gamma z|y,z)\\
&= \mathbb{E}y - 2\mathbb{E}_{y,z}y\left(F_{v|y,z}(0|y,z) + f_{v|y,z}(0|y,z)(\Delta\alpha + \Delta\beta y + \Delta\gamma z|y,z) + o(\Delta\alpha + \Delta\beta y + \Delta\gamma z)\right)\\
&= -2f_u(0)\mathbb{E}y(\Delta\alpha + \Delta\beta y + \Delta\gamma z + o(\Delta\alpha + \Delta\beta y + \Delta\gamma z)).
\end{aligned}$$

Thus, first order condition with respect to β becomes

$$-2f_u(0)\mathbb{E}y(\Delta\alpha + \Delta\beta y + \Delta\gamma z + o(\Delta\alpha + \Delta\beta y + \Delta\gamma z)) = o(1). \tag{E.9}$$

Similarly, the first order condition with respect to γ becomes

$$-2f_u(0)\mathbb{E}z(\Delta\alpha + \Delta\beta y + \Delta\gamma z + o(\Delta\alpha + \Delta\beta y + \Delta\gamma z)) = o(1). \tag{E.10}$$

Combining Eq. (E.7), (E.9), and (E.10), we get

$$\begin{pmatrix} 1 & \mathbb{E}y & \mathbb{E}z \\ \mathbb{E}y & \mathbb{E}yz & \mathbb{E}yz \\ \mathbb{E}z & \mathbb{E}yz & \mathbb{E}z^2 \end{pmatrix}\begin{pmatrix} \Delta\alpha \\ \Delta\beta \\ \Delta\gamma \end{pmatrix} = \begin{pmatrix} o(1) + o(\Delta\alpha) + o(\Delta\beta) + o(\Delta\gamma) \\ o(1) + o(\Delta\alpha) + o(\Delta\beta) + o(\Delta\gamma) \\ o(1) + o(\Delta\alpha) + o(\Delta\beta) + o(\Delta\gamma) \end{pmatrix}. \tag{E.11}$$

Matrix $\begin{pmatrix} 1 & \mathbb{E}y & \mathbb{E}z \\ \mathbb{E}y & \mathbb{E}yz & \mathbb{E}yz \\ \mathbb{E}z & \mathbb{E}yz & \mathbb{E}z^2 \end{pmatrix}$ is nonsingular. If it were singular, it would have an eigenvector $(\lambda_1, \lambda_2, \lambda_2)$ corresponding to the zero eigenvalue. Then the random variable $\lambda_1 + \lambda_2 y_t + \lambda_3 z_t$ must have zero second moment. That is, $\lambda_1 + \lambda_2 y_t + \lambda_3 z_t$ must be zero, which contradicts conditions of the theorem.

Thus, Eq. (E.11) has the solution which satisfies $\Delta\alpha = o(1)$, $\Delta\beta = o(1)$, $\Delta\gamma = o(1)$. Moreover, the solution which satisfies $\Delta\alpha = o(1)$, $\Delta\beta = o(1)$, $\Delta\gamma = o(1)$ is unique. That is, the LAD estimator is consistent and converges to the true value of the parameters as $T \to \infty$. □

Proof of Theorem 8. Denote the LAD estimate as $(\hat{\alpha}, \hat{\beta}, \hat{\gamma})$. Because LAD is consistent (Theorem 7), $\hat{\alpha} = \alpha + o(1), \hat{\beta} = \beta + o(1), \hat{\gamma} = \gamma + o(1)$. We are going to use results from the proof of Theorem 7. The difference is that now we will apply the martingale central limit theorem instead of the law of large numbers (see e.g., Hall and Heyde (1980, Section 3) for the formulation of the martingale CLT). Define the filtration $\mathcal{F}_t = \{y_t, z_t, y_{t-1}, z_{t-1}, \ldots\}$. Then

$$\xi_t := \begin{pmatrix} \operatorname{sgn}(v_{t+1} - \Delta\alpha - \Delta\beta y_t - \Delta\gamma z_t) - \mathbb{E}\left(\operatorname{sgn}(v_{t+1} - \Delta\alpha - \Delta\beta y_t - \Delta\gamma z_t) | \mathcal{F}_t\right) \\ y_t \operatorname{sgn}(v_{t+1} - \Delta\alpha - \Delta\beta y_t - \Delta\gamma z_t) - \mathbb{E}\left(y_t \operatorname{sgn}(v_{t+1} - \Delta\alpha - \Delta\beta y_t - \Delta\gamma z_t) | \mathcal{F}_t\right) \\ z_t \operatorname{sgn}(v_{t+1} - \Delta\alpha - \Delta\beta y_t - \Delta\gamma z_t) - \mathbb{E}\left(z_t \operatorname{sgn}(v_{t+1} - \Delta\alpha - \Delta\beta y_t - \Delta\gamma z_t) | \mathcal{F}_t\right) \end{pmatrix}$$

is a martingale difference with respect to filtration $\mathcal{F}_t$.

Let us calculate the corresponding asymptotic covariance matrix. First, observe that

$$\mathbb{E}\left(\operatorname{sgn}(v_{t+1} - \Delta\alpha - \Delta\beta y_t - \Delta\gamma z_t) | \mathcal{F}_t\right) = 1 - 2\mathbb{P}_v\left(v_{t+1} < \Delta\alpha + \Delta\beta y_t + \Delta\gamma z_t\right).$$

When $\Delta\alpha + \Delta\beta y_t + \Delta\gamma z_t = 0$, the probability that v_{t+1} is smaller than $\Delta\alpha + \Delta\beta y_t + \Delta\gamma z_t$ equals to $1/2$, as $\operatorname{med}(v) = 0$. Thus, $\mathbb{E}\left(\operatorname{sgn}(v_{t+1} - \Delta\alpha - \Delta\beta y_t - \Delta\gamma z_t) | \mathcal{F}_t\right) = o(1)$ for $\Delta\alpha = o(1)$, $\Delta\beta = 0(1)$, $\Delta\gamma = o(1)$.

Second, note that $\operatorname{sgn}^2(v_{t+1} - \Delta\alpha - \Delta\beta y_t - \Delta\gamma z_t) \equiv 1$, so that

- $\mathbb{E}\left(y_t \operatorname{sgn}^2(v_{t+1} - \Delta\alpha - \Delta\beta y_t - \Delta\gamma z_t) | \mathcal{F}_t\right) \equiv y_t$,
- $\mathbb{E}\left(z_t \operatorname{sgn}^2(v_{t+1} - \Delta\alpha - \Delta\beta y_t - \Delta\gamma z_t) | \mathcal{F}_t\right) \equiv z_t$,
- $\mathbb{E}\left(y_t z_t \operatorname{sgn}^2(v_{t+1} - \Delta\alpha - \Delta\beta y_t - \Delta\gamma z_t) | \mathcal{F}_t\right) \equiv y_t z_t$,
- $\mathbb{V}\left(\operatorname{sgn}(v_{t+1} - \Delta\alpha - \Delta\beta y_t - \Delta\gamma z_t) | \mathcal{F}_t\right) = 1 - o(1)$,
- $\mathbb{V}\left(y_t \operatorname{sgn}(v_{t+1} - \Delta\alpha - \Delta\beta y_t - \Delta\gamma z_t) | \mathcal{F}_t\right) = y_t^2(1 - o(1))$,
- $\mathbb{V}\left(z_t \operatorname{sgn}(v_{t+1} - \Delta\alpha - \Delta\beta y_t - \Delta\gamma z_t) | \mathcal{F}_t\right) = z_t^2(1 - o(1))$.

Under stationarity $\frac{1}{T}\sum_t y_t \to \mathbb{E}y$, $\frac{1}{T}\sum_t z_t \to \mathbb{E}z$, $\frac{1}{T}\sum_t y_t^2 \to \mathbb{E}y^2$, $\frac{1}{T}\sum_t y_t z_t \to \mathbb{E}yz$, $\frac{1}{T}\sum_t z_t^2 \to \mathbb{E}z^2$. We want to apply the martingale CLT. Thus, we need to check the Lindeberg condition. Because

$$\sum_{t=1}^{T} \mathbb{V}\left(\frac{1}{\sqrt{T}}\xi_t | \mathcal{F}_t\right) = \begin{pmatrix} 1 - o(1) & \frac{1}{T}\sum_t y_t & \frac{1}{T}\sum_t z_t \\ \frac{1}{T}\sum_t y_t & \frac{1-o(1)}{T}\sum_t y_t^2 & \frac{1}{T}\sum_t y_t z_t \\ \frac{1}{T}\sum_t z_t & \frac{1}{T}\sum_t y_t z_t & \frac{1-o(1)}{T}\sum_t z_t^2 \end{pmatrix},$$

the variance matrix is $O(1)$. So we need to check that for any numbers λ_1, λ_2, λ_3 and $\varepsilon > 0$

$$\sum_{t=1}^{T} \mathbb{E}\left(\frac{1}{T}(\lambda_1\xi_{1t} + \lambda_2\xi_{2t} + \lambda_3\xi_{3t})^2 \mathbf{1}\left(\frac{1}{\sqrt{T}}|\lambda_1\xi_{1t} + \lambda_2\xi_{2t} + \lambda_3\xi_{3t}| > \varepsilon\right)\right) \xrightarrow[T\to\infty]{} 0. \tag{E.12}$$

Because y_t and z_t are stationary, Eq. (E.12) reduces to

$$\mathbb{E}(\lambda_1\xi_{1t}+\lambda_2\xi_{2t}+\lambda_3\xi_{3t})^2\mathbf{1}\left(\frac{1}{\sqrt{T}}|\lambda_1\xi_{1t}+\lambda_2\xi_{2t}+\lambda_3\xi_{3t}|>\varepsilon\right)\xrightarrow[T\to\infty]{}0,$$

which is true as $(1, y_t, z_t)$ and, thus, $\xi_t = (\xi_{1t}, \xi_{2t}, \xi_{3t})'$ and any linear combination of ξ_t's coordinates, have finite second moments.

So by the martingale CLT with the Lindeberg condition,

(E.13)

$$\frac{1}{\sqrt{T}}\begin{pmatrix}\sum_t\left[\operatorname{sgn}\left(v_{t+1}-\Delta\alpha-\Delta\beta y_t-\Delta\gamma z_t\right)-\mathbb{E}\left(\operatorname{sgn}\left(v_{t+1}-\Delta\alpha-\Delta\beta y_t-\Delta\gamma z_t\right)|\mathcal{F}_t\right)\right]\\ \sum_t\left[y_t\operatorname{sgn}\left(v_{t+1}-\Delta\alpha-\Delta\beta y_t-\Delta\gamma z_t\right)-\mathbb{E}\left(y_t\operatorname{sgn}\left(v_{t+1}-\Delta\alpha-\Delta\beta y_t-\Delta\gamma z_t\right)|\mathcal{F}_t\right)\right]\\ \sum_t\left[z_t\operatorname{sgn}\left(v_{t+1}-\Delta\alpha-\Delta\beta y_t-\Delta\gamma z_t\right)-\mathbb{E}\left(z_t\operatorname{sgn}\left(v_{t+1}-\Delta\alpha-\Delta\beta y_t-\Delta\gamma z_t\right)|\mathcal{F}_t\right)\right]\end{pmatrix}$$
$$\xrightarrow[T\to\infty]{d}\mathcal{N}\left(\begin{pmatrix}0\\0\\0\end{pmatrix},\begin{pmatrix}1 & \mathbb{E}y & \mathbb{E}z\\ \mathbb{E}y & \mathbb{E}y^2 & \mathbb{E}yz\\ \mathbb{E}z & \mathbb{E}yz & \mathbb{E}z^2\end{pmatrix}\right).$$

By Eq. (E.13), first order conditions $H(\Delta\alpha, \Delta\beta, \Delta\gamma) = (0, 0, 0)'$ (see Eq. (E.4)) can be rewritten as

(E.14)

$$\mathcal{N}\left(\begin{pmatrix}0\\0\\0\end{pmatrix},\begin{pmatrix}1 & \mathbb{E}y & \mathbb{E}z\\ \mathbb{E}y & \mathbb{E}y^2 & \mathbb{E}yz\\ \mathbb{E}z & \mathbb{E}yz & \mathbb{E}z^2\end{pmatrix}\right)+o(1)$$
$$+\frac{1}{\sqrt{T}}\begin{pmatrix}\sum_t\mathbb{E}\left(\operatorname{sgn}\left(v_{t+1}-\Delta\alpha-\Delta\beta y_t-\Delta\gamma z_t\right)|\mathcal{F}_t\right)\\ \sum_t\mathbb{E}\left(y_t\operatorname{sgn}\left(v_{t+1}-\Delta\alpha-\Delta\beta y_t-\Delta\gamma z_t\right)|\mathcal{F}_t\right)\\ \sum_t\mathbb{E}\left(z_t\operatorname{sgn}\left(v_{t+1}-\Delta\alpha-\Delta\beta y_t-\Delta\gamma z_t\right)|\mathcal{F}_t\right)\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}.$$

Following the linearization argument in the proof of theorem 7, where we need to replace stationary v, y, z with v_{t+1}, y_t, z_t, we get

(E.15)

$$\begin{aligned}\mathbb{E}\left(\operatorname{sgn}\left(v_{t+1}-\Delta\alpha-\Delta\beta y_t-\Delta\gamma z_t\right)|\mathcal{F}_t\right)&=-2f_u(0)(\Delta\alpha+\Delta\beta y_t+\Delta\gamma z_t)\\&\quad+o(\Delta\alpha)+o(\Delta\beta)+o(\Delta\gamma),\\ \mathbb{E}\left(y_t\operatorname{sgn}\left(v_{t+1}-\Delta\alpha-\Delta\beta y_t-\Delta\gamma z_t\right)|\mathcal{F}_t\right)&=-2f_u(0)y_t(\Delta\alpha+\Delta\beta y_t+\Delta\gamma z_t)\\&\quad+o(\Delta\alpha)+o(\Delta\beta)+o(\Delta\gamma),\\ \mathbb{E}\left(z_t\operatorname{sgn}\left(v_{t+1}-\Delta\alpha-\Delta\beta y_t-\Delta\gamma z_t\right)|\mathcal{F}_t\right)&=-2f_u(0)z_t(\Delta\alpha+\Delta\beta y_t+\Delta\gamma z_t)\\&\quad+o(\Delta\alpha)+o(\Delta\beta)+o(\Delta\gamma).\end{aligned}$$

Plugging Eq. (E.15) into Eq. (E.14), we get

$$\mathcal{N}\left(\begin{pmatrix}0\\0\\0\end{pmatrix},\begin{pmatrix}1 & \mathbb{E}y & \mathbb{E}z\\ \mathbb{E}y & \mathbb{E}y^2 & \mathbb{E}yz\\ \mathbb{E}z & \mathbb{E}yz & \mathbb{E}z^2\end{pmatrix}\right)-\frac{2f_u(0)}{\sqrt{T}}\sum_t\begin{pmatrix}\Delta\alpha+\Delta\beta y_t+\Delta\gamma z_t\\ \Delta\alpha y_t+\Delta\beta y_t^2+\Delta\gamma y_t z_t\\ \Delta\alpha z_t+\Delta\beta y_t z_t+\Delta\gamma z_t^2\end{pmatrix}$$
$$=o(1)+\sqrt{T}\left(o(\Delta\alpha)+o(\Delta\beta)+o(\Delta\gamma)\right). \tag{E.16}$$

Because $y_t \xrightarrow[t\to\infty]{d} y$, we have $\frac{1}{T}\sum_t y_t = \mathbb{E}y + o(1)$, and similarly with other terms in the summations in Eq. (E.16). Thus,

$$2\sqrt{T}f_u(0)\begin{pmatrix}1 & \mathbb{E}y & \mathbb{E}z\\ \mathbb{E}y & \mathbb{E}y^2 & \mathbb{E}yz\\ \mathbb{E}z & \mathbb{E}yz & \mathbb{E}z^2\end{pmatrix}\begin{pmatrix}\Delta\alpha\\ \Delta\beta\\ \Delta\gamma\end{pmatrix}$$
$$=\mathcal{N}\left(\begin{pmatrix}0\\0\\0\end{pmatrix},\begin{pmatrix}1 & \mathbb{E}y & \mathbb{E}z\\ \mathbb{E}y & \mathbb{E}y^2 & \mathbb{E}yz\\ \mathbb{E}z & \mathbb{E}yz & \mathbb{E}z^2\end{pmatrix}\right)+o(1)+\sqrt{T}\left(o(\Delta\alpha)+o(\Delta\beta)+o(\Delta\gamma)\right),$$

or, equivalently,

$$\sqrt{T}\begin{pmatrix}\Delta\alpha\\ \Delta\beta\\ \Delta\gamma\end{pmatrix}=\mathcal{N}\left(\begin{pmatrix}0\\0\\0\end{pmatrix},\frac{1}{4f_u^2(0)}\begin{pmatrix}1 & \mathbb{E}y & \mathbb{E}z\\ \mathbb{E}y & \mathbb{E}y^2 & \mathbb{E}yz\\ \mathbb{E}z & \mathbb{E}yz & \mathbb{E}z^2\end{pmatrix}^{-1}\right)$$
$$+o(1)+\sqrt{T}\left(o(\Delta\alpha)+o(\Delta\beta)+o(\Delta\gamma)\right).$$

The solution to the above equation satisfies

$$\Delta\alpha=O(T^{-1/2}),\Delta\beta=O(T^{-1/2}),\Delta\gamma=O(T^{-1/2})$$

and

$$\sqrt{T}\begin{pmatrix}\Delta\alpha\\ \Delta\beta\\ \Delta\gamma\end{pmatrix}=\sqrt{T}\begin{pmatrix}\hat{\alpha}-\alpha\\ \hat{\beta}-\beta\\ \hat{\gamma}-\gamma\end{pmatrix}\xrightarrow[T\to\infty]{d}\mathcal{N}\left(\begin{pmatrix}0\\0\\0\end{pmatrix},\frac{1}{4f_u^2(0)}\begin{pmatrix}1 & \mathbb{E}y & \mathbb{E}z\\ \mathbb{E}y & \mathbb{E}y^2 & \mathbb{E}yz\\ \mathbb{E}z & \mathbb{E}yz & \mathbb{E}z^2\end{pmatrix}^{-1}\right). \quad \square$$

Proof of Theorem 6. We again proceed as in the proof of Theorem 7 and replace sums with expectations and then linearize expectations. The first order conditions are $h(a,b,c) = (0,0,0)'$, where

$$h(a,b,c)=\sum_{t=1}^{T}\begin{pmatrix}1\\ y_{t-1}\\ z_{t-1}\end{pmatrix}\operatorname{sgn}\left(y_t-a-by_{t-1}-cz_{t-1}\right)\mathbf{1}(a+by_{t-1}+cz_{t-1}>0).$$

Fix some a, b, c and denote $\Delta\alpha = a - \alpha$, $\Delta\beta = b - \beta$, $\Delta\gamma = c - \gamma$, and

$$\begin{aligned} H(\Delta\alpha, \Delta\beta, \Delta\gamma) =& h(a,b,c) \\ =& \frac{1}{T}\sum_t \begin{pmatrix} 1 \\ y_{t-1} \\ z_{t-1} \end{pmatrix} \operatorname{sgn}\left(v_t - \Delta\alpha - \Delta\beta y_{t-1} - \Delta\gamma z_{t-1}\right) \\ & \cdot \mathbf{1}(\alpha + \beta y_{t-1} + \gamma z_{t-1} + \Delta\alpha + \Delta\beta y_{t-1} + \Delta\gamma z_{t-1}), \end{aligned}$$

where $v_{t+1} = \max(u_{t+1}, -\alpha - \beta y_t - \gamma z_t)$. We know form Theorem 5 that the solution to $H(\Delta\alpha, \Delta\beta, \Delta\gamma) = (0,0,0)'$ satisfies $\Delta\alpha = o(1), \Delta\beta = o(1), \Delta\gamma = o(1)$.

We now follow the lines of Theorem 8: we add and subtract conditional expectations from each term in $H(\Delta\alpha, \Delta\beta, \Delta\gamma)$. We then apply the martingale CLT for the differences and linearize the expectations. The only difference is that everything is multiplied by an indicator $\mathbf{1}(\alpha + \beta y_{t-1} + \gamma z_{t-1} + \Delta\alpha + \Delta\beta y_{t-1} + \Delta\gamma z_{t-1} > 0)$. Let us show how the linearization works in the presence of the indicator.

Because signum takes only values -1 and 1,

$$\begin{aligned} &\mathbb{E}\left[\begin{pmatrix} 1 \\ y_t \\ z_t \end{pmatrix} \operatorname{sgn}\left(y_{t+1} - a - by_t - cz_t\right) \mathbf{1}(a + by_t + cz_t > 0) \middle| \mathcal{F}_t \right] \\ &= \begin{pmatrix} 1 \\ y_t \\ z_t \end{pmatrix} \left(1 - 2F_{v_{t+1}|y_t,z_t}(\Delta\alpha + \Delta\beta y_t + \Delta\gamma z_t | y_t, z_t)\right) \mathbf{1}(\alpha + \beta y_t + \gamma z_t + \Delta\alpha + \Delta\beta y_t + \Delta\gamma z_t > 0). \end{aligned} \tag{E.17}$$

Instead of Taylor expansion, we are going to use the direct formula: $f(x+\Delta x)g(x+\Delta x) = f(x)g(x) + f(x)(g(x+\Delta x) - g(x)) + g(x+\Delta x)(f(x+\Delta x) - f(x))$. Thus, the first expectation in Eq. (E.17) can be rewritten as

$$\begin{aligned} &(1 - 2F_{v_{t+1}|y_t,z_t}(0|y_t, z_t))\mathbf{1}(\alpha + \beta y_t + \gamma z_t > 0) \\ &- 2f_{v_{t+1}|y_t,z_t}(0|y_t, z_t)(\Delta\alpha + \Delta\beta y_t + \Delta\gamma z_t)\mathbf{1}(\alpha + \beta y_t + \gamma z_t + \Delta\alpha + \Delta\beta y_t + \Delta\gamma z_t > 0) \\ &+ (1 - 2F_{v_{t+1}|y_t,z_t} 0|y_t, z_t))\Big(\mathbf{1}(\alpha + \beta y_t + \gamma z_t + \Delta\alpha + \Delta\beta y_t + \Delta\gamma z_t > 0) \\ &- \mathbf{1}(\alpha + \beta y_t + \gamma z_t > 0)\Big) + o(\Delta\alpha + \Delta\beta y_t + \Delta\gamma z_t) \\ &= -2f_u(0)(\Delta\alpha + \Delta\beta y_t + \Delta\gamma z_t)\mathbf{1}(\alpha + \beta y_t + \gamma z_t > 0) + o(\Delta\alpha) + o(\Delta\beta) + o(\Delta\gamma), \end{aligned} \tag{E.18}$$

where we used the fact that $F_{v_{t+1}|y_t,z_t}(0|y_t, z_t) = F_u(0) = 0.5$ and $f_{v_{t+1}|y_t,z_t}(0|y_t, z_t) = f_u(0)$ when $\alpha + \beta y_t + \gamma z_t > 0$, and continuity of $\mathbf{1}(\alpha + \beta y_t + \gamma z_t + \Delta\alpha + \Delta\beta y_t + \Delta\gamma z_t > 0)$ with respect to $\Delta\alpha + \Delta\beta y_t + \Delta\gamma z_t = 0$.

Similarly, the other two expectations in Eq. (E.17) can be rewritten as

$$-2f_u(0)y_t(\Delta\alpha + \Delta\beta y_t + \Delta\gamma z_t)\mathbf{1}(\alpha + \beta y_t + \gamma z_t > 0) + o(\Delta\alpha) + o(\Delta\beta) + o(\Delta\gamma)$$

and

$$-2f_u(0)z_t(\Delta\alpha + \Delta\beta y_t + \Delta\gamma z_t)\mathbf{1}(\alpha + \beta y_t + \gamma z_t > 0) + o(\Delta\alpha) + o(\Delta\beta) + o(\Delta\gamma).$$

Thus, duplicating the arguments from the proof of Theorem 8,

$$\sqrt{T}\begin{pmatrix} \Delta\alpha \\ \Delta\beta \\ \Delta\gamma \end{pmatrix} = \sqrt{T}\begin{pmatrix} \hat{\alpha}-\alpha \\ \hat{\beta}-\beta \\ \hat{\gamma}-\gamma \end{pmatrix}$$

$$\xrightarrow[T\to\infty]{d} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{4f_u^2(0)}\left(\mathbb{E}\left[\begin{pmatrix} 1 & y & z \\ y & y^2 & yz \\ z & yz & z^2 \end{pmatrix}\mathbf{1}(\alpha + \beta y + \gamma z > 0)\right]\right)^{-1}\right). \qquad \square$$

E.2. Explosive LAD.

Lemma E.2. *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

$$y_{t+k} \geq \beta^k \left(y_t - \sum_{i=1}^{k} v_{t+i-1}\beta^{-i}\right) \tag{E.19}$$

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 (E.19):

$$y_{T+k} \geq \beta^k (y_T/2), \qquad \text{for all } k = 1, 2, \ldots. \tag{E.20}$$

Let us call the event where (E.20) 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

$$y_{t+1} \geq \beta y_t - v_t = \beta' y_t + (\beta - \beta') y_t \left(1 - \frac{v_t}{y_t}\right) \tag{E.21}$$

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 (E.21) 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. □

Proof of Theorem 9. Denote the LAD estimate as $(\hat{\alpha}, \hat{\beta})$. The estimate is the solution to minimization problem

$$\min_{a,b} \sum_t |y_{t+1} - [a + by_t]_+| \, . \tag{E.22}$$

The corresponding first order conditions are

$$\sum_t \operatorname{sgn}(y_{t+1} - a - by_t)\, \mathbf{1}(a + by_t > 0) = 0, \tag{E.23}$$

$$\sum_t y_t \operatorname{sgn}(y_{t+1} - a - by_t)\, \mathbf{1}(a + by_t > 0) = 0. \tag{E.24}$$

Note that formally the summations may not equal to zero, as they involve discrete increments. Thus, we need to find point, where the summations switch signs from minus to plus.

The proof differs depending on the behavior of y_t ($\beta > 1$ or $\beta = 1$, $\alpha + \mathbb{E}u_t > 0$ or $\beta = 1$, $\alpha + \mathbb{E}u_t = 0$).

Let us analyze Eq. (E.23). Define $\xi_t = \operatorname{sgn}(y_{t+1} - a - by_t)\, \mathbf{1}(a + by_t > 0)$. Then Eq. (E.23) can be rewritten as

$$\sum_t \mathbb{E}(\xi_t | y_t) + \sum_t (\xi_t - \mathbb{E}(\xi_t | y_t)) \, . \tag{E.25}$$

The second term, $\sum\limits_t (\xi_t - \mathbb{E}(\xi_t|y_t))$ is of order $O(\sqrt{T})$. This follows from the fact that

$$\mathbb{E} \sum_t (\xi_t - \mathbb{E}(\xi_t|y_t)) = 0$$

and

$$\mathbb{E}\left(\sum_t (\xi_t - \mathbb{E}(\xi_t|y_t))\right)^2 = \sum_t \mathbb{E}\left(\xi_t - \mathbb{E}(\xi_t|y_t)\right)^2 < const \cdot T,$$

as each $\xi_t \in \{-1, 0, 1\}$.

Define $\Delta\alpha = \hat{\alpha} - \alpha$, $\Delta\beta = \hat{\beta} - \beta$, $v_{t+1} = \max(u_{t+1}, -\alpha - \beta y_t)$. Then

$$\mathbb{E}(\xi_t|y_t) = \mathbb{E}\left(\mathrm{sgn}(v_{t+1} - \Delta\alpha - \Delta\beta y_t)\mathbf{1}(\Delta\alpha + \Delta\beta y_t + \alpha + \beta y_t > 0)|y_t\right)$$
$$= (1 - 2F_v(\Delta\alpha + \Delta\beta y_t))\mathbf{1}(\Delta\alpha + \Delta\beta y_t + \alpha + \beta y_t > 0).$$

We want to linearize $\mathbb{E}(\xi_t|y_t)$. When $\beta > 1$ or $\alpha + \mathbb{E}u_t > 0, \beta = 1$ we know from Theorem D.6, that $y_t \xrightarrow{a.s.} \infty$. Thus, for T' large enough, starting from $t > T'$ we get $\alpha + \beta y_t \gg 0$, so that the indicator does not bind.

• Case I: $\beta > 1$.

We want to linearize $\mathbb{E}(\xi_t|y_t)$ around $\Delta\alpha + \Delta\beta y_t = 0$. Because y_t goes to infinity, we cannot assume $\Delta\beta y_t \approx 0$ for all t. However, we can assume that $\Delta\beta y_t \approx 0$ for $t < T - \sqrt{T}$. At the end of the proof we will find solution to first order conditions, which indeed satisfies this assumption, and has $\Delta\beta y_T = O(1)$ (so that $\Delta\beta y_{T-\sqrt{T}} = o(1)$) and $\Delta\alpha = o(1)$.

Thus, for $t \in (T', T - \sqrt{T})$,

$$(1 - 2F_v(\Delta\alpha + \Delta\beta y_t))\mathbf{1}(\Delta\alpha + \Delta\beta y_t + \alpha + \beta y_t > 0) = (1 - 2F_v(\Delta\alpha + \Delta\beta y_t))$$
$$\approx -2f_u(0)(\Delta\alpha + \Delta\beta y_t),$$

so that

$$\sum_t \mathbb{E}(\xi_t|y_t) \approx -2Tf_u(0)\Delta\alpha - 2f_u(0)\Delta\beta \sum_t y_t + O(\sqrt{T}),$$

where we again used the fact that $\xi_t \in \{-1, 0, 1\}$ to bound terms with $t \notin (T', T - \sqrt{T})$.

Therefore, Eq. (E.23) can be rewritten as

$$-2Tf_u(0)\Delta\alpha - 2f_u(0)\Delta\beta \sum_{t=T'}^{T-\sqrt{T}} y_t + O(\sqrt{T}) = 0. \tag{E.26}$$

Let us now analyze the second first order condition, Eq. (E.24).

Suppose that $\beta > 1$. Then terms corresponding to $t \approx T$ dominate the summation, as they have the largest y_{t+1}'s, and the indicator is no longer binding. By Lemma E.2 we know that almost surely there exists T_0 such that $y_{t+1} > \beta' y_t$ for some $\beta' \in (1, \beta)$ and any $t \geq T_0$. Thus, $y_{T-k} \leq \beta'^{-k} y_T$. Simultaneously we assume $y_{T_0} > 1$ to avoid pathologies.

Choose an additional integer τ to be fixed later and split the sum in Eq. (E.24) into three:

$$\begin{aligned}\sum_{t=1}^{T_0} y_t \mathbf{1}(\hat\alpha + \hat\beta y_t > 0)\mathrm{sgn}\big(u_t - \Delta\alpha - \Delta\beta\big) + \sum_{t=T_0+1}^{T-\tau} y_t \mathrm{sgn}\big(u_t - \Delta\alpha - \Delta\beta y_t\big)\\ + \sum_{t=T-\tau+1}^{T} y_t \mathrm{sgn}\big(u_t - \Delta\alpha - \Delta\beta\big)\end{aligned} \tag{E.27}$$

The first sum is bounded by a (random) number M as $T \to \infty$. Now choose (random, independent from T) τ such that $(\beta')^\tau > \frac{2}{1-1/\beta'}$ and $(\beta')^\tau > 2M$. In this case whenever all the signs in the third sum are positive, (E.27) is positive, and whenever all the signs in the third sum are negative, (E.27) is negative. Indeed, the first sum is bounded by M, and the very last term with $t = T$ is at least twice larger due to our choice of τ and the bound $y_{T_0} > 1$ that we started from. Since y_t grows faster than geometric series with denominator β' for $t > T_0$, the second sum can be bounded from above by $y_{T-\tau}$ multiplied by geometric series with denominator $1/\beta'$. Hence, by our choice of τ and inequality $y_T > (\beta')^\tau y_{T-\tau}$, the last term with $t = T$ is at least twice as large as the sum.

The conclusion is that the value of $\Delta\beta$ lies between the maximum value of b which makes all $\operatorname{sgn}\big(u_t - \Delta\alpha - by_t\big)$ positive for $t = T - \tau + 1, \dots, T$, and the minimum value of b which makes all $\operatorname{sgn}\big(u_t - \Delta\alpha - by_t\big)$ negative for $t = T - \tau + 1, \dots, T$. Thus, looking at the points, where $\operatorname{sgn}\big(u_t - \Delta\alpha - by_t\big)$ changes, we obtain

$$\min_{t=T-\tau-1,\dots,T} \frac{u_t - \Delta\alpha}{y_t} \le \Delta\beta \le \max_{t=T-\tau-1,\dots,T} \frac{u_t - \Delta\alpha}{y_t}. \tag{E.28}$$

Expressing $\Delta\beta$ via $\Delta\alpha$ using (E.26) and plugging into (E.28), we get

(E.29)

$$\frac{1}{T}\left(\sum_{t=1}^{T-\sqrt{T}} y_t\right) \min_{t=T-\tau-1,\dots,T} \frac{u_t - \Delta\alpha}{y_t} \le O\left(T^{-\frac{1}{2}}\right) - \Delta\alpha \le \frac{1}{T}\left(\sum_{t=1}^{T-\sqrt{T}} y_t\right) \max_{t=T-\tau-1,\dots,T} \frac{u_t - \Delta\alpha}{y_t}.$$

The last inequality implies that $\Delta\alpha \to 0$ almost surely as $T \to \infty$, since $\frac{\sum_{t=T'}^{T-\sqrt{T}} y_t}{y_{t'}}$ stays bounded as $T \to \infty$ for all $t' = T - \tau + 1, \dots, T$ (the numerator is at most $y_{T-\sqrt{T}}$ multiplied by a geometric series with denominator $1/\beta'$ and the denominator is at least $y_{T-\sqrt{T}}\beta'^{\sqrt{T}-\tau}$, so that their ratio is bounded by $1/(1-\beta'^{-1})$). In fact, we see that the speed of decay is $\frac{1}{T}$. Since $\Delta\alpha \to 0$, it stays bounded, and therefore, (E.28) implies that $\Delta\beta \to 0$ as fast as $1/y_T$, i.e. exponentially fast.

Finally, the solution to the first order conditions is a global minimum, not a local one, because y_t grows exponentially fast in t ($y_{t+1} > \beta' y_t$). Thus, for all large t, $\max(a + by_t, 0) = a + by_t$. This implies that most of the terms in (E.22) are convex functions of a and b (if the parameters are taken in a compact set), and therefore, the solution to the minimization problem can be found (up to a small error) as a point satisfying the first order conditions.

- Case II: $\beta = 1$, $\alpha + \mathbb{E}u_t > 0$.

In this case we do not need to consider separately observations with $t > T - \sqrt{T}$, and we will find solution to first order conditions, which has $\Delta\beta y_T = o(1)$ and $\Delta\alpha = o(1)$. Thus, for $t > T'$,

$$\begin{aligned}(1 - 2F_v(\Delta\alpha + \Delta\beta y_t))\mathbf{1}(\Delta\alpha + \Delta\beta y_t + \alpha + \beta y_t > 0) &= (1 - 2F_v(\Delta\alpha + \Delta\beta y_t)) \\ &\approx -2f_u(0)(\Delta\alpha + \Delta\beta y_t),\end{aligned}$$

and Eq. (E.23) can be rewritten as

$$-2Tf_u(0)\Delta\alpha - 2f_u(0)\Delta\beta\sum_{t=T'}^{T} y_t + O(\sqrt{T}) = 0. \tag{E.30}$$

Let us now analyze the second first order condition, Eq. (E.24). Define

$$\eta_t = y_t \mathrm{sgn}(y_{t+1} - (\hat\alpha + \hat\beta y_t)).$$

Then for $t > T'$ the sum in Eq. (E.24) can be rewritten as

$$\sum_t \mathbb{E}(\eta_t|y_t) + \sum_t(\eta_t - \mathbb{E}(\eta_t|y_t)),$$

where $\mathbb{E}\sum_t(\eta_t - \mathbb{E}(\eta_t|y_t)) = 0$ and

$$\begin{aligned}\mathbb{E}\left(\sum_t(\eta_t - \mathbb{E}(\eta_t|y_t))\right)^2 &= \mathbb{E}\sum_t\left(\eta_t - \mathbb{E}(\eta_t|y_t)\right)^2 = \mathbb{E}\sum_t\left(\eta_t^2 - (\mathbb{E}(\eta_t|y_t))^2\right)\\ &= \mathbb{E}\sum_t y_t^2\left(1 - \left(\mathbb{E}(\mathrm{sgn}(y_{t+1} - (\hat\alpha + \hat\beta y_t))|y_t)\right)^2\right) \le \mathbb{E}\sum_t y_t^2,\end{aligned}$$

because sgn $\in [-1, 1]$.

Random variable y_t grows linearly in t (i.e. $y_T = O(T)$). To see this, note that $y_t \ge \alpha + y_{t-1} + u_t \ge \alpha t + y_0 + \sum_{s=1}^{t} u_s$ and $y_t \le \alpha + y_{t-1} + |u_t| \le \alpha t + y_0 + \sum_{s=1}^{t}|u_s|$, so that

$$\alpha \xleftarrow[t\to\infty]{} \alpha + y_0/t + \frac{1}{t}\sum_{s=1}^{t} u_s \le y_t/t \le \alpha + y_0/t + \frac{1}{t}\sum_{s=1}^{t}|u_s| \xrightarrow[t\to\infty]{} \alpha + \mathbb{E}|u_t|.$$

Thus, $\mathbb{E}\sum_t y_t^2$ is of order $\sum_t t^2$, so that $\mathbb{E}\sum_t y_t^2 = O(T^3)$ and $\sum_t(\eta_t - \mathbb{E}(\eta_t|y_t)) = O(T^{3/2})$.

Define $v_{t+1} = \max(u_{t+1}, -\alpha - \beta y_t)$. We are left with analyzing

$$\begin{aligned}\sum_t \mathbb{E}(\eta_t|y_t) &= \sum_t y_t\mathbb{E}(\mathrm{sgn}(y_{t+1} - (\hat\alpha + \hat\beta y_t))|y_t) = \sum_t y_t\mathbb{E}(\mathrm{sgn}(v_{t+1} - \Delta\alpha - \Delta\beta y_t)|y_t)\\ &= \sum_t y_t(1 - 2F_{v_{t+1}|y_t}(-\Delta\alpha - \Delta\beta y_t)).\end{aligned}$$

Taylor expanding $1 - 2F_{v_{t+1}|y_t}(-\Delta\alpha - \Delta\beta y_t)$ around $\Delta\alpha + \Delta\beta y_t = 0$ and using the fact that $F_{v|y}(0) = F_u(0) = 0.5$, $f_{v|y}(0) = f_u(u)$ as for $t > T'$ $\alpha + \beta y_t > 0$, we get

$$\sum_t \mathbb{E}(\eta_t|y_t) = -2f_u(0)\sum_t y_t(\Delta\alpha + \Delta\beta y_t + o(\Delta\alpha + \Delta\beta y_t)),$$

so that Eq. (E.24) can be rewritten as

$$\Delta\alpha\sum_{t=T'}^{T} y_t + \Delta\beta\sum_{t=T'}^{T} y_t^2 = O(T^{3/2}). \tag{E.31}$$

Solving Eq. (E.30) and (E.31), and using the fact that y_t grows linearly in T so that $\sum_t y_t = O(\sum_t t) = O(T^2)$, we get $\Delta\alpha = O(T^{-0.5})$, $\Delta\beta = O(T^{-1.5})$. Thus, $\Delta\alpha \xrightarrow[T\to\infty]{} 0$, $\Delta\beta \xrightarrow[T\to\infty]{} 0$, and the LAD estimator is consistent. Similarly to the Case I, the solution to the first order conditions is a global minimum because y_t grows linearly in t. Thus, for all large t, $\max(a+by_t, 0) = a+by_t$. This implies that most of the terms in (E.22) are convex functions of a and b (if the parameters are taken in a compact set), and therefore, the solution to the minimization problem can be found (up to a small error) as a point satisfying the first order conditions.

- Case III: $\beta = 1$, $\alpha + \mathbb{E}u_t = 0$.

Let us calculate the order of terms, where indicator binds. By Theorem D.9, $\frac{1}{\sqrt{T}} y_{\lfloor Ts \rfloor} \to \sigma|W(s)|$ for any $s \in (0,1]$, where $\sigma = \mathbb{E}u_t^2$ and W is a standard Brownian motion. Thus, the indicator can be rewritten as $\mathbf{1}\left(b\frac{y_{\lfloor Ts \rfloor}}{\sqrt{T}} > -\frac{a}{\sqrt{T}}\right)$.

As T goes to infinity, $\frac{a}{\sqrt{T}}$ goes to zero and $b\frac{y_{\lfloor Ts \rfloor}}{\sqrt{T}}$ goes to $b\sigma|W(s)|$. The (random) time Brownian motion spends inside interval $[-\varepsilon, \varepsilon]$ goes to zero almost surely as $\varepsilon \to 0$ (see Section 3.6 in Karatzas and Shreve (2012)). Thus, for $b > 0$ the time $b\sigma|W(s)|$ is smaller than $-\frac{a}{\sqrt{T}}$ goes to zero as $T \to \infty$, i.e. it is $o(1)$. Therefore, $\sharp\{t : b\frac{y_{\lfloor Ts \rfloor}}{\sqrt{T}} \le -\frac{a}{\sqrt{T}}\} = o(T)$.

Note also that $b \le 0$ cannot solve the minimization problem (E.22). When $b < 0$, $b\sigma|W(s)| \le 0$ and the indicator starts to bind all the time, so that we get $\sum_t |y_{t+1}| = O(T^{3/2})$, as $\frac{1}{T\sqrt{T}} \sum_t |y_{t+1}| \to \sigma \int_0^1 |W(s)|ds$. Yet, $\sum_t |y_{t+1} - y_t| = \sum_t |u_t| = O(T)$. When $b = 0$, we again get $\sum_t |y_{t+1} - [a]_+| = O(T^{3/2})$ as $\frac{1}{T\sqrt{T}} \sum_t |y_{t+1} - [a]_+| = \frac{1}{T} \sum_t |y_{t+1}/\sqrt{T} - [a]_+/\sqrt{T}| \to \sigma \int_0^1 |W(s)|ds$.

Thus, we can linearize Eq. (E.23) to get

$$-2Tf_u(0)\Delta\alpha - 2f_u(0)\Delta\beta \sum_t y_t + o(T) + O(\sqrt{T}) = 0. \tag{E.32}$$

To analyze the second first order condition, Eq. (E.24), we proceed as in Case II. The only difference is that now we use Theorem D.9 to calculate the order of different sums involving y_t. That is, $\frac{1}{T^2}\mathbb{E}\sum_t y_t^2 \xrightarrow[T\to\infty]{} \mathbb{E}\sigma^2 \int_0^1 W^2(s)ds$ implying $\mathbb{E}\sum_t y_t^2 = O(T^2)$ and $\frac{1}{T^{3/2}}\mathbb{E}\sum_t y_t \xrightarrow[T\to\infty]{} \mathbb{E}\sigma \int_0^1 |W(s)|ds$ implying $\mathbb{E}\sum_t y_t = O(T^{3/2})$. Thus, the second first order condition, Eq. (E.24) can be rewritten as

$$\Delta\alpha O(T^{3/2}) + \Delta\beta O(T^2) + o(T^{3/2}) = O(T). \tag{E.33}$$

Solving Eq. (E.32) and (E.33), we get $\Delta\alpha = o(1)$, $\Delta\beta = o(T^{-0.5})$. Thus, $\Delta\alpha \xrightarrow[T\to\infty]{} 0$, $\Delta\beta \xrightarrow[T\to\infty]{} 0$, and the LAD estimator is consistent. The solution to the first order conditions is a global minimum because the indicator in the maximization problem binds in $o(T)$ terms.

This implies that the dominant terms in (E.22) are convex functions of a and b (if the parameters are taken in a compact set), and therefore, the solution to the minimization problem can be found (up to a small error) as a point satisfying the first order conditions. □

Appendix F. Tables

Table F.1 shows relative performance of LAD with peer effects (at $t-1$) versus LAD without peer effects. The main text contains analogues table (Table 3) for LAD with peer effects at $t-4$.

$\frac{R_{w.p.e.}-R_{w/o\ p.e.}}{R_{w/o\ p.e.}}$

From \ To	FR	NL	DE	IT	UK	IE	DK	EL	PT	ES	BE	LU
FR	·	0.07	-0.00	0.02	0.04	-0.02	0.01	-0.12	0.05	0.00	-0.01	0.03
NL	-0.13	·	-0.32	-0.02	-0.10	0.00	-0.03	-0.04	-0.10	-0.06	-0.02	0.03
DE	0.02	-0.04	·	-0.09	-0.02	-0.03	-0.09	-0.05	-0.05	0.01	0.02	-0.04
IT	-0.03	-0.06	-0.03	·	-0.03	-0.03	-0.18	0.00	0.02	-0.02	-0.04	-0.17
UK	-0.09	-0.02	-0.08	-0.01	·	0.08	0.00	-0.10	-0.02	-0.01	-0.03	0.22
IE	0.02	-0.09	0.02	-0.18	-0.14	·	-0.03	-0.03	-0.09	0.04	-0.00	0.10
DK	-0.12	-0.09	-0.01	-0.14	-0.01	0.00	·	-0.03	-0.08	-0.13	0.06	0.14
EL	-0.09	0.07	-0.07	-0.00	-0.03	0.04	-0.09	·	0.00	0.01	-0.01	-0.13
PT	-0.06	-0.20	0.02	-0.04	-0.02	-0.03	0.04	0.11	·	0.04	-0.05	-0.17
ES	-0.09	-0.07	0.03	-0.14	-0.02	-0.01	-0.03	-0.07	-0.00	·	-0.03	0.08
BE	0.01	0.02	0.04	-0.01	-0.02	0.08	0.00	0.03	-0.08	0.01	·	-0.03
LU	0.01	0.19	0.02	0.01	-0.17	0.05	-0.07	0.07	-0.06	0.16	0.03	·

TABLE F.1. Performance of LAD with peer effects (at $t-1$) relative to LAD without peer effects for all country pairs. Negative numbers mean LAD with peer effects is better, positive numbers mean LAD without peer effects is better. For each pair of countries (i, j) we calculate $R_{w.(w/o)p.e.} = \frac{1}{T-T'} \sum_{t=1}^{T-T'} |y_{ij,t+T'} - \hat{y}_{ij,t+T'}|$, where $\hat{y}_{ij,t+T'}$ is the prediction based on LAD with (without) peer effects and window size is $T' = 115$.

References

Feller, W., *An introduction to probability theory and its applications*, Vol. 2, John Wiley & Sons, 2008.

Hall, P. and C.C. Heyde, *Martingale limit theory and its application*, Academic press, 1980.

Hayashi, F., *Econometrics*, Princeton University Press, 2000.

Kallenberg, O., *Foundations of modern probability*, 2 ed., Springer Science & Business Media, 2002.

Karatzas, I. and S. Shreve, *Brownian motion and stochastic calculus*, Springer Science & Business Media, 2012.

Lévy, P., *Processus Stochastiques et Mouvement Brownien*, Gauthier-Villars,Paris, 1948.

Michel, J. and R.M. de Jong, "Mixing properties of the dynamic Tobit model with mixing errors," *Economics Letters*, 2018, *162*, 112–115.

Powell, J.L., "Least absolute deviations estimation for the censored regression model," *Journal of Econometrics*, 1984, *25* (3), 303–325.

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