Online Supplement

"Econometric Analysis of Production Networks with Dominant Units"

by

M. Hashem Pesaran and Cynthia Fan Yang

S1 Extremum estimator for large N and T panels

In this appendix, we extend the extremum estimator to large N and T panels where the errors are allowed to be serially correlated and heteroskedastic. Consider the network $\mathbf{W}_t = (w_{ij,t})$, for $i, j = 1, 2, ..., N$, where $w_{ij,t} \geq 0$, and $\mathbf{W}_t \boldsymbol{\tau}_N = 1$, and denote the outdegrees by d_{it}, and note that $\mathbf{d}_t = (d_{1t}, d_{2t}, .., d_{Nt})' = \mathbf{W}_t' \boldsymbol{\tau}_N$. Suppose that d_{it}, for $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$, are generated according to the following exponent specification

$$d_{it} = \kappa N^{\delta_i} \exp(\upsilon_{it}),\ i = 1, 2, ..., N;\ t = 1, 2, ..., T, \tag{S.1}$$

and assume that υ_{it} follows a covariance stationary process with absolutely summable autocovariances, $\upsilon_{it} = \sum_{l=0}^{\infty} a_{il} \varsigma_{i,t-l}$, where $\varsigma_{it} \sim IID(0, 1)$ with finite fourth-order moments. Let $\gamma_i(h)$ denote the h-order autocovariance of υ_{it}, $\gamma_i(h) = E(\upsilon_{it}\upsilon_{i,t+h}) = \sum_{l=0}^{\infty} a_{il} a_{i,l+|h|}$, for $h = 0, 1, 2, ...$ Also note that δ_i are fixed constants in the range $0 \leq \delta_i \leq 1$ that satisfy the summability condition $\sum_{i=1}^{N} \delta_i < K < \infty$, and the following constraint for each time period

$$\sum_{i=1}^{N} d_{it} = \boldsymbol{\tau}_N' \mathbf{d}_t = N = \kappa \sum_{i=1}^{N} N^{\delta_i} \exp(\upsilon_{it}). \tag{S.2}$$

As with the case of short T panels considered in Section 7.2 of the paper, consider the following estimator of δ_i

$$\hat{\delta}_i = \frac{T^{-1} \sum_{t=1}^{T} \ln d_{it} - (TN)^{-1} \sum_{t=1}^{T} \sum_{j=1}^{N} \ln d_{jt}}{\ln N}, \tag{S.3}$$

and note that

$$\hat{\delta}_i - \delta_i = \bar{\delta} + \frac{\bar{\upsilon}_i - \bar{\upsilon}}{\ln N}, \tag{S.4}$$

where $\bar{\delta} = N^{-1} \sum_{i=1}^{N} \delta_i$, $\bar{\upsilon}_i = T^{-1} \sum_{t=1}^{T} \upsilon_{it}$, and $\bar{\upsilon} = N^{-1} \sum_{i=1}^{N} \bar{\upsilon}_i$. Also

$$Var(\bar{\upsilon}_i) = \frac{1}{T} \left[\gamma_i(0) + 2 \sum_{h=1}^{T-1} \left(1 - \frac{h}{T} \right) \gamma_i(h) \right],\ \text{for } i = 1, 2, ..., N, \tag{S.5}$$

and since $Cov(\bar{v}_i, \bar{v}_j) = 0$ for $i \neq j$ we have

$$Var(\bar{v}) = \frac{1}{NT}\left[\bar{\gamma}(0) + 2\sum_{h=1}^{T-1}\left(1 - \frac{h}{T}\right)\bar{\gamma}(h)\right], \tag{S.6}$$

where $\bar{\gamma}(h) = N^{-1}\sum_{k=1}^{N}\gamma_k(h)$, $h = 0, 1, 2, ...$ Using (S.5) and (S.6) in (S.4) yields, for all $i \neq j$,

$$\begin{aligned}
Cov(\hat{\delta}_i, \hat{\delta}_j) &= \frac{1}{(\ln N)^2} Cov(\bar{v}_i - \bar{v}, \bar{v}_j - \bar{v}) \\
&= \frac{1}{(\ln N)^2}\left[Cov(\bar{v}_i, \bar{v}_j) + Var(\bar{v}) - Cov(\bar{v}_i, \bar{v}) - Cov(\bar{v}_j, \bar{v})\right] \\
&= \frac{1}{(\ln N)^2}\left[Var(\bar{v}) - \frac{1}{N}Var(\bar{v}_i) - \frac{1}{N}Var(\bar{v}_j)\right] \\
&= \frac{1}{(\ln N)^2 NT}\left\{\bar{\gamma}(0) - \gamma_i(0) - \gamma_j(0) + 2\sum_{h=1}^{T-1}\left(1 - \frac{h}{T}\right)\left[\bar{\gamma}(h) - \gamma_i(h) - \gamma_j(h)\right]\right\},
\end{aligned}$$

and

$$\begin{aligned}
Var\left(\hat{\delta}_i\right) &= \frac{1}{(\ln N)^2} Var(\bar{v}_i - \bar{v}) \\
&= \frac{1}{(\ln N)^2}\left[Var(\bar{v}_i) + Var(\bar{v}) - 2Cov(\bar{v}_i, \bar{v})\right] \\
&= \frac{1}{(\ln N)^2}\left[\left(1 - \frac{2}{N}\right)Var(\bar{v}_i) + Var(\bar{v})\right] \\
&= \frac{1}{(\ln N)^2 T}\left(1 - \frac{2}{N}\right)\left[\gamma_i(0) + 2\sum_{h=1}^{T-1}\left(1 - \frac{h}{T}\right)\gamma_i(h)\right] \\
&\quad + \frac{1}{(\ln N)^2 NT}\left[\bar{\gamma}(0) + 2\sum_{h=1}^{T-1}\left(1 - \frac{h}{T}\right)\bar{\gamma}(h)\right].
\end{aligned} \tag{S.7}$$

It follows that $Var\left(\hat{\delta}_i\right)$ can be estimated by

$$\begin{aligned}
\widehat{Var}\left(\hat{\delta}_i\right) &= \frac{1}{(\ln N)^2 T}\left(1 - \frac{2}{N}\right)\left[\hat{\gamma}_i(0) + 2\sum_{h=1}^{L}\left(1 - \frac{h}{L}\right)\hat{\gamma}_i(h)\right] \\
&\quad + \frac{1}{(\ln N)^2 NT}\left[\hat{\bar{\gamma}}(0) + 2\sum_{h=1}^{L}\left(1 - \frac{h}{L}\right)\hat{\bar{\gamma}}(h)\right],
\end{aligned} \tag{S.8}$$

where

$$\begin{aligned}
\hat{\gamma}_i(h) &= \frac{\sum_{t=h+1}^{T}\hat{v}_{it}\hat{v}_{i,t-h}}{T}, \\
\hat{\bar{\gamma}}(h) &= \frac{\sum_{k=1}^{N}\hat{\gamma}_k(h)}{N},
\end{aligned}$$

and

$$\hat{\upsilon}_{it} = \ln d_{it} - \widehat{\ln \kappa} - \hat{\delta}_i \ln N, \tag{S.9}$$

as $N \to \infty$, and $L/T \to 0$, as $T \to \infty$. The value of L is often set to $T^{1/3}$, which ensures that $L/T \to 0$, as $T \to \infty$.

The extremum estimator of $\delta_{\max} = \max(\delta_1, \delta_2, ..., \delta_N)$ is given by

$$\hat{\delta}_{\max} = \frac{T^{-1}\sum_{t=1}^{T} \ln d_{\max,t} - (TN)^{-1}\sum_{t=1}^{T}\sum_{j=1}^{N} \ln d_{jt}}{\ln N}, \tag{S.10}$$

where $d_{\max,t}$ is the largest value of d_{it} for period t. The asymptotic normality of $\hat{\delta}_{\max}$ can be established by applying standard central limit theorems for stationary processes to $\bar{\upsilon}_i - \bar{\upsilon}$, which leads to

$$\frac{\left(\hat{\delta}_{\max} - \delta_{\max} - \bar{\delta}\right)}{\left[Var\left(\hat{\delta}_{\max}\right)\right]^{1/2}} \to_d N(0,1), \text{ as } N, T \to \infty \text{ jointly.}$$

To eliminate the nuisance parameter $\bar{\delta}$, the condition

$$\bar{\delta}(\ln N)\sqrt{T} = \left(\sum_{i=1}^{N} \delta_i\right)\frac{(\ln N)\sqrt{T}}{N} \to 0, \tag{S.11}$$

has to hold as N and $T \to \infty$, and given the summability condition it is sufficient that the following condition on the relative expansion rates of N and T is satisfied

$$\frac{(\ln N)\sqrt{T}}{N} \to 0, \tag{S.12}$$

as $N, T \to \infty$, jointly, which implies that when T takes moderate to large values, N needs to be sufficiently large relative to T. It is clear that N and T can rise at the same rate. But by setting $T = \ominus\left(N^{\phi}\right)$, it also follows that condition (S.12) can be rewritten as

$$\frac{(\ln N)\sqrt{T}}{N} = \exp\left[\ln(\ln N) + \left(\frac{\phi}{2} - 1\right)\ln N\right] \to 0,$$

as $N \to \infty$, which holds if and only if

$$\ln(\ln N) + \left(\frac{\phi}{2} - 1\right)\ln N < 0.$$

as $N \to \infty$. Therefore, (S.12) will be satisfied so long as $\phi < 2$, which allows T to rise faster than N.

Hence, the statistic for testing $\delta_{\max} = \delta^0_{\max}$, where $\delta^0_{\max} > 1/2$, is given by

$$\mathfrak{D}_{\max} = \frac{\left(\hat{\delta}_{\max} - \delta^0_{\max}\right)}{\left[\widehat{Var}\left(\hat{\delta}_{\max}\right)\right]^{1/2}}, \tag{S.13}$$

and $\mathfrak{D}_{\max} \to_d N(0,1)$, if $(\ln N) N^{-1}\sqrt{T} \to 0$, as $N, T \to \infty$ jointly, and $L/T \to 0$, as $T \to \infty$.

S2 Monte Carlo supplement

S2.1 Experiments with exponentially decaying $\delta_{(i)}$

The observations on the outdegrees, d_{it}, are generated by the following exponent specification

$$\ln d_{it} = \ln \kappa + \delta_i \ln N + \upsilon_{it},\ \ i = 1, 2, ..., N;\ \ t = 1, 2, ..., T, \tag{S.14}$$

where $\upsilon_{it} \sim IIDN(0, 1)$, and

$$\kappa = \frac{exp\left(-\frac{1}{2}\right)}{N^{-1}\sum_{i=1}^{N} N^{\delta_i}} > 0, \tag{S.15}$$

such that d_{it} sum up to N across i for each t. We consider balanced panels and allow all units to be weakly dominant. To ensure that $\bar{\delta} = N^{-1}\sum_{i=1}^{N}\delta_i \to 0$ at a sufficiently fast rate, we assume that individual $\delta_{(i)}$ decays exponentially, where the degree of dominance of unit i is denoted by δ_i, and the associated ordered values are denoted by $\delta_{(i)}$, namely, $\delta_{\max} = \delta_{(1)} \geq \delta_{(2)} \geq ... \geq \delta_{(N)}$.[S1] In particular, we consider $\delta_{(i)} = 0.9^i$, for $i = 1, 2, ..., N$, and combinations of $N = 100, 300, 500, 1,000$, and $450,000$, and $T = 1, 2, 6, 10$, and 20. The number of replications is set to $2,000$. We report the top four largest estimates of δ, denoted by $\hat{\delta}_{\max} = \hat{\delta}_{(1)} \geq \hat{\delta}_{(2)} \geq \hat{\delta}_{(3)} \geq \hat{\delta}_{(4)}$, which are computed by (S.3). When $T > 1$, the variance of $\hat{\delta}_{(i)}$ is estimated by

$$\widehat{Var}\left(\hat{\delta}_i\right) = \frac{\hat{\sigma}_{\upsilon}^2}{(\ln N)^2 T}\left(1 - \frac{1}{N}\right), \tag{S.16}$$

where

$$\hat{\sigma}_{\upsilon}^2 = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{\upsilon}_{it}^2}{N(T-1)}, \tag{S.17}$$

and $\hat{\upsilon}_{it}$ is given by (S.9).

We also carry out misspecification experiments by generating the outdegrees by (S.14) with $\delta_{(i)} = 0.75^i$ and 0.9^i, for $i = 1, 2, ..., N$, and comparing the performance of the extremum estimator with that of the three power law estimators, namely, the Gabaix-Ibragimov estimator ($\hat{\beta}_{GI}$), the maximum likelihood estimator ($\hat{\beta}_{MLE}$), and the Clauset, Shalzi and Newman (2009, CSN) estimator ($\hat{\beta}_{CSN}$).[S2] The sample sizes under consideration are combinations of $N = 100, 300, 500, 1,000$, and $450,000$, and $T = 1$ and 2.

Table S.1 summarizes the estimation results for the four largest values of δ, namely $0.9, 0.9^2, 0.9^3$, and 0.9^4. For other values of $\delta_{(i)}$, for $i = 5, 6, ..., N$, the estimates fall below $1/2$ and have no consequence for the shock diffusion within the network. These results confirm the validity of our theoretical derivations for the case where the degrees of dominance of units in a network decay exponentially.

Table S.2 reports the frequencies with which each of the top four dominant units are selected across $2,000$ Monte Carlo replications. The probability of correct identification

[S1] Note that the denominator of (S.15), $N^{-1}\sum_{i=1}^{N} N^{\delta_i}$, converges to a finite positive constant.

[S2] See Section 7.1 of the paper for details.

is lower, compared with the results for the experiments with a finite number of dominant units (Table 2 in the paper). As expected, the more clustered are the degrees of pervasiveness across units, the more difficult it is to differentiate one unit from another.

Finally, turning to Tables S.3 and S.4, we observe that the three power law estimators all suffer from severe biases when the DGP follows the exponent specification, especially when N is large.

S2.2 Experiments with unbalanced panels

The data generating process (DGP) is given by the exponent specification, (S.14), for $i = 1, ..., N$, and $t = T_i^0, T_i^0 + 1, ..., T_i^1$, $(T_i^1 \geq T_i^0)$. We generate an unbalanced panel where the number of time series observations for unit i, namely $T_i = T_i^1 - T_i^0 + 1$, lies between 2 and 4. To ensure that the most important dominant units are present across the years, only units in the bottom 95^{th} percentile of the distribution of δ were subject to missing observations. In the case of these units, we dropped the first and the last observations with a 50% probability. This randomization process is repeated for all the $2,000$ replications.

We consider networks with a finite number of dominant units, and a large number of non-dominant units. Specifically,

- A.1. One strongly dominant unit: $\delta_{\max} = \delta_{(1)} = 1$, with $\delta_{(i)} = 0$ for $i = 2, 3, ..., N$.
- A.2. Two strongly dominant units: $\delta_{\max} = \delta_{(1)} = \delta_{(2)} = 1$, with $\delta_{(i)} = 0$ for $i = 3, 4, ..., N$.
- A.3. One strongly dominant unit and one weakly dominant unit: $\delta_{\max} = \delta_{(1)} = 1$ and $\delta_{(2)} = 0.75$, with $\delta_{(i)} = 0$ for $i = 3, 4, ..., N$.

The estimates of δ_i are computed using

$$\hat{\delta}_i = \frac{T_i^{-1} \sum_{t=T_i^0}^{T_i^1} \ln d_{it} - N^{-1} \sum_{i=1}^{N} \left(T_i^{-1} \sum_{t=T_i^0}^{T_i^1} \ln d_{it} \right)}{\ln N}, \tag{S.18}$$

and their variances (when $T > 1$) are estimated by

$$\widehat{Var}\left(\hat{\delta}_i\right) = \frac{\hat{\sigma}_v^2}{(\ln N)^2} \left(\frac{1}{T_i} - \frac{1}{NT_i} \right), \tag{S.19}$$

where

$$\hat{\sigma}_v^2 = \frac{\sum_{i=1}^{N} (T_i - 1)^{-1} \sum_{t=T_i^0}^{T_i^1} \hat{v}_{it}^2}{N},$$

and $\hat{v}_{it}$ is given by (S.9).

The results are presented in Table S.5. It can be seen that the extremum estimator continues to perform well. Note, however, that in the case of unbalanced panels, we need to assume that the outdegrees of the units with the highest degrees of dominance are observed for at least two time periods.

S2.3 Experiments with heteroskedastic and serially correlated errors

The DGP is given by

$$\ln d_{it} = \ln \kappa + \delta_i \ln N + \upsilon_{it},\ \ i = 1, 2, ..., N;\ \ t = 1, 2, ..., T, \tag{S.20}$$

where the idiosyncratic errors, υ_{it}, are generated as

$$\begin{aligned} \upsilon_{it} &= \sigma_i e_{it}, \quad \text{for } i = 1, 2, ..., N;\ t = 1, 2, ..., T, \\ \sigma_i^2 &= \frac{1}{2} + \frac{3}{4} z_i, \text{ where } z_i \sim IID\left[\chi^2(2)\right], \\ e_{it} &= \rho_{e,i} e_{i,t-1} + \sqrt{1-\rho_{e,i}^2}\,\eta_{it}, \text{ for } t = -49, ..., 0, 1, ..., T, \\ e_{i,-50} &= 0, \quad \eta_{it} \sim IIDN(0,1), \\ \rho_{e,i} &\sim IIDU(0.05, 0.95), \end{aligned} \tag{S.21}$$

where σ_i^2 are generated following Bailey et al. (2016) such that all σ_i^2 are bounded away from zero and $N^{-1}\sum_{i=1}^{N}\sigma_i^2 \to 2$, as $N \to \infty$.

To ensure that

$$\kappa \sum_{i=1}^{N} N^{\delta_i} \exp(\upsilon_{it}) = N,$$

for N sufficiently large, κ is set to

$$\kappa = \frac{1}{N^{-1}\sum_{i=1}^{N} N^{\delta_i} \exp\left[\frac{1}{2}Var(\upsilon_{it})\right]}, \tag{S.22}$$

where $Var(\upsilon_{it}) = \sigma_i^2$.

Under (S.20) we consider Experiments A.1-A.3, which are described in Section S2.2, and all experiments are replicated $2,000$ times for combinations of $N = 500, 1,000, 2,000, 5,000$, and $10,000$, and $T = 50$, 100, 200, and 500. The values of δ_i, $\rho_{e,i}$ and σ_i^2 are fixed across replications. The test statistic is computed by (S.13), where $L = T^{1/3}$.

The results are summarized in Table S.6. As can be seen the bias of the extremum estimator is very small and its RMSE declines with N and/or T as predicted by the theory. The empirical sizes are close to the 5% nominal size if both N and T are large, and if N is large enough relative to T, which is in line with condition (S.12). The test based on (S.13) has a high power, which improves with both N and T.

Table S.1: Bias, RMSE, size and power of the extremum estimator for the top four dominant units under Exponent DGP with exponentially decaying $\delta_{(i)}$

	Bias(×100)					RMSE(×100)					Size(×100)					Power(×100)				
$T \backslash N$	100	300	500	1,000	450,000	100	300	500	1,000	450,000	100	300	500	1,000	450,000	100	300	500	1,000	450,000
										$\delta_{(1)} = 0.9$										
1	4.04	5.60	5.44	4.93	1.28	15.42	13.87	13.08	12.00	6.70	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2	-2.43	1.17	1.63	1.76	0.37	11.82	9.84	9.24	8.54	5.01	1.15	1.95	2.30	2.45	3.20	14.45	35.70	44.85	55.50	97.85
6	-7.24	-2.10	-1.14	-0.45	-0.03	10.34	6.60	5.95	5.36	3.05	7.80	2.75	2.20	2.50	4.15	25.25	71.90	84.50	94.25	100.00
10	-8.19	-2.67	-1.60	-0.81	-0.09	10.21	5.79	5.03	4.44	2.45	19.15	5.70	4.30	4.00	5.40	37.30	89.80	97.00	99.60	100.00
20	-8.80	-2.96	-1.80	-0.93	-0.05	9.88	4.77	3.91	3.29	1.72	44.65	10.80	6.75	4.80	4.90	64.15	99.60	100.00	100.00	100.00
										$\delta_{(2)} = 0.9^2 = 0.81$										
1	-1.74	1.40	1.78	1.85	0.39	11.46	10.04	9.57	8.92	5.70	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2	-6.33	-1.54	-0.68	-0.10	0.08	10.89	7.90	7.35	6.80	4.36	0.25	0.30	0.30	0.25	1.75	4.85	22.40	32.70	48.10	98.50
6	-8.46	-2.79	-1.67	-0.84	-0.10	10.45	6.11	5.43	4.90	3.02	6.85	1.85	1.45	1.45	4.35	17.55	72.15	84.95	94.05	100.00
10	-8.74	-2.87	-1.70	-0.82	0.01	10.10	5.29	4.55	4.04	2.40	17.55	3.50	3.00	2.75	4.40	33.95	91.75	97.45	99.75	100.00
20	-8.87	-2.92	-1.73	-0.85	0.02	9.73	4.52	3.68	3.12	1.69	43.90	8.60	5.65	4.90	5.20	66.45	99.85	100.00	100.00	100.00
										$\delta_{(3)} = 0.9^3 = 0.729$										
1	-3.59	-0.05	0.52	0.82	0.35	10.32	8.70	8.32	7.84	5.24	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2	-6.86	-1.81	-0.86	-0.20	0.17	10.43	7.19	6.66	6.21	4.28	0.15	0.05	0.20	0.25	1.05	2.55	19.10	31.80	48.25	99.10
6	-8.45	-2.67	-1.52	-0.68	0.05	10.21	5.78	5.09	4.60	2.94	6.55	1.15	0.95	1.15	3.30	16.40	74.45	87.25	95.10	100.00
10	-8.61	-2.73	-1.57	-0.70	0.06	9.96	5.22	4.49	3.99	2.39	17.50	3.25	2.60	2.40	4.35	34.15	92.50	97.85	99.75	100.00
20	-8.75	-2.81	-1.63	-0.75	0.07	9.66	4.52	3.71	3.17	1.74	43.70	9.00	5.10	4.20	4.80	65.30	99.80	100.00	100.00	100.00
										$\delta_{(4)} = 0.9^4 = 0.6561$										
1	-4.04	-0.36	0.23	0.57	0.28	9.30	7.78	7.54	7.21	5.03	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2	-7.01	-1.83	-0.88	-0.21	0.00	10.19	6.90	6.42	6.01	4.18	0.05	0.10	0.05	0.05	0.90	1.45	18.30	30.45	47.05	99.05
6	-8.36	-2.56	-1.43	-0.60	0.11	9.98	5.52	4.86	4.38	2.82	4.55	1.00	0.80	0.85	3.15	15.20	74.85	89.65	96.85	100.00
10	-8.73	-2.84	-1.66	-0.79	0.00	9.85	4.92	4.16	3.67	2.29	15.60	2.55	1.70	1.30	3.80	32.60	94.35	98.75	99.85	100.00
20	-8.97	-3.02	-1.82	-0.94	-0.05	9.77	4.53	3.68	3.12	1.71	46.00	8.10	5.35	3.85	5.10	64.50	99.65	100.00	100.00	100.00

Notes: The DGP is given by (S.14), where the true values of δ are generated as $\delta_{(i)} = 0.9^i$, for $i = 1, 2, ..., N$. $\delta_{(i)}$ denotes the i^{th} largest δ, namely, $\delta_{\max} = \delta_{(1)} \geq \delta_{(2)} \geq \delta_{(3)} \geq \delta_{(4)} \geq ...$, which are estimated by (S.3). The standard errors of $\hat{\delta}_{(i)}$ cannot be computed when $T = 1$, otherwise they are computed by (S.16). The power is calculated at $\delta_{(i)} - 0.2$ when the true value is $\delta_{(i)}$, for $i = 1, 2, 3, 4$. The number of replications is $2,000$.

Table S.2: Frequencies with which each of the top four dominant units are selected under Exponent DGP with exponentially decaying $\delta_{(i)}$

	Empirical frequency (percent)				
$T \backslash N$	100	300	500	1,000	450,000
	$\delta_{(1)} = 0.9$				
1	40.40	48.40	51.85	56.15	78.05
2	54.00	60.65	64.00	68.00	87.65
6	72.20	80.20	83.30	86.50	97.85
10	79.60	85.85	88.25	90.85	99.80
20	89.95	94.85	95.85	97.25	99.95
	$\delta_{(2)} = 0.9^2 = 0.81$				
1	23.10	27.15	29.15	31.25	54.95
2	31.55	38.20	41.60	46.00	73.20
6	48.05	59.00	63.65	69.15	94.20
10	59.45	69.95	75.15	80.40	98.95
20	78.20	87.60	89.70	92.90	99.90
	$\delta_{(3)} = 0.9^3 = 0.729$				
1	14.80	18.85	21.80	25.25	48.30
2	21.90	28.55	31.75	35.35	65.45
6	40.40	50.80	55.15	61.00	90.90
10	54.95	65.40	70.95	76.50	97.70
20	73.20	83.50	86.75	90.75	99.95
	$\delta_{(4)} = 0.9^4 = 0.6561$				
1	10.90	14.50	17.05	18.80	41.80
2	17.75	23.85	25.90	28.80	59.80
6	37.30	47.50	51.55	56.55	87.80
10	50.80	60.65	65.55	71.30	95.75
20	68.00	78.00	82.30	87.15	99.85

Notes: This table complements Table S.1 and reports the frequencies with which each of the top four dominant units are selected across $2,000$ replications. The DGP is given by (S.14), where the true values of δ are generated as $\delta_{(i)} = 0.9^i$, for $i = 1, 2, ..., N$. See also the notes to Table S.1.

Table S.3: Estimates of the shape parameter, β, of the power law and inverse of the exponent, $\delta_{\max}$, under Exponent DGP with exponentially decaying $\delta_{(i)}$, where $\delta_{(i)} = 0.9^i$ ($\beta = 1/0.9 = 1.11$)

		$T=1$					$T=2$				
	N	100	300	500	1,000	450,000	100	300	500	1,000	450,000
Assumed cut-off value		**Log-log regression** $\left(\widehat{\beta}_{GI}\right)$									
10%		1.06	0.92	0.95	1.04	2.34	1.08	0.93	0.96	1.05	2.34
		(0.48)	(0.24)	(0.19)	(0.15)	(0.02)	(0.34)	(0.17)	(0.14)	(0.11)	(0.01)
20%		0.99	1.00	1.06	1.19	2.09	1.01	1.01	1.07	1.19	2.09
		(0.31)	(0.18)	(0.15)	(0.12)	(0.01)	(0.23)	(0.13)	(0.11)	(0.08)	(0.01)
30%		0.98	1.04	1.12	1.25	1.90	0.99	1.05	1.12	1.25	1.90
		(0.25)	(0.16)	(0.13)	(0.10)	(0.01)	(0.18)	(0.11)	(0.09)	(0.07)	(0.01)
Assumed cut-off value		**Maximum Likelihood Estimation** $\left(\widehat{\beta}_{MLE}\right)$									
10%		1.13	1.10	1.22	1.44	2.11	1.05	1.07	1.21	1.42	2.11
		(0.36)	(0.20)	(0.17)	(0.14)	(0.01)	(0.23)	(0.14)	(0.12)	(0.10)	(0.01)
20%		1.04	1.18	1.29	1.45	1.79	1.01	1.16	1.28	1.44	1.79
		(0.23)	(0.15)	(0.13)	(0.10)	(0.01)	(0.16)	(0.11)	(0.09)	(0.07)	(0.00)
30%		1.01	1.16	1.25	1.36	1.57	0.98	1.15	1.25	1.36	1.57
		(0.18)	(0.12)	(0.10)	(0.08)	(0.00)	(0.13)	(0.09)	(0.07)	(0.06)	(0.00)
Estimated		**Feasible MLE** $\left(\widehat{\beta}_{CSN}\right)$									
cut-off value		49%	39%	36%	30%	1%	46%	38%	35%	28%	1%
		0.96	1.13	1.24	1.40	2.82	0.95	1.12	1.23	1.40	2.84
		(0.15)	(0.11)	(0.10)	(0.08)	(0.04)	(0.11)	(0.08)	(0.07)	(0.06)	(0.03)
		Inverse of $\widehat{\delta}_{\max}$									
		1.09	1.06	1.06	1.07	1.10	1.16	1.11	1.10	1.10	1.11
		(N/A)	(N/A)	(N/A)	(N/A)	(N/A)	(0.21)	(0.15)	(0.14)	(0.12)	(0.07)

Notes: The DGP is given by (S.14). The true values of δ are generated as $\delta_{(i)} = 0.9^i$, for $i = 1, 2, ..., N$, where $\delta_{(i)}$ denotes the i^{th} largest δ. The true value of β is $\beta = 1/0.9 = 1.11$. $\hat{\beta}_{GI}$ denotes the Gabaix-Ibragimov estimate, $\widehat{\beta}_{MLE}$ denotes the maximum likelihood estimate, $\hat{\beta}_{CSN}$ denotes the feasible maximum likelihood estimate developed in Clauset et al. (2009). $\hat{\delta}_{\max}$ is the exponent estimate computed by (S.10). The standard error for the inverse of $\hat{\delta}_{\max}$ is computed by the delta method. (N/A) indicates that the standard error of $\hat{\delta}_{\max}$ cannot be computed when $T = 1$.

Table S.4: Estimates of the shape parameter, β, of the power law and inverse of the exponent, $\delta_{\max}$, under Exponent DGP with exponentially decaying $\delta_{(i)}$, where $\delta_{(i)} = 0.75^i$ ($\beta = 1/0.75 = 1.33$)

	$T=1$					$T=2$				
N	100	300	500	1,000	450,000	100	300	500	1,000	450,000
Assumed cut-off value	**Log-log regression** $\left(\widehat{\beta}_{GI}\right)$									
10%	1.36	1.39	1.47	1.60	2.39	1.34	1.39	1.47	1.60	2.39
	(0.61)	(0.36)	(0.29)	(0.23)	(0.02)	(0.42)	(0.25)	(0.21)	(0.16)	(0.01)
20%	1.34	1.45	1.54	1.66	2.11	1.34	1.46	1.54	1.67	2.11
	(0.42)	(0.27)	(0.22)	(0.17)	(0.01)	(0.30)	(0.19)	(0.15)	(0.12)	(0.01)
30%	1.31	1.45	1.52	1.63	1.91	1.32	1.45	1.53	1.63	1.91
	(0.34)	(0.22)	(0.18)	(0.13)	(0.01)	(0.24)	(0.15)	(0.13)	(0.09)	(0.01)
Assumed cut-off value	**Maximum Likelihood Estimation** $\left(\widehat{\beta}_{MLE}\right)$									
10%	1.61	1.67	1.75	1.86	2.11	1.48	1.63	1.72	1.85	2.11
	(0.51)	(0.30)	(0.25)	(0.19)	(0.01)	(0.33)	(0.21)	(0.17)	(0.13)	(0.01)
20%	1.46	1.56	1.61	1.68	1.79	1.40	1.54	1.60	1.67	1.79
	(0.33)	(0.20)	(0.16)	(0.12)	(0.01)	(0.22)	(0.14)	(0.11)	(0.08)	(0.00)
30%	1.35	1.43	1.47	1.51	1.58	1.31	1.42	1.46	1.50	1.58
	(0.25)	(0.15)	(0.12)	(0.09)	(0.00)	(0.17)	(0.11)	(0.08)	(0.06)	(0.00)
Estimated	**Feasible MLE** $\left(\widehat{\beta}_{CSN}\right)$									
cut-off value	41%	29%	24%	18%	1%	35%	25%	21%	16%	1%
	1.31	1.52	1.62	1.78	2.83	1.33	1.53	1.64	1.80	2.87
	(0.23)	(0.18)	(0.16)	(0.14)	(0.04)	(0.17)	(0.13)	(0.12)	(0.11)	(0.03)
	Inverse of $\widehat{\delta}_{\max}$									
	1.32	1.33	1.33	1.34	1.34	1.39	1.36	1.36	1.35	1.34
	(N/A)	(N/A)	(N/A)	(N/A)	(N/A)	(0.30)	(0.23)	(0.21)	(0.19)	(0.10)

Notes: The DGP is given by (S.14). The true values of δ are generated as $\delta_{(i)} = 0.75^i$, for $i = 1, 2, ..., N$, where $\delta_{(i)}$ denotes the i^{th} largest δ. The true value of β is $\beta = 1/0.75 = 1.33$. See the notes to Table S.3.

Table S.5: Bias, RMSE, size and power of the extremum estimator for the dominant units under Exponent DGP for unbalanced panels

	Experiment	A.1: One strongly dominant unit $\delta_{\max} = 1$	A.2: Two strongly dominant units $\delta_{(1)} = \delta_{(2)} = 1$		A.3: One strongly and one weakly dominant units $\delta_{(1)} = 1$, $\delta_{(2)} = 0.75$	
	N		$\delta_{(1)} = 1$	$\delta_{(2)} = 1$	$\delta_{(1)} = 1$	$\delta_{(2)} = 0.75$
Bias ($\times$100)	100	-1.25	3.87	-8.41	-1.64	-2.40
	300	-0.52	4.08	-5.83	-0.65	-0.94
	500	-0.38	3.95	-5.15	-0.45	-0.64
	1,000	-0.26	3.71	-4.47	-0.30	-0.41
	450,000	-0.08	2.08	-2.27	-0.08	-0.10
RMSE ($\times$100)	100	10.74	9.64	12.22	10.34	10.72
	300	8.65	8.27	9.25	8.46	8.71
	500	7.94	7.69	8.37	7.81	8.03
	1,000	7.14	7.01	7.43	7.06	7.24
	450,000	3.79	3.78	3.88	3.79	3.87
Size ($\times$100)	100	5.10	3.40	7.25	3.95	4.40
	300	4.55	3.85	5.55	3.80	4.55
	500	4.25	4.15	5.20	3.90	4.60
	1,000	4.15	4.05	4.65	3.75	4.30
	450,000	4.25	4.40	4.55	4.25	4.70
Power ($\times$100)	100	13.00	20.35	2.30	11.30	71.85
	300	18.80	33.30	3.75	18.00	83.90
	500	22.65	38.85	5.35	22.10	88.80
	1,000	26.50	45.65	7.20	26.10	94.35
	450,000	73.75	93.10	53.80	73.75	100.00

Notes: The unbalanced panels are generated with $T_{\max} = 4$. For each Monte Carlo replication, the top 5% of the units in terms of the true degree of dominance do not have missing observations, whereas the rest will have missing data for the first and the last periods with a 50% probability. The DGP is given by (S.14). For Experiment A.1, there is one strongly dominant unit and the rest are non-dominant: $\delta_{\max} = 1$, with $\delta_{(i)} = 0$ for $i = 2, 3, ..., N$. For Experiment A.2, there are two strongly dominant units and the rest are non-dominant: $\delta_{(1)} = \delta_{(2)} = 1$, with $\delta_{(i)} = 0$ for $i = 3, 4, ..., N$. For Experiment A.3, there are one strongly dominant unit and one weakly dominant unit, and the rest are non-dominant: $\delta_{(1)} = 1$ and $\delta_{(2)} = 0.75$, with $\delta_{(i)} = 0$ for $i = 3, 4, ..., N$. $\delta_{(i)}$ denotes the i^{th} largest δ, i.e., $\delta_{\max} = \delta_{(1)} \geq \delta_{(2)} \geq \delta_{(3)} \geq ...$, which are estimated by (S.18), and the standard errors of $\hat{\delta}_{(i)}$ are computed by (S.19). The power is calculated at 0.9 if true value is 1, and at 1 if true value is 0.75. The number of replications is $2,000$.

Table S.6: Bias, RMSE, size and power of the extremum estimator for the dominant units under Exponent DGP with heteroskedastic and serially corrleated errors

	Bias(×100)				RMSE(×100)				Size(×100)				Power(×100)			
$N \backslash T$	50	100	200	500	50	100	200	500	50	100	200	500	50	100	200	500
Experiment A.1: One strongly dominant unit, $\delta_{\max} = 1$																
500	-0.20	-0.18	-0.19	-0.19	3.92	2.77	1.95	1.23	10.15	8.05	6.90	5.45	76.15	95.55	99.95	100.00
1,000	-0.09	-0.08	-0.09	-0.09	3.52	2.48	1.74	1.10	9.85	7.85	6.90	5.70	84.70	98.35	100.00	100.00
2,000	-0.04	-0.03	-0.04	-0.04	3.20	2.25	1.58	1.00	9.70	7.80	6.60	5.60	90.90	99.55	100.00	100.00
5,000	-0.02	0.00	-0.01	-0.01	2.86	2.01	1.41	0.89	9.95	7.65	6.45	5.60	95.25	100.00	100.00	100.00
10,000	-0.01	0.01	0.00	0.00	2.64	1.86	1.31	0.82	10.10	7.70	6.55	5.50	97.40	100.00	100.00	100.00
Experiment A.2: Two strongly dominant units, $\delta_{(1)} = \delta_{(2)} = 1$																
$\delta_{(1)} = 1$																
500	1.53	0.92	0.53	0.20	3.11	2.11	1.47	0.87	7.80	5.50	5.05	2.85	98.25	100.00	100.00	100.00
1,000	1.54	0.99	0.64	0.34	2.88	1.97	1.38	0.83	8.25	5.75	5.70	3.60	99.60	100.00	100.00	100.00
2,000	1.48	0.99	0.66	0.39	2.66	1.83	1.29	0.79	8.85	6.35	5.85	4.70	99.85	100.00	100.00	100.00
5,000	1.37	0.93	0.64	0.40	2.40	1.66	1.18	0.73	9.25	6.65	6.50	5.40	100.00	100.00	100.00	100.00
10,000	1.28	0.88	0.61	0.39	2.23	1.55	1.10	0.69	9.35	6.70	6.65	5.70	100.00	100.00	100.00	100.00
$\delta_{(2)} = 1$																
500	-2.30	-1.70	-1.31	-0.96	3.57	2.55	1.90	1.29	12.05	10.10	10.55	13.15	73.35	94.60	99.90	100.00
1,000	-1.90	-1.36	-1.02	-0.70	3.10	2.18	1.60	1.05	10.60	8.70	8.70	8.65	83.60	98.15	100.00	100.00
2,000	-1.65	-1.15	-0.85	-0.56	2.77	1.93	1.40	0.90	10.30	7.80	7.70	7.10	90.35	99.55	100.00	100.00
5,000	-1.42	-0.98	-0.71	-0.45	2.44	1.69	1.22	0.77	9.95	6.95	7.00	5.90	95.25	100.00	100.00	100.00
10,000	-1.30	-0.89	-0.64	-0.40	2.25	1.55	1.12	0.70	9.80	7.10	6.75	5.80	97.40	100.00	100.00	100.00
Experiment A.3: One strongly dominant unit and one weakly dominant unit, $\delta_{(1)} = 1$, $\delta_{(2)} = 0.75$																
$\delta_{(1)} = 1$																
500	-0.35	-0.33	-0.34	-0.34	3.93	2.78	1.97	1.26	10.45	8.20	7.55	6.30	75.05	94.75	99.90	100.00
1,000	-0.17	-0.15	-0.16	-0.17	3.53	2.49	1.75	1.11	9.95	8.25	6.80	5.65	84.50	98.15	100.00	100.00
2,000	-0.08	-0.07	-0.08	-0.08	3.20	2.25	1.58	1.00	9.85	8.00	6.55	5.70	90.75	99.55	100.00	100.00
5,000	-0.03	-0.02	-0.03	-0.03	2.86	2.01	1.41	0.89	9.95	7.70	6.50	5.45	95.25	100.00	100.00	100.00
10,000	-0.01	0.00	-0.01	-0.01	2.64	1.86	1.31	0.82	10.15	7.70	6.50	5.55	97.40	100.00	100.00	100.00
$\delta_{(2)} = 0.75$																
500	-0.31	-0.34	-0.35	-0.32	2.63	1.77	1.34	0.86	9.40	6.55	8.05	8.30	99.95	100.00	100.00	100.00
1,000	-0.14	-0.16	-0.17	-0.15	2.34	1.56	1.17	0.73	8.85	6.15	7.45	6.35	100.00	100.00	100.00	100.00
2,000	-0.02	-0.08	-0.08	-0.06	2.22	1.41	1.06	0.66	9.15	6.00	7.05	5.95	99.80	100.00	100.00	100.00
5,000	0.00	-0.03	-0.03	-0.01	1.90	1.26	0.94	0.58	9.10	5.90	6.90	5.80	100.00	100.00	100.00	100.00
10,000	0.03	-0.01	-0.02	0.00	1.84	1.16	0.87	0.54	9.10	6.05	6.80	5.90	99.90	100.00	100.00	100.00

Notes: The DGP is given by (S.20), where the errors are generated by (S.21). $\delta_{\max} = \delta_{(1)}$ and $\delta_{(2)}$ are estimated by (S.3), and the standard errors are computed by (S.8). See also the notes to Table S.5.