Econometrica Supplementary Material

SUPPLEMENT TO "LONG MEMORY VIA NETWORKING"
(*Econometrica*, Vol. 86, No. 6, November 2018, 2221–2248)

SUSANNE M. SCHENNACH
Economics Dept., Brown University

This Supplemental Material includes various extensions of the paper's main results, namely, (i) deviations from power laws in the c_n coefficients, (ii) the presence of multiple sources of noise in the network, (iii) the possibility of non-integrable limiting power spectra, and (iv) heterogeneity in the agents' responses. It also includes the description of a simple and stylized variant of the Loss–Plosser model as well as a "toy" application based on the "input-output accounts" database compiled by the Bureau of Economic Analysis.

S1. SOME EXTENSIONS

S1.1. *Deviations From Power Laws*

THE ASSUMED POWER LAW BEHAVIOR for c_n in Theorem 1 may seem specific, but other natural possibilities yield either uninteresting or implausible results. One obvious generalization is $c_n = e^{\beta n} n^{-(1-\alpha)}$ for $\alpha, \beta \in \mathbb{R}$. However, the $\beta < 0$ case falls under case (ii) of Theorem 1 and yields a short-memory process. The case $\beta > 0$ yields a spectrum that diverges at all λ such that $|\tilde{r}(\lambda)| > e^{-\beta}$ and not just at $\lambda = 0$. In that case, even a perturbation of a finite duration would be magnified by the network to such an extent that the overall economy would leave the local equilibrium considered in a finite time and visit another equilibrium. The process would then presumably repeat itself until a stable equilibrium (with non-explosive c_n) is found. In a sense, the economy should plausibly self-organize to rule out cases where $\tilde{z}^\infty(\lambda)$ diverges for $\lambda \neq 0$. In this sense, $\beta = 0$ is the only nontrivial and plausible case. It is straightforward to extend Theorem 1 to allow for $\alpha > 1$, thus covering cointegrated processes (e.g., Avarucci and Velasco (2009)) or "mildly explosive" processes (e.g., Phillips and Magdalinos (2007)). (The necessary adjustments are outlined in footnote 11 in the Appendix of the main text, to avoid cluttering the main proof with lengthy manipulations.)

While the results of Theorem 1 are already robust to deviations from exact power laws that are absolutely summable, we can also handle deviations of the c_n coefficients from a power law that are bigger than absolutely summable. For instance, consider the case where the c_n (for $n \geq 1$) admit an expansion of the form

$$c_n = \sum_{i=1}^{\bar{\imath}} A_i n^{-(1-\alpha_i)} + c'_n, \tag{S1}$$

where $\alpha_1 > \alpha_2 > \cdots > \alpha_{\bar{\imath}}$ and $\sum_{n=1}^{\infty} |c'_n| < \infty$. One can apply Theorem 1 to each individual term to yield the conclusion that the resulting power spectrum $|\tilde{z}^\infty(\lambda)|^2$ would then have

Susanne M. Schennach: smschenn@brown.edu

The author would like to thank Lars Hansen, Eric Renault, Adam McCloskey, Florian Gunsilius, seminar participants at Cemmap/UCL, at the University of Cambridge, and at the Harvard/MIT Econometrics seminar, and anonymous referees for useful comments, and acknowledges support from NSF Grants SES-1061263/1156347, SES-1357401, and SES-1659334.

 https://doi.org/10.3982/ECTA11930

the behavior

$$\left|\tilde{z}^{\infty}(\lambda)\right|^{2}=\sum_{i=1}^{\bar{\imath}}O\left(|\lambda|^{-2\alpha_i}\right)=O\left(|\lambda|^{-2\alpha_1}\right)\quad\text{as }|\lambda|\to 0,$$

since $\alpha_1>\alpha_i$ for $i=2,\dots,\bar{\imath}$. Taking $\bar{\imath}$ finite is without much loss of generality, since eventually, for some α_i, the power law would become absolutely integrable (if consecutive exponents α_i are at least some finite distance from each other). Expansions of the form (S1) can be obtained, for instance, if the c_n coefficients can be written as $c_n=g(n^{-1})$, where $g(\cdot)$ is a function such that $(g(u))^a$ admits a Taylor expansion around $u=0$ for some real a, so this extension brings considerable generality.

S1.2. *Multiple Sources of Noise*

In this section, we consider the effect of multiple sources of noise with an arbitrary covariance structure introduced at multiple points of the network. We maintain the Gaussian assumption. It turns out that the general covariance case can always be reduced to the uncorrelated noise case (across the spatial dimension) by a suitable redefinition of the network. Specifically, consider again our general vector autoregressive setup $X_t=\sum_{s=0}^{\infty}W_sX_{t-s}+V^{1/2}u_t$, but where the noise now has the general form $V^{1/2}u_t$ for some general correlation matrix V and with u_t being an $N(0,I)$ noise vector. This model can equivalently be written via an augmented state vector $(X_t',X_t^{*\prime})'$ as

$$\begin{bmatrix}X_t\\X_t^*\end{bmatrix}=\sum_{s=0}^{\infty}\begin{bmatrix}W_s & V^{1/2}\mathbf{1}\{s=0\}\\0 & 0\end{bmatrix}\begin{bmatrix}X_{t-s}\\X_{t-s}^*\end{bmatrix}+\begin{bmatrix}0\\u_t\end{bmatrix},$$

which has the same basic form as Equation (1) with a noise that is spatially uncorrelated. This construction amounts to building a network with twice the number of nodes containing the original network (as modeled via W_s) and an additional network (modeled via V) whose role is solely to propagate each component of the uncorrelated noise vector u_t to multiple nodes of the original network.

For uncorrelated noise sources, we can easily compute the c_n coefficients via Equation (4) associated with one source node i at the time (setting all but one element of e^o to zero) while considering a given fixed set of destination nodes (via e^d). Let $|\tilde{z}_i^{\infty}(\lambda)|^2$ denote the power spectrum obtained when only source node i is active. Since the noise sources are independent, the overall power spectrum is simply the sum of the individual power spectra $\sum_i|\tilde{z}_i^{\infty}(\lambda)|^2$.

S1.3. *Non-integrable Power Spectra*

One can also establish a convergence result similar to Theorem 2 that covers both integrable ($\alpha<1/2$) and non-integrable ($\alpha\geq 1/2$) limiting power spectra $|\tilde{z}^{\infty}(\lambda)|^2$ by focusing on increments of the processes. Working with increments is a standard technique (see Mandelbrot and Ness (1968) and Comte and Renault (1996), for instance) that offers the advantage of providing finite-variance quantities even in the presence of nonstationarity in the process.

THEOREM S1: *Let the Assumptions of Theorem* 1 *hold*. *Assume that* $|\tilde{r}(\lambda)|<1$ *for* $\lambda\in\,]0,\pi]$, *that* $|\tilde{y}(\lambda)|$ *is uniformly bounded for* $\lambda\in[0,\pi]$, *and consider the differenced process*

$$\Delta Z_t^n\equiv Z_t^n-Z_{t-\Delta t}^n$$

for a given $\Delta t \in \mathbb{Z}$ *and any* $n \in \mathbb{N}$ (*with corresponding moving average representation* $\Delta z_t^n \equiv z_t^n - z_{t-\Delta t}^n$ *and spectrum* $\Delta \tilde{z}^n(\lambda) \equiv (1 - e^{i\lambda\Delta t})\tilde{z}^n(\lambda)$). *Let* $\tilde{z}^\infty(\lambda) \equiv \lim_{n\to\infty} \tilde{z}^n(\lambda)$ *with a corresponding moving average representation* z_t^∞. *Then, there exists a stationary process* ΔZ_t^∞ *with moving average representation*[1] $\Delta z_t^\infty \equiv z_t^\infty - z_{t-\Delta t}^\infty$ *and spectrum* $\Delta \tilde{z}^\infty(\lambda) \equiv (1 - e^{i\lambda\Delta t})\tilde{z}^\infty(\lambda)$ *satisfying* $\int_0^\pi |\Delta \tilde{z}^n(\lambda) - \Delta \tilde{z}^\infty(\lambda)|^2\, d\lambda \to 0$, $\sum_{t=0}^\infty |\Delta z_t^n - \Delta z_t^\infty|^2 \to 0$, *and* $E[|\Delta Z_t^n - \Delta Z_t^\infty|^2] \to 0$ *for any given* $t \in \mathbb{Z}$ *and* $\sum_{t=-\infty}^\infty E[|\Delta Z_t^n - \Delta Z_t^\infty|^2] w_t \to 0$ *for a given absolutely summable weighting sequence* w_t.

PROOF: The proof is similar to the one of Theorem 2 and we focus here on the differences. It is clear that the differenced process ΔZ_t^n admits the moving average representation:

$$\Delta Z_t^n = \sum_{s=-\infty}^{t} \left(z_{t-s}^n - z_{t-\Delta t-s}^n\right) G_s,$$

where the kernel $z_{t-s}^n - z_{t-\Delta t-s}^n$ is absolutely summable since it is a difference of two absolutely summable terms. Its Fourier transform is thus well-defined and equal to

$$\Delta \tilde{z}^n(\lambda) = \sum_{t=0}^{\infty} \left(z_t^n - z_{t-\Delta t}^n\right) e^{i\lambda t} = \tilde{z}^n(\lambda) - e^{i\lambda\Delta t}\tilde{z}^n(\lambda) = \left(1 - e^{i\lambda\Delta t}\right)\tilde{z}^n(\lambda).$$

The pointwise limit of $\Delta \tilde{z}^n(\lambda)$ also poses no problem (as in Theorem 2):

$$\Delta \tilde{z}^\infty(\lambda) \equiv \lim_{n\to\infty} \left(1 - e^{i\lambda\Delta t}\right)\tilde{z}^n(\lambda) = \left(1 - e^{i\lambda\Delta t}\right)\tilde{z}^\infty(\lambda),$$

with the additional advantage that $\Delta \tilde{z}^n(0) = 0$ and therefore $\Delta \tilde{z}^\infty(0) = 0$ (so the $\lambda = 0$ point is no longer exceptional).

Now observe that, for some sufficiently small $\bar{\lambda} > 0$,

$$\begin{aligned}
\int_0^\pi \left|\Delta \tilde{z}^\infty(\lambda)\right|^2 d\lambda &= \int_{\lambda \le \bar{\lambda}} \left|\left(1 - e^{i\lambda\Delta t}\right)\tilde{z}^\infty(\lambda)\right|^2 d\lambda + \int_{\bar{\lambda}}^\pi \left|\left(1 - e^{i\lambda\Delta t}\right)\tilde{z}^\infty(\lambda)\right|^2 d\lambda \\
&\le \int_{\lambda \le \bar{\lambda}} C_1 |\Delta t \lambda|^2 |\lambda|^{-2\alpha}\, d\lambda + \int_{\bar{\lambda}}^\pi 2\left|\tilde{z}^\infty(\lambda)\right|^2 d\lambda \\
&\le \int_{\lambda \le \bar{\lambda}} C_1 |\lambda|^{2(1-\alpha)}\, d\lambda + \int_{\bar{\lambda}}^\pi 2\left|\tilde{y}(\lambda)\right|^2 d\lambda < \infty
\end{aligned}$$

for some finite constant $C_1 > 0$ and where $1 - \alpha \ge 0$. Hence $\Delta \tilde{z}^\infty \in \mathcal{L}_2(\mathbb{R})$ and therefore the corresponding Δz^∞ is also in ℓ^2 and the corresponding process ΔZ_t^∞ is stationary.

Next, we again make use of Lebesgue's dominated convergence theorem to show that $\int_0^\pi |\Delta \tilde{z}^n(\lambda) - \Delta \tilde{z}^\infty(\lambda)|^2\, d\lambda \to 0$, which requires the existence of a square-integrable $\bar{z}(\lambda)$ such that $|\Delta \tilde{z}^n(\lambda) - \Delta \tilde{z}^\infty(\lambda)| \le \bar{z}(\lambda)$. For $|\lambda| \ge \bar{\lambda}$, we proceed as in Theorem 2 after noting that the prefactor $(1 - e^{i\lambda\Delta t})$ is bounded in magnitude by 2. For $|\lambda| \le \bar{\lambda}$, we proceed as in Theorem 2, after noting that the prefactor $(1 - e^{i\lambda\Delta t})$ is bounded in magnitude by $C_2|\lambda|$ for some finite $C_2 > 0$. This leads to a $\bar{z}(\lambda)$ that has the form $|\lambda|^{1-\alpha}$ (instead of $|\lambda|^{-\alpha}$), which is clearly square-integrable for $|\lambda| \le \bar{\lambda}$ for any $\alpha \in [0, 1]$. *Q.E.D.*

[1] We take the convention that $z_t^\infty = 0$ for $t < 0$.

S1.4. *Heterogeneity*

To allow for heterogeneity in the agents' responses, we relax Assumption 1 as follows.

ASSUMPTION S1: *The autoregressive coefficient matrix in Equation* (1) *factors as* $W_{s,ij} = r_{s,ij}W_{ij}$ *where the* W_{ij} *are fixed constants* (*satisfying* $\sum_{j=1}^{N} W_{ij} = 1$ *for* $i = 1, \dots, N$), *while the impulse response function* $r_{s,ij}$ *of each agent is chosen at random once at* $t = -\infty$ *and kept constant thereafter.*

The assumption allows for the effect of each input j on the output of each node i of the network to be characterized by a *different* convolution operation. We view the network structure as fixed (via the deterministic W_{ij}) and allow for heterogeneity in the agents (via the random impulse response functions $r_{s,ij}$). We place no specific assumption regarding the covariance structure between the different elements of $r_{s,ij}$, although we will need to constrain the amount of possible dependence.

This section provides conditions under which the conclusion of Theorem 1 actually holds with probability 1 for such randomly constructed networks. A key feature of the result is the existence of an average spectral representation denoted $\bar{r}(\lambda)$. In essence, there are so many very long pathways that connect the origin and the destination, that the fluctuations in the $r_{s,ij}$ across the different i, j quickly average out to a single effective value representative of the whole network. To state our result, we introduce a few convenient definitions that are heterogeneous analogues of previously defined quantities.

DEFINITION S1: Let $\tilde{r}_{ij}(\lambda) = \sum_{s=0}^{\infty} r_{s,ij} e^{i\lambda s}$. Let $\mathcal{P}_n$ denote the set of paths connecting the origin nodes to the destination nodes in n steps (each element p of $\mathcal{P}_n$ is an $(n+1)$-dimensional vector of integer specifying which sequence of nodes are visited by the path). For any maximum path length $\bar{n} \in \mathbb{N}$, the spectral representation of the aggregate output of the destination nodes is given by

$$\tilde{z}^{\bar{n}}(\lambda) = \tilde{y}(\lambda) \sum_{n=0}^{\bar{n}} \sum_{p \in \mathcal{P}_n} \prod_{\ell=1}^{n} \left(\tilde{r}_{p_\ell p_{\ell+1}}(\lambda) W_{p_\ell p_{\ell+1}}\right), \tag{S2}$$

and we let $c_n = \sum_{p \in \mathcal{P}_n} \prod_{\ell=1}^{n} W_{p_\ell p_{\ell+1}}$ (which coincides with the earlier definition via Equation (4) after expanding the matrix product).

Equation (S2) merely states that the output is the sum of the effect of the input noise (modeled via $\tilde{y}(\lambda)$) through the various possible pathways p, of lengths up to $\bar{n}$, joining the origin and the destination nodes. Along each path, the noise is filtered as it goes through the network. Going from node p_ℓ to node $p_{\ell+1}$, its spectral representation is multiplied by $\tilde{r}_{p_\ell p_{\ell+1}}(\lambda)$ (the spectral response of node $p_{\ell+1}$) and weighted by the link strength $W_{p_\ell p_{\ell+1}}$.

THEOREM S2: *Let* $\tilde{y}$ *satisfy Assumption* 2 *and let Assumption* S1 *hold. Let* $\bar{r}(\lambda) \equiv \lim_{n\to\infty} (\sum_{p \in \mathcal{P}_n} (\prod_{\ell=1}^{n} W_{p_\ell p_{\ell+1}}) E[\prod_{\ell=1}^{n} \tilde{r}_{p_\ell p_{\ell+1}}(\lambda)])^{1/n}$. *Assume that* $\bar{r}(\lambda)$ *exists, satisfies Assumption* 3, *and is such that*

$$E\left[\left(\sum_{p \in \mathcal{P}_n} \left(\prod_{\ell=1}^{n} W_{p_\ell p_{\ell+1}}\right)\left(\prod_{\ell=1}^{n} \frac{\tilde{r}_{p_\ell p_{\ell+1}}(\lambda)}{\bar{r}(\lambda)} - 1\right)\right)^2\right] \le D n^{-3-\varepsilon} \tag{S3}$$

for some $D, \varepsilon > 0$ *for all* λ *in some neighborhood of the origin. Then, the conclusion of Theorem* 1 *for* $\tilde{z}^{\bar{n}}(\lambda)$ *holds with probability* 1.

To prove this result, we first need a simple lemma.

LEMMA S1: *Let* c_n *be a deterministic sequence and let the corresponding* $\tilde{z}^\infty(\lambda)$ *satisfy* $\tilde{z}^\infty(\lambda) = A(i\lambda)^{-\alpha} + o((i\lambda)^{-\alpha})$ *(for* $A \in \mathbb{R}$ *and* $\alpha \in \mathbb{R}^+$*). Let* c'_n *be a random sequence such that* $E[(c'_n - c_n)^2] \le D(1+n)^{-3-\varepsilon}$ *for some* $\varepsilon, D > 0$*; then the corresponding* $\tilde{z}^{\infty\prime}(\lambda)$ *satisfies* $\tilde{z}^{\infty\prime}(\lambda) = A(i\lambda)^{-\alpha} + o((i\lambda)^{-\alpha})$ *with probability* 1.

PROOF: To simplify the notation, let the sequence start at index $n = 1$ instead of 0. By Lemma 5, it suffices to show that $\sum_{n=1}^\infty |c'_n - c_n|$ is finite with probability 1, that is, $P[\sum_{n=1}^\infty |c'_n - c_n| \ge C] \to 0$ as $C \to \infty$. Let $\Delta c_n = c'_n - c_n$, and for a given C, let $c = C(\sum_{n=1}^\infty n^{-1-\varepsilon/3})^{-1}$. Note that $\sum_{n=1}^\infty n^{-1-\varepsilon/3} < \infty$ and that $C \to \infty \Longrightarrow c \to \infty$. Then note that $|\Delta c_n| \le cn^{-1-\varepsilon/3}$ for all $n \in \mathbb{N}^*$ implies that $\sum_{n=1}^\infty |\Delta c_n| \le C$. Taking the contrapositive of that statement yields that the event $\sum_{n=1}^\infty |\Delta c_n| \ge C$ implies the event $|\Delta c_n| \ge cn^{-1-\varepsilon/3}$ for some $n \in \mathbb{N}^*$. Then write

$$
\begin{aligned}
P\left[\sum_{n=1}^\infty |\Delta c_n| \ge C\right] &\le P\big[|\Delta c_n| \ge cn^{-1-\varepsilon/3} \text{ for some } n \in \mathbb{N}^*\big] \\
&\le \sum_{n=1}^\infty P\big[|\Delta c_n| \ge cn^{-1-\varepsilon/3}\big] = \sum_{n=1}^\infty P\big[|\Delta c_n|^2 \ge c^2 n^{-2-2\varepsilon/3}\big] \\
&\le \sum_{n=1}^\infty \frac{E\big[|\Delta c_n|^2\big]}{c^2 n^{-2-(2/3)\varepsilon}} \le \sum_{n=1}^\infty \frac{Dn^{-3-\varepsilon}}{c^2 n^{-2-(2/3)\varepsilon}} = \frac{D}{c^2} \sum_{n=1}^\infty n^{-1-\varepsilon/3},
\end{aligned}
$$

where we have used, in turn, (i) the fact that if two events are such that $A \Longrightarrow B$, then $P[B] \ge P[A]$, (ii) for any sequence of events A_i, we have $P[\bigcup_i A_i] \le \sum_i P[A_i]$, (iii) monotonicity of the function x^2 for $x \ge 0$, (iv) Markov's inequality $P[X \ge x] \le E[X]/x$ applied to the random variable $X = |\Delta c_n|^2$, (v) the assumption $E[|\Delta c_n|^2] \le Dn^{-3-\varepsilon}$. Since $\sum_{n=1}^\infty n^{-1-\varepsilon/3} < \infty$, it follows that, as $C \to \infty$, $c \to \infty$ and $P[\sum_{n=1}^\infty |X_n| \ge C] \to 0$, as desired. *Q.E.D.*

PROOF OF THEOREM S2: From Definition S1, we have $c_n = \sum_{p \in \mathcal{P}_n} \prod_{\ell=1}^n W_{p_\ell p_{\ell+1}}$ and thus

$$
\begin{aligned}
\tilde{z}^{\bar{n}}(\lambda) &= \tilde{y}(\lambda) \sum_{n=0}^{\bar{n}} \sum_{p \in \mathcal{P}_n} \prod_{\ell=1}^n \big(\tilde{r}_{p_\ell p_{\ell+1}}(\lambda) W_{p_\ell p_{\ell+1}}\big) \\
&= \tilde{y}(\lambda) \sum_{n=0}^{\bar{n}} \big(\bar{r}(\lambda)\big)^n \sum_{p \in \mathcal{P}_n} \left(\prod_{\ell=1}^n W_{p_\ell p_{\ell+1}}\right) \left(\prod_{\ell=1}^n \frac{\tilde{r}_{p_\ell p_{\ell+1}}(\lambda)}{\bar{r}(\lambda)}\right) \\
&= \tilde{y}(\lambda) \sum_{n=0}^{\bar{n}} \big(\bar{r}(\lambda)\big)^n \sum_{p \in \mathcal{P}_n} \prod_{\ell=1}^n W_{p_\ell p_{\ell+1}} + \sum_{p \in \mathcal{P}_n} \left(\prod_{\ell=1}^n W_{p_\ell p_{\ell+1}}\right) \left(\prod_{\ell=1}^n \frac{\tilde{r}_{p_\ell p_{\ell+1}}(\lambda)}{\bar{r}(\lambda)} - 1\right) \\
&= \tilde{y}(\lambda) \sum_{n=0}^{\bar{n}} \big(\bar{r}(\lambda)\big)^n \left(c_n + \sum_{p \in \mathcal{P}_n} \left(\prod_{\ell=1}^n W_{p_\ell p_{\ell+1}}\right) \left(\prod_{\ell=1}^n \frac{\tilde{r}_{p_\ell p_{\ell+1}}(\lambda)}{\bar{r}(\lambda)} - 1\right)\right) \\
&= \tilde{y}_0(\lambda) \sum_{n=0}^{\bar{n}} (c_n + \Delta c_n)\big(\bar{r}(\lambda)\big)^n,
\end{aligned}
$$

where

$$\Delta c_n = \sum_{p \in \mathcal{P}_n} \left(\prod_{\ell=1}^{n} W_{p_\ell p_{\ell+1}} \right) \left(\prod_{\ell=1}^{n} \frac{\tilde{r}_{p_\ell p_{\ell+1}}(\lambda)}{\bar{r}(\lambda)} - 1 \right).$$

Hence, Lemma S1 applies directly when Δc_n satisfies the variance bound assumed in the present theorem. *Q.E.D.*

Condition (S3) is stated in somewhat high-level form for maximum generality, but it is relatively easy to realize that it is a weak restriction. This condition places a limit on the order of magnitude of the variance of a certain average. (The weighting factor $\prod_{\ell=1}^{n} W_{p_\ell p_{\ell+1}}$ sums up to 1 over all paths in $\mathcal{P}_n$, so the sum is a weighted average.) This average is taken over all possible pathways and effectively samples the spectral representation of the impulse response of a large number of agents. Typically, the number of possible pathways of length n is an exponentially increasing function of n (because at each node, there are a certain number of possible ways to go and these alternatives multiply to give the number of paths). Hence, unless the covariance of the summand across two pathways is extremely strong, the decrease of the variance of the average with n should often satisfy the bound (S3).

Note that (S3) bounds the heterogeneity in the response of paths, while placing only weak restrictions on the heterogeneity in the response of individual agents. Even if the economy is characterized by agents whose response $\tilde{r}_{ij}(\lambda)$ varies significantly with i and j, it is still plausible that the response $\prod_{\ell=1}^{n} \tilde{r}_{p_\ell p_{\ell+1}}(\lambda)$ of most paths $p \in \mathcal{P}_n$ could be very similar due to an averaging effect over the responses of many different agents sampled along the path. This assumption is plausible even in an economy with a mixture of very large firms (e.g., banks that are "too big to fail," such as some banks in the recent banking crisis) and very small firms. In that case, as most paths will likely go through some of the same large firms, the responses $\prod_{\ell=1}^{n} \tilde{r}_{p_\ell p_{\ell+1}}(\lambda)$ of two paths would tend to be quite similar, since they would often include some identical $\tilde{r}_{p_\ell p_{\ell+1}}(\lambda)$ terms. The fact that only the average $\bar{r}(\lambda)$ needs to satisfy Assumption 3, and not the individual $\tilde{r}_{ij}(\lambda)$, brings considerable generality to the result. In particular, the constant results to scale assumption need not hold at the node level but only at a global level.

S2. A SIMPLIFIED LONG AND PLOSSER MODEL

S2.1. *Model*

In this section, we show how the Long and Plosser model (hereafter LP) and its solution can be specialized to our setup where there are no separate labor inputs. LP's production function has the form

$$q_{it} = \eta_{it} \ell_{it-1}^{b_i} \prod_{j=1}^{N} q_{ij,t-1}^{W_{ij}}, \tag{S4}$$

where ℓ_{it} is labor inputs for the production of good i and b_i is a parameter such that the constant returns to scale $b_i + \sum_{j=1}^{N} W_{ij} = 1$ constraint holds. All other variables are as in our model. LP's representative consumer maximizes his expected discounted utility:

$$u_t = E\left[\sum_{s=t}^{\infty} \beta^{t-s} Z_t^{\theta_0} \prod_{i=1}^{N} c_{is}^{\theta_i} \;\middle|\; q_{t-1}, \eta_{t-1} \right], \tag{S5}$$

where Z_t is leisure, equal to $H - \sum_{j=1}^{N} \ell_{js}$ where H is the total labor available, and θ_0 is a parameter and all other variables are as in our model. Defining

$$\gamma_j \equiv \theta_j + \beta \sum_{i=1}^{N} \gamma_i W_{ij},$$

LP showed that the solution to this model is

$$\begin{aligned}
c_{it} &= \left(\frac{\theta_i}{\gamma_i}\right) q_{it}, \\
Z_t &= \theta_0 \left(\theta_0 + \beta \sum_{i=1}^{N} \gamma_i b_i\right)^{-1} H, \\
q_{ij,t} &= \left(\frac{\beta \gamma_i W_{ij}}{\gamma_j}\right) q_{jt}, \\
\ell_{it} &= \beta \gamma_i b_i \left(\theta_0 + \beta \sum_{j=1}^{N} \gamma_j b_j\right)^{-1} H, \\
\ln q_t &= W \ln q_{t-1} + k + \ln \eta_t,
\end{aligned}$$

where k is a vector of constants and the ln function is applied element-by-element.

Our production function is a special case of Equation (S4) obtained in the limit as $b_i \to 0$ while adjusting W_{ij} to preserve the constant returns to scale constraint. As a result, the solution to our model reduces to

$$\begin{aligned}
c_{it} &= \left(\frac{\theta_i}{\gamma_i}\right) q_{it}, \\
Z_t &= H, \\
q_{ij,t} &= \left(\frac{\beta \gamma_i W_{ij}}{\gamma_j}\right) q_{jt}, \\
\ell_{it} &= 0, \\
\ln q_t &= W \ln q_{t-1} + k + \ln \eta_t,
\end{aligned}$$

and substituting the solution $Z_t = H$ into the utility yields

$$u_t = E\left[\sum_{s=t}^{\infty} \beta^{t-s} H^{\theta_0} \prod_{i=1}^{N} c_{is}^{\theta_i} \,\middle|\, q_{t-1}, \eta_{t-1}\right],$$

which is equivalent to our utility (Equation (S5)) up to an irrelevant multiplicative constant H^{θ_0}. Observe that the solution remains well-behaved in the limit of $b_i \to 0$. In particular, the form of the time-evolution of $\ln q_t$ is preserved; the only difference is that the coefficients W_{ij} must now satisfy $\sum_{j=1}^{N} W_{ij} = 1$ instead of $\sum_{j=1}^{N} W_{ij} = 1 - b_i < 1$. Within the original Long–Plosser model, when $b_i > 0$, labor's ability to adjust instantaneously effectively dampens the noise and always yields exponentially decaying c_n coefficients (since $\sum_{j=1}^{N} W_{ij} < 1$) and thus short-memory processes as solutions. The limit $b_i \to 0$ leads to more interesting long-memory dynamics in the large-network limit.

It should be noted that the absence of a separate labor input ($b_i \to 0$ limit) does not mean that the model does not allow for labor inputs. Labor can be supplied via the network and treated symmetrically as part of the remaining inputs q_t. The limit $b_i \to 0$ then implies that the fraction of labor input that can adjust instantaneously to shocks is infinitesimal, which is arguably no less plausible than assuming that the entire labor force can adjust instantaneously to shocks.

S2.2. *Empirical Example*

One way to empirically assess if the proposed mechanism for long-memory generation is plausible is to verify if the c_n coefficients in a toy model based on real economic network data indeed obey a power law with the appropriate exponent. For this purpose, we use the so-called "input-output accounts" database compiled by the Bureau of Economic Analysis describing interactions between sectors of the U.S. economy. We use the most disaggregated version of these data since it already contains all the information about information propagation (or "diffusion") over all scales, small and large. This strategy enables a plot of $(\ln(c_n), \ln(n))$ over as many orders of magnitude as possible, thus facilitating the identification of a linear trend.

We construct the network following the same procedure as in Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012), using a reconstructed Commodity-by-Commodity Direct Requirements table for year 2002, available in their supplementary material. These represent the equilibrium cost shares of each commodity j in the production of another commodity i. (Following Acemoglu et al. (2012), we use the terms industries and commodities interchangeably.) In the Long and Plosser-type model, these shares are equal to the Cobb–Douglas parameters W_{ij} of the production function (Equation (S4) with $b_i = 0$). We include an additional node ℓ in the network to model labor supply. In the same spirit as in Acemoglu et al. (2012) (see p. 1998), and in accordance with our constant return to scale assumption, we set the labor share in the production of good i to $W_{i\ell} = 1 - \sum_{j \neq \ell} W_{ij}$.

To close the loop, the labor force must take input from the economy for their livelihood. We do not have quantitative data on this; hence, we assume that the workers take inputs from all industries $j = 1, \dots, (N-1)$ with equal equilibrium share $W_{\ell j} = \rho/(N-1)$ and from each other with share $W_{\ell\ell} = 1 - \rho$. We used $\rho = 0.75$, but the results are not very sensitive to this parameter.

In this empirical example, there is no reason to expect that the c_n coefficients should be the same for every choice of source and destination node. As an example, we pick the group of industries that are numbered, according to North American Industry Classification System (NAICS), with a leading "2". These correspond largely to primary sector industries (such as mining and utilities). We compute the c_n coefficients via Equation (4), setting both the destination vector e^d and origin vector e^o to be a vector selecting all industries in this group. This corresponds to computing the spectrum of the aggregate response of this group of industries to a common shock.

The resulting c_n coefficients are shown in Figure 1 and reveal evidence of a power law $c_n = n^{-\gamma}$ in this industry group with an exponent of $\gamma \approx 0.58$, as obtained with a standard linear least squares regression of the data in logarithmic form. This corresponds to $\alpha = 1 - \gamma \approx 0.42$, that is, a power spectrum behaving as $|\lambda|^{-2\alpha} = |\lambda|^{-0.84}$ near the origin, resulting in a long-memory network behavior of a fractionally integrated nature of order $\alpha \approx 0.42$. Although this is, strictly speaking, a finite network, one can still observe a behavior that would be expected from an infinite network for "short" paths, because "short" paths do not "feel" the boundary of the network. Of course, if we increased the range

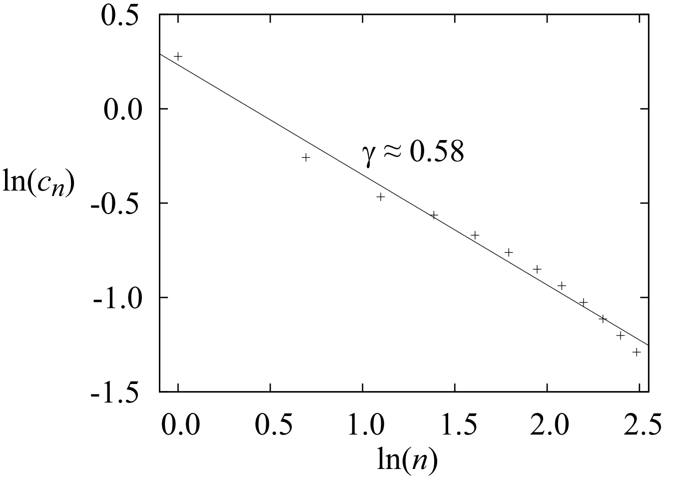

FIGURE 1.—Evidence of power law scaling $n^{-\gamma}$ with $\gamma \approx 0.58$ in the c_n coefficients (i.e., the probability of reaching a given point of the network after n steps of a random walk) in a network representing the U.S. economy as 418 "sectors."

of n, the graph would flatten out, as would be expected for a finite network (since the c_n would be asymptotically constant in that case).

We can pursue this example a bit further and explicitly calculate the spectrum associated with the power law $c_n \propto n^{-0.58}$ for our simplified Long–Plosser model. We employ the expression $\tilde{z}^{\bar{n}}(\lambda) = \sum_{n=0}^{\bar{n}} c_n(\tilde{r}(\lambda))^n \tilde{y}(\lambda)$, in which $\tilde{r}(\lambda) = e^{i\lambda}$ (since there is a single lag in the autoregressive representation in this model) and $\tilde{y}(\lambda) = 1$ (assuming a standard white noise as noise source). Figure 2 illustrates how $\tilde{z}^{\bar{n}}(\lambda)$ converges to a power law $\lambda^{-\alpha}$ as $\bar{n}$ increases. One can see that, as $\bar{n} \to \infty$, the oscillations around the limiting power law decrease in magnitude and the interval over which the spectrum is well described by a power law expands towards zero frequency.

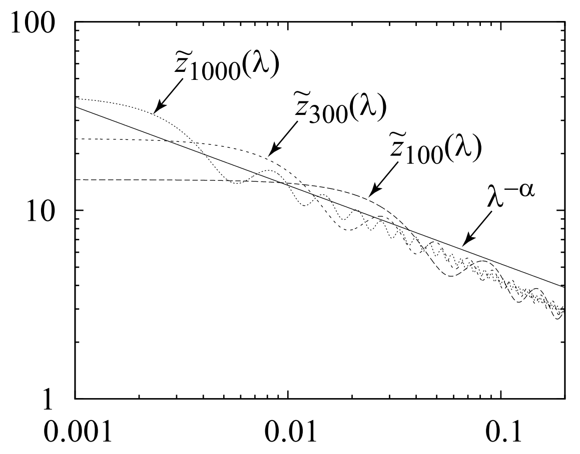

FIGURE 2.—Convergence of the simulated spectrum $z^{\overline{n}}(\lambda)$ to a power law ($\lambda^{-\alpha}$, with $\alpha = 0.42$), as the maximum path length $\overline{n}$ increases to infinity.

REFERENCES

ACEMOGLU, D., V. M. CARVALHO, A. OZDAGLAR, AND A. TAHBAZ-SALEHI (2012): "The Network Origins of Aggregate Fluctuations," *Econometrica*, 80, 1977–2016.[8]

AVARUCCI, M., AND C. VELASCO (2009): "A Wald Test for the Cointegration Rank in Nonstationary Fractional Systems," *Journal of Econometrics*, 151, 178–189.[1]

COMTE, F., AND E. RENAULT (1996): "Long Memory Continuous Time Models," *Journal of Econometrics*, 73, 101–149.[2]

MANDELBROT, B. B., AND J. W. V. NESS (1968): "Fractional Brownian Motions, Fractional Noises and Applications," *SIAM Review*, 10, 422–437.[2]

PHILLIPS, P., AND T. MAGDALINOS (2007): "Limit Theory for Moderate Deviations From a Unit Root," *Journal of Econometrics*, 136, 115–130.[1]

Co-editor Lars Peter Hansen handled this manuscript.

Manuscript received 25 September, 2013; final version accepted 15 June, 2018; available online 26 July, 2018.