Online Appendix to "The role of granularity in the variance and tail probability of aggregate output"

Yoshiyuki Arata*

March 28, 2020

Factset

I test the empirical validity of the granular hypothesis by using firm-level data compiled by Factset and the formulas obtained in Sections 3 and 4. I find that consistent with the results in Section 5, the empirical granularity is an important factor for the variance of aggregate output but not for the macro tail probability.

Data description

The dataset provided by Factset is covers listed and non-listed firms (mostly large firms) across firms. An important feature of this dataset is that it collects the information of customer-supplier relationships. This information is collected from public source such as financial reports of firms, websites, and news articles.

To begin with, I perform several filtering steps. First, I focus on samples in 2016-2017. Next, by excluding firms in financial and banking sector, I focus on firms in industrial, utility, and transportation sectors. The number of firms in my samples is reduced to 29776. Summary statistics of sales and the number of employees are given in Table 1.

Figure 1 shows the complementary cumulative distribution function (CCDF) o firm sales in the log-log scale, where I remove firms reporting sales equal to 0. As in Section 5, the tail of the CCDF is close a straight line in the log-log scale, implying that a Pareto tail is a good approximation. Hill's estimate of exponent α for the firm size distribution is $\widehat{\alpha} = 1.28$(s.e.0.051).

I construct the productivity growth of firm i, $\widehat{\epsilon}_{it}$, as in Section 5. Summary statistics is also given in Table 1. Figure 2 show that the density estimation of $\widehat{\epsilon}_{it}$. Hill's estimate of the left tail of $\widehat{\epsilon}_{it}$ is given by 3.26(s.e. $= 0.081$).

Network and Domar weights

Next, I calculate the Domar weight λ_i by using Equation (2). I assume that $\beta_i = 0$ for all i, i.e. the estimated λ_i represent network heterogeneity only. Before giving the estimates of λ_i, I give summary statistics of the customer-supplier network in my sample. The network consists of 29776 firms and 124167 network ties among them. In particular, the number of firms having at least one network tie is 19098. Figures 3 and 4 show the CCDFs of in- and out-degree of network ties in the log-log scale, i.e. the number of customers and suppliers for each firm, respectively. As expected, the CCDFs are close to a straight line, i.e. a Pareto tail. This implies that there exist a

*Research Institute of Economy, Trade and Industry (RIETI), 1-3-1 Kasumigaseki Chiyoda-ku, Tokyo, 100-8901 Japan. E-mail: arata-yoshiyuki@rieti.go.jp.

var.	# of obs.	mean	1st Q.	3rd Q.	max
s_i	29776	1633.9	30.4	640.8	485144
L_i	29776	6991	249	3860	2300000

var.	# of obs.	mean	sd
Δs_i	24529	0.0830	0.243
ΔL_i	24529	0.0469	0.302
$\hat{\epsilon}_i$	24529	0.0760	0.231

Table 1: Summary statistics

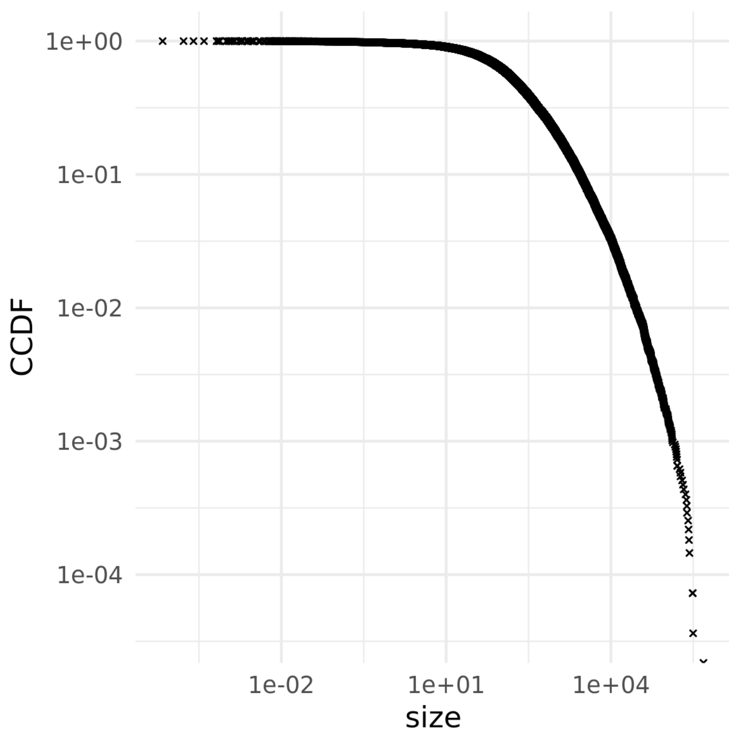

Figure 1: The CCDF of firm sales

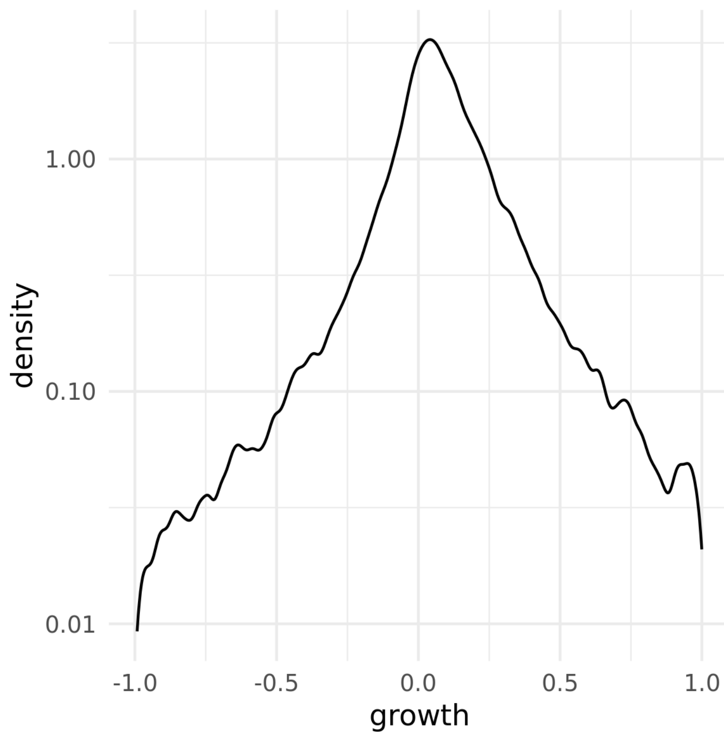

Figure 2: The density estimation of productivity growth

small number of firms which play a role of hubs in the network and have high Domar weights. To check the connectivity of the network, I also calculate the largest (weakly and strongly) connected components. The largest weakly (strongly) connected components contains 18842 (9085) firms.

Table 1 shows summary statistics of λ_i calculated from Equation (2). Figure 5 shows the CCDF of λ_i in the log-log scale, which implies the high heterogeneity of λ_i. Hill's estimate of the tail of the distribution of λ_i is $\widehat{\alpha} = 1.86$(s.e. $= 0.061$), which is in the range of $1 < \alpha < 2$. Hence, as in Section 5, I apply Propositions 2 and 4.

First, my calculation shows that the largest $\|\lambda\|_\infty$ is 0.0224. Since the sample standard deviation of $\widehat{\epsilon}_i$ is 23%, this implies that the larges firm leads to 0.52% standard deviation of GDP growth rate. The total of the standard deviation of GDP growth rate induced by microeconomic shocks is $\sigma_\epsilon \times \|\lambda\|_2 = 0.97\%$. Consistent with the analysis in Section 5, the granularity is an important factor for the variance of GDP growth rate.

On the other hand, the contribution of microeconomic shocks to the macroeconomic tail probability is negligible. I calculate $\sum_i \lambda_i^\beta$ in the case of a Pareto tail, which turns out to be 4.04×10^{-6}. This shows that the granularity is too low to lead to the large deviation of GDP growth rate.

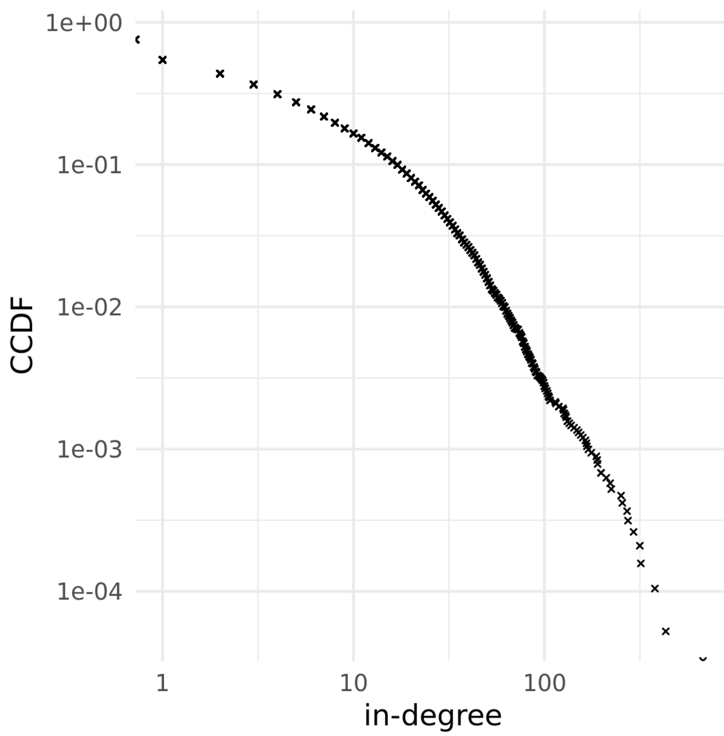

Figure 3: The CCDF of in-degree

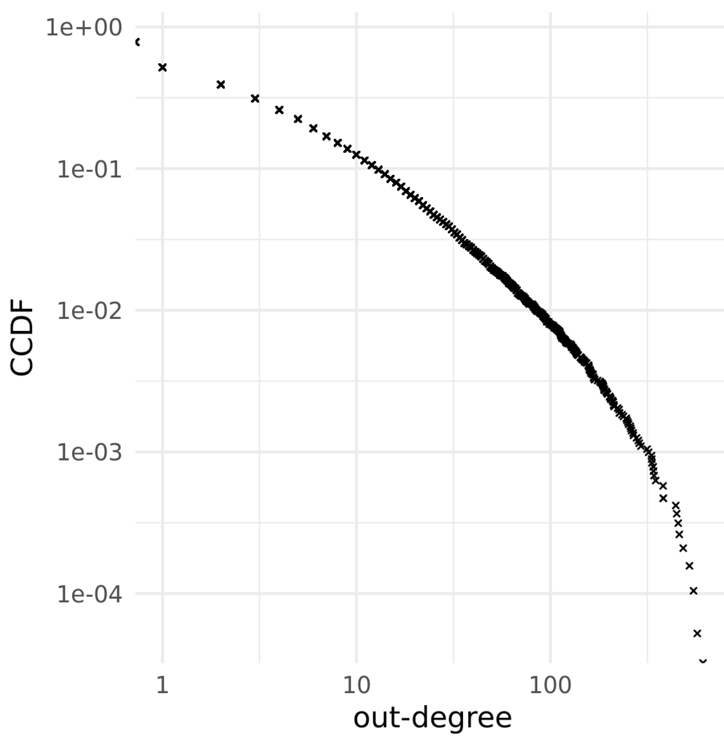

Figure 4: The CCDF of out-degree

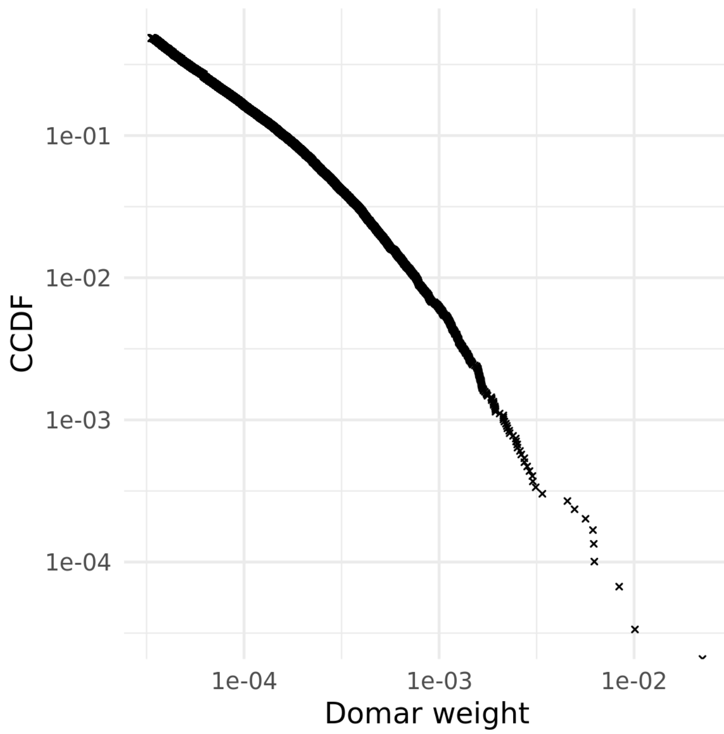

Figure 5: The CCDF of Domar weights