ECON641 – Problem Set 2

Anirudh Yadav

December 9, 2018

Contents

1 The Firm Size Distribution

1.1 Power law in firm size

A random variable S follows a power law if

$$\Pr[S > s] = Cs^{-\zeta}, \text{ with } C, s > 0 \tag{1}$$
$$\implies \log \Pr[S > s] = \log C - \zeta s.$$

Note that Zipf's law refers to a power law distribution with exponent $\zeta \approx 1$. Recent research has shown that the distribution of firm size is approximately described by Zipf's law. Below, we're going to explore the application of these ideas to Compustat data.

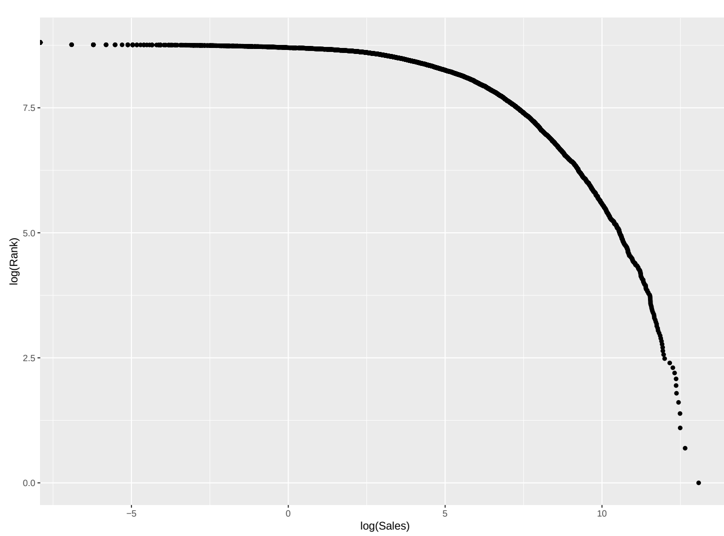
Before I get into the data work, let's briefly try to understand why we're doing all of this log-rank/log-size stuff. Suppose S follows a power law as in (1). Then, draw N realizations and rank them in descending order $S_{(1)} > S_{(2)} > ... > S_N$. Because of the ranking, we get $i/N = 1 - F(S_{(i)})$ (since $F(S_{(i)})$ is distributed standard uniform). Thus, $i/N = CS_{(i)}^{-\zeta}$, or

$$\text{Rank} = NCS_{(i)}^{-\zeta}$$
$$\implies \log \text{Rank} = \text{constant} - \zeta \log S_{(i)}.$$

1.1.1 Firm sales

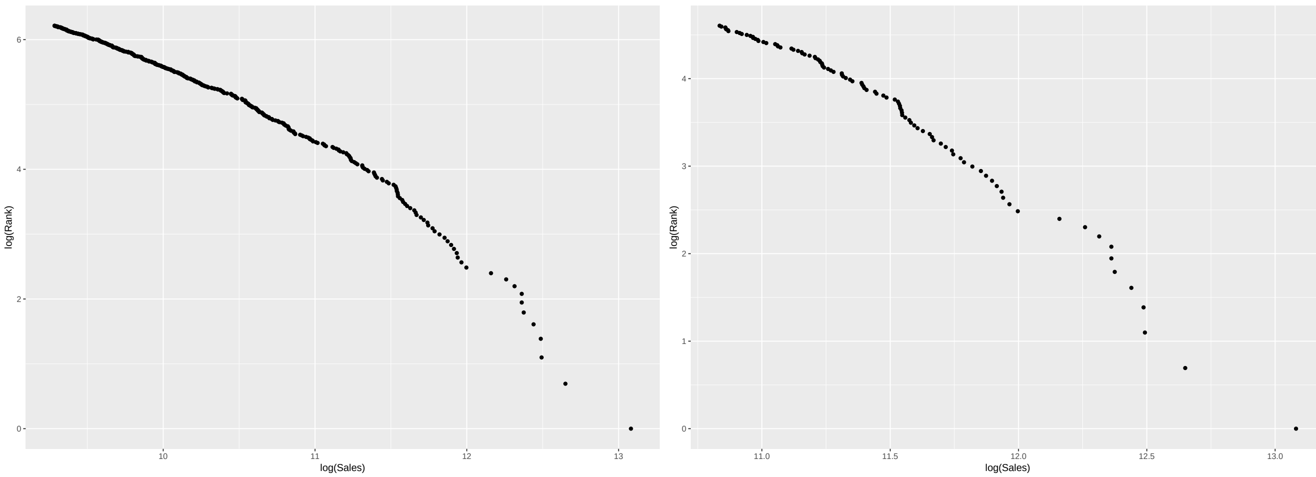
First, I plot the log–log plot of rank vs. sales for all firms in the Compustat database for 2015. Clearly, the relationship is not linear, suggesting that the firm size distribution does not follow a power law. This result is consistent with Stanley, *et. al.* (1995), who also use Compustat data. Note that Axtell (2001) uses US Census data (which obviously contains a far more comprehensive sample of firms compared to Compustat) and finds that the firm size distribution does indeed follow Zipf's law.

Figure 1: **Sales Distribution of All Firms in Compustat in 2015**



Next, I plot the log–log sales plots for the top 500 and 100 firms in the Compustat database. In both cases, there seems to be a linear relationship, suggesting that the tail of the firm sales distribution follows a power law.

Figure 2: **Sales Distribution for Top 500 and Top 100 Firms in 2015**



(a) Top 500 firms (b) Top 100 firms

Next, I estimate the power law coefficient ζ using the estimator proposed by Gabaix and Ibragimov (2011) for the samples of firms above. That is, I estimate the model

$$\log(\text{Rank}_i - 1/2) = \text{constant} - \zeta \log S_i + \epsilon_i$$

using OLS. I compute standard errors as in Gabaix and Ibragimov (2011). Table 1 shows the estimation results – suggesting that the sales distribution of the top 500 firms is approximately follows Zipf's law.

Table 1: **Estimated Power Law Coefficients for Firm Sales in 2015**

	For different lower size cutoffs		
	All firms	Top 500	Top 100
$\hat{\zeta}$	-0.27	-1.28	-2.06
Std. err.	0.00	0.08	0.29

1.1.2 Firm sales by industry

1.1.3 Employment

1.1.4 Employment by industry

1.1.5 Robustness check: 1985 data