

企業間関係と企業規模

齊藤(梅野)有希子・渡辺 努*

本稿では日本の法人企業約82万社(全法人企業の約3分の1)をカバーするデータセットを用いて、企業の資本・取引関係数が企業規模とどのように関係するかを調べた。第1に、企業の資本・取引関係数の分布はロングテールである。資本・取引関係数の多い企業の上位1%がもつ関係数は全関係数の約50%であり、偏在している。第2に、取引関係数が多いハブ企業だけを取り出しその相互関係をみると、一部の超ハブ企業に関係数が集中しており、偏在が一層顕著である。第3に、企業規模が大きいほど関係数は多く全体として両者は比例関係にある。しかし既に多くの関係をもつ企業では、規模が拡大するほどには関係数を増やしていない。これは関係の維持コストを企業が節約しているためと解釈できる。

1. はじめに

企業活動は様々な相互関係性の上に成り立っている。第1は物流である。企業は素原材料を川上の企業から購入する一方、生産物を川下の企業に対して中間投入として販売したり、流通企業に対して最終財として販売したりする。第2に、物流はしばしば企業間の与信関係を伴う。商品取引の決済時点を遅らせるために手形が発行されるのはその代表例であるが、より一般的に、商品の受渡しと資金の受渡しのタイミングがずればそこに企業間の与信が発生する。さらに企業間与信が銀行与信に置き換えられることも少なくない。第3は資本関係である。親会社の子会社を作るというだけでなく、物流で密接な関係にある企業が資本関係を結ぶことも少なくない。第4は役員派遣などの人的な交流である。

これらの企業間関係についてはこれまで経済学や経営学の様々な分野で多くの研究がなされてきた。企業と銀行の関係を例にとると、メインバンクやリレーションシップ・バンキングといった文脈で、企業と銀行の間に持続的な関係が成立するのはなぜか、持続的な関係をもつことは企業と銀行それぞれにどのようなメリットをもたらすのかといった点を巡って議論が展開

されてきた¹⁾。

このような企業間あるいは企業と金融機関の間の関係性は、企業(金融機関を含む)の間に成立するネットワークとみることができる。本稿の目的は、ネットワーク分析の手法を適用することにより企業間関係について新たな視点を提示することである。

ネットワーク分析は主として自然科学の分野で発展し、最近では社会学や経営学などにも広く応用されている²⁾。企業間関係の視点から特に興味深いのは「スケールフリー」とよばれるネットワーク構造である。このタイプのネットワークは、持てるもの(極端に多くの関係をもつ主体)と持たざるもの(ごく限られた関係をもつ主体)から構成される。有名な例はインターネットのホームページであり、非常に多くのリンクが貼られている少数の人気サイトと、リンクの少ない無数のサイトという偏在、つまり中心(ハブ)と周辺の二極化がその特徴である。タンパク質相互作用や代謝経路など生物学におけるネットワークもスケールフリー型の例である。

企業の場合も物流や資本関係の中核となるごく少数のハブ企業と、それにぶら下がる多数の中小企業で構成されており、スケールフリー型に近い性質をもつと考えられる。スケールフリー型のネットワークについては、少数のノード

にリンクが集中しているため、リンク数(関係数)の分布を描くとロングテールになり、分布の裾はパレート分布(または、べき分布)に従うことが知られている。パレート分布はもともと富の偏在を描写するためのものであったが、その同じ分布が関係性の偏在を描写する上でも有用なのである。本稿では、こうした性質が企業間関係についても満たされるか否かをデータを用いて検証する。

ネットワーク分析の手法を企業間関係に適用する際には関係数について十分なデータ量を確保する必要がある。本稿では、日本の法人企業約82万社(全法人企業の約3分の1)について他の法人企業との資本・取引関係を記述しているユニークなデータセットを用いる。このデータセットでは、各企業について、(1)仕入先企業、(2)販売先企業、(3)資本関係先企業の3種類の関係先企業名がリストアップされている。企業間ネットワークに関する内外の先行研究は、上場企業など特定の企業群や特定地域・産業に絞って分析されてきたが(例えば Souma *et al.* (2005, 2006), Garlaschelli *et al.* (2005), Caldarelli *et al.* (2006) など)、これらの研究と比較すると本稿で用いるデータセットのカバレッジは高く企業間ネットワークの分析に適している。

本稿の構成は以下のとおりである。第2節では本稿で用いるデータについて解説する。第3節では本稿で用いる分析手法について説明する。第4節では分析結果を説明する。第5節は本稿の結論である。

2. データ

本稿では東京商工リサーチが作成したデータセットを用いる。このデータセットには中小企業を含む826169社の法人企業が収録されている。それぞれの企業について、企業間関係データと企業情報データがある。企業間関係データ

は、各企業について関係を有する企業コードを列挙したものである。「関係」の種類としては、「仕入先」、「販売先」、「大株主」の3種類があり、それぞれについて最大24社まで列挙できる。関係の調査時点は2003年であり、全体として約400万の関係が収録されている。また、企業情報データには、各企業コードに対応する企業名、産業分類、地域、売上高、営業利益、創業年度が収録されている。

このデータセットの第1の特徴はそのカバレッジである。対象企業数82万社は日本に存在する全法人企業の約1/3に相当し、企業間関係のデータセットとしては国内はもとより海外でも例のない規模であり、実際の企業間ネットワークにより近いものが再現できている³⁾。ただし、個々の企業が関係先として列挙する企業の数には24社という上限が設けられているため、必ずしも全取引関係を網羅しているわけではない点には注意が必要である。

第2の特徴は物流関係(「仕入先」と「販売先」)と資本関係(「大株主」)という2つの異なるタイプの関係を捉えていることである。これら2つの関係が同じ傾向をもつのか、それとも異なるのかを見ることができる。

3. 分析手法

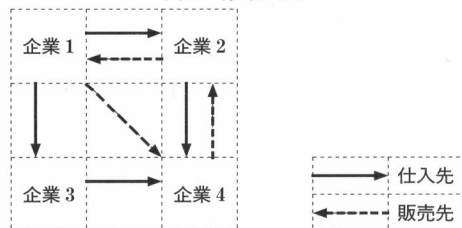
3.1 関係数の数え方

まず関係数を数えるところから始めよう。いま「仕入先」と「販売先」が表1のように与えられているとする。通常の物流関係では、企業1にとって企業2が仕入先であるならば、企業2は企業1にとって販売先である。しかし本稿で用いるデータセットではこのような双方向性が必ず満たされるわけではない。表1の例では、企業1は企業3を仕入先として列挙しているが、企業3は企業1を販売先として列挙していない。こうしたことが生じるのは、仕入先、販売先と

表1. 企業間関係データの例示

	#1	#2	#3	#4	#5	#6	#7
指名する企業	企業1	企業1	企業1	企業2	企業2	企業3	企業4
指名される企業	企業2	企業3	企業4	企業1	企業4	企業4	企業2
関係区分	「仕入先」	「仕入先」	「販売先」	「販売先」	「仕入先」	「仕入先」	「販売先」

図1. 有向グラフ



して列挙する企業数に上限があるためである。つまり、企業1にとって企業3は大事な仕入先であり、仕入先の上位24社に入る。しかし、企業3にとっては企業1は数ある販売先のうちのひとつに過ぎず、上位24社に入らない⁴⁾。

このように、本稿のデータセットでは、企業間の関係は双方向ではなく、関係の向きが意味をもつ。図1はこれを明示したものである。図1では、矢印の向きは、物流の向き(仕入先から販売先)ではなく、どの企業がどの企業を関係先として指名しているかを示している。このように関係に向きのあるネットワークは有向グラフとよばれている。なお、表1と図1は物流関係のみに注目しているが、同様の図は資本関係についても描くことができる。

有向グラフでは、関係数の数え方は2種類ある。第1は、ある企業が関係先として列挙した企業の数进行するという方法である。これは図1で言えば、ある企業から出て行く矢の数を数えることに相当するため「出次数(outdegree)」とよばれている。第2の数え方は、ある企業が他企業から関係先として指名された数を数えるという方法である。これは図1で言えば、ある企業に入ってくる矢の数を数えることに相当するため「入次数(indegree)」とよばれている⁵⁾。

本稿で用いるデータセットでは、各企業が関係先として列挙できる企業数には上限があるため、出次数の分布は切断されたものとなっている。一方、自分を除く全企業から指名されることがあり得るという意味で入次数には上限がなく真の分布に近いと考えられる⁶⁾。以下本稿では、入次数によって関係数を数えることとする。

3.2 スケールフリーネットワーク

次数(特に入次数)はどのような性質をもつだろうか。次節以降ではこの点について分析を行うが、その前に分析に最小限必要な用語を説明しておこう。

資本関係にしる物流関係にしる、特定の企業が多く次数をもっている。例えば、トヨタの販売先や仕入先企業を列挙するとすれば、その数は零細企業の比ではない。この意味で次数は少数の企業に偏在している。このように次数の偏在しているネットワークは企業間のネットワークだけではない。先述のインターネットの例は有名な例であるが、それ以外にも、学術論文の引用ネットワークでは、多くの論文に引用される少数の論文がある一方でほとんど引用されない無数の論文が存在するという偏在が指摘されている。このように次数の偏在しているネットワークはスケールフリーネットワークとよばれている。スケールフリーネットワークでは、各ノード(企業)に出入りするリンク数はべき分布に従うことが知られている。いま次数を x と表記することにして、あるノードの次数が x となる頻度(確率)が $x^{-\gamma}(\gamma>1)$ で与えられているとする。これがべき分布の定義であり、その名前はべき乗に由来している。

次数の確率がこのようにべき分布で与えられるとき次数は次の2つの性質をもつ。第1に、密度関数がべき乗であるが故に次数 x が非常に大きな値をとる確率は、密度関数が指数関数で与えられる指数分布などと比べて大きい。つまり、極端に大きな次数をとるノード(企業)が出現する確率が相対的に大きく、次数の分布を描くと裾が長い。第2に、次数 x と次数 x' の相対的な起こりやすさは x と x' の比にのみ依存し、 x と x' の絶対的な値には依存しない。つまりスケールフリーである。次数 x と次数 x' の相対的な起こりやすさは $x^{-\gamma}/x'^{-\gamma}=(x/x')^{-\gamma}$ であり、 x と x' の比率のみに依存しそれぞれの絶対水準には依存していない。

次数がべき分布に従うのはなぜだろうか。確率モデルで解釈する立場からは次のように説明できる⁷⁾。いま0期から始まって每期1社ずつ企業が生まれるとする。第 t 期における企業数

は t 社である。また、 t 期において j の次数をもつ企業の社数を $Z_j(t)$ と表記する。

t 期において新たに誕生した 1 社の企業は既存の企業 1 社と関係を結ぶとする。どの企業と関係を結ぶかについては、ランダムに企業を選ぶ場合と、各企業が既に持っている次数に比例して選ぶ場合があるとする。前者の確率は α 、後者の確率は $1-\alpha$ とする。後者の場合には、既に次数を多くもっている企業が選ばれることになり、持てるものがさらに富むメカニズムである。

この設定の下で $Z_j(t)$ が増加する確率は

$$\alpha Z_{j-1}(t)/t + (1-\alpha)(j-1)Z_{j-1}(t)/t$$

である。ここで第 1 項はランダムに企業が選ばれる場合、第 2 項は次数に比例して選ばれる場合に相当する。また、 $Z_j(t)$ が減少する確率は

$$\alpha Z_j(t)/t + (1-\alpha)jZ_j(t)/t$$

である。この 2 つの式から $Z_j(t)$ の遷移式として

$$\frac{dZ_j(t)}{dt} = \frac{\alpha(Z_{j-1}(t) - Z_j(t)) + (1-\alpha)((j-1)Z_{j-1}(t) - jZ_j(t))}{t} \quad (1)$$

が得られる。さらに、定常状態が存在すると仮定し、そこでは次数 j の企業が全体に占める割合は d_j とする。つまり、 d_j は定常状態における次数 j の頻度であり、 $Z_j(t) = d_j \cdot t$ である。これを (1) 式に代入すると d_j の満たすべき式として

$$d_j(1 + \alpha + j(1-\alpha)) = d_{j-1}(\alpha + (j-1)(1-\alpha))$$

$$\frac{d_j}{d_{j-1}} = 1 - \frac{2-\alpha}{1+\alpha+j(1-\alpha)} \sim 1 - \left(\frac{2-\alpha}{1-\alpha}\right)\left(\frac{1}{j}\right) \quad (2)$$

が成立する。この式の $\sim$ は、 j が大きくなるにつれて右辺と左辺の比が 1 に収束するという意味である(つまり、 j が大きいときに近似的に成立する)。いま

$$d_j \sim d_j^{\frac{2-\alpha}{1-\alpha}} \quad (3)$$

とすると、このときには

$$\frac{d_j}{d_{j-1}} \sim \left(\frac{j-1}{j}\right)^{\frac{2-\alpha}{1-\alpha}} \sim 1 - \left(\frac{2-\alpha}{1-\alpha}\right)\left(\frac{1}{j}\right)$$

が成立することから、(3)式が定常状態におけ

る次数 j の頻度である。つまり、ある次数の出現する確率が次数のべき乗となっており、次数がべき分布に従っている。 j の肩に乗っている値 $\left(\frac{2-\alpha}{1-\alpha}\right)$ はべき指数とよばれている。(3)式からわかるように、 α の値が 0 のときにはべき指数は 2 である。 α が大きくなるにつれてべき指数は単調に上昇し α が 1 に十分近いときにはべき指数は大きな値をとる。つまり、新たな企業が既に多くの次数をもつ企業と関係をもつというメカニズムが強ければ強いほどべき指数は小さく、定常状態における次数の偏在が顕著である。

次節以降では次数がべき分布に従うか否かを検証する。すなわち、次数を X として

$$\Pr(X = x) \sim \xi c/x^{\xi+1} \quad (4)$$

または、その累積密度関数である

$$\Pr(X \geq x) \sim c/x^{\xi} \quad (5)$$

が成立するか否かを調べていく⁸⁾。 c は任意の定数であり、(4)式と(5)式の ξ がべき指数である。べき指数の定義には、密度関数(PDF)の x の肩に乗っている数字((4)式の $\xi+1$)と累積密度関数(CDF)の x の肩に乗っている数字((5)式の ξ)の 2 種類がある。以下本稿では後者の定義を用いる。

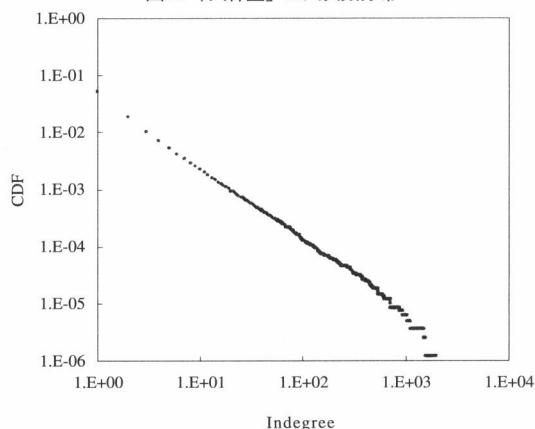
4. 分析結果

4.1 資本関係

図 2 は資本関係について入次数の累積密度関数(CDF)を示したものである。横軸は入次数、縦軸は累積密度であり、両軸とも対数表示されている。例えば、横軸の入次数 100 に対応する縦軸の数値はおよそ 0.0001 となっているが、これは 100 以上の入次数をもつ企業数が全体の 0.01% であることを示している。

図 2 からは次のことがわかる。第 1 に、入次数 1 以上に対応する CDF の値は 1 を大きく下回っており、入次数 0 の企業が多いことを示している。サンプル企業 826169 社中で入次数 0 の企業は 781912 社であり、94.6% を占めている。資本関係がない、あるいは、あるとしても「大株主」と認知されるほどの関係を持たない企業が大多数であることを示している。

図2.「大株主」の入次数分布



第2に, CDFは直線上にのっている。(5)式からわかるように, べき分布に従っているときには, CDFの対数値と x の対数値は線形関係にある。したがって, CDFが直線に乗っているということは, 次数がべき分布に従っていることを示しており, さらに, 企業間の資本関係がスケールフリーネットワークであることを示している。直線の傾きを調べるために

$$\log \Pr(X_i \geq x_i) = \text{const.} - \xi \log x_i + \nu \quad (6)$$

をOLSで推計する。ここで x_i は, 入次数の多い順に企業を並べたときに i 番目に位置する企業の入次数を表し,

$$x_1 > x_2 \geq \dots \geq x_n$$

を満たす。この回帰式は“log rank log size regression”とよばれるものであり, ξ の推計値がべき指数である⁹⁾。 ξ の漸近標準誤差は標本数を n として $\hat{\xi}(n/2)^{-1/2}$ で計算できる(Gabaix and Ibragimov(2006))。

$x=1$ に対応する点と $x=0$ に対応する点が不連続であることからわかるように, x のとり得る全ての値についてCDFが直線上にのっ

ているわけではなく, x の分布の裾の形状がべき分布に近いということである。したがって, べき指数の推計に際しては, 予め分布の裾を定義する必要がある。これは, (6)式の x_n の値をどのように選ぶかということに相当する。べき指数の推計に際しては, x の上位 $p\%$ を抽出し, そのサンプルで(6)式を推計するという方法がしばしば用いられる。ここではこの方法に従って, 入次数の上位5%の企業を抽出した場合と上位1%の企業を抽出した場合についてべき指数を推計する。

表2に示した推計結果をみると, 上位5%を抽出した場合(OBS=44257), べき指数は1.414であり, 標準誤差は0.009である。一方, 上位1%を抽出した場合は1.420である(標準誤差0.009)。(3)式で, 理論的にはPDFのべき指数が2以上(CDFのべき指数が1以上)であることを見たが, ここの結果はこれと整合的である。

既存研究の推計結果と比較すると, 日本の上場企業の株主関係を分析したSouma *et al.* (2006)では2003年時点で0.86との推計値を得ており, それと比べると本稿の推計値は大きく, 次数の偏在の度合いが低いという結果になっている。本稿で用いているデータセットは, 上場企業だけでなく非上場の中小企業もカバーしており, これがSouma *et al.*(2006)との大きな違いである。本稿の結果は, 資本関係が特定の企業に集中するという性質が, 中小企業では上場企業ほどに顕著でないことを示唆している。一方, 海外の上場企業を対象とした研究ではべき指数が1を上回る結果が多い。例えば, ミラノ証券取引所の上場企業では1.97, NYSEでは1.37, NASDAQでは1.22であり(Garlaschelli *et al.*(2005)), 本稿の推計値と近い値となっている。

表2. べき指数の推計(OLS)

	上位5%の企業				上位1%の企業			
	「大株主」	「仕入」	「販売」	全関係	「大株主」	「仕入」	「販売」	全関係
傾き(べき指数)	1.414	1.472	1.293	1.325	1.420	1.535	1.308	1.366
標準誤差	(0.009)	(0.009)	(0.008)	(0.009)	(0.009)	(0.023)	(0.020)	(0.021)
切片	-1.005	0.457	-0.079	0.437	-0.988	0.640	-0.036	0.557
OBS	44257	44137	46001	42066	44257	8279	8295	8290

表3. べき指数の推計(ヒル推計量)

	上位5%の企業				上位1%の企業			
	「大株主」	「仕入」	「販売」	全関係	「大株主」	「仕入」	「販売」	全関係
べき指数	1.773	2.036	2.375	2.344	1.089	2.594	0.992	2.719
標準誤差	(0.008)	(0.009)	(0.011)	(0.011)	(0.005)	(0.028)	(0.010)	(0.029)
OBS	44257	44137	46001	42066	44257	8279	8295	8290

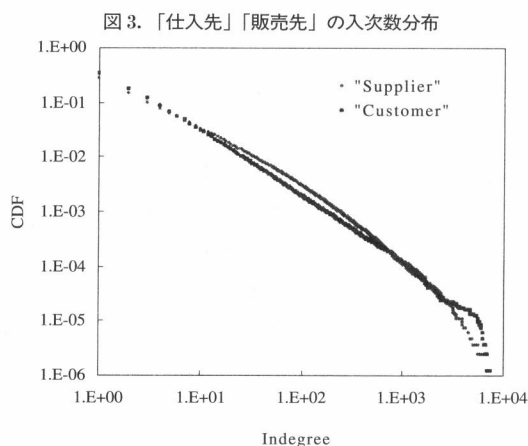
べき指数を推計するもうひとつの方法としてはヒル推計量(Hill's estimator)がある。これは

$$(n-1) / \sum_{i=1}^{n-1} [\log x_i - \log x_n]$$

で定義される推計量である。仮に次数が裾だけでなく全域でべき分布に従っているとすれば(つまり $\Pr(X \geq x) = c/x^k$) であるとすればヒル推計量は k の最尤推定量である。べき分布に従うのが裾に限られる場合でも同じ推計量が有益と考えられている(例えば Gavaix and Ioannides(2004)を参照)。表3の推計結果をみると、上位5%の企業を抽出した場合は1.773で、OLSによる推計値よりも高めである。ただし、上位1%の企業を抽出した場合はOLSの推計値を下回っており(1.089)、分布の裾をどのように定義するかによって結果が不安定である¹⁰⁾。

4.2 物流関係

次に物流の面から企業間関係を見てみよう。図3は「仕入先」と「販売先」のそれぞれについて入次数の分布を示したものである。図の見方は図2と同じであり、横軸・縦軸とも対数表

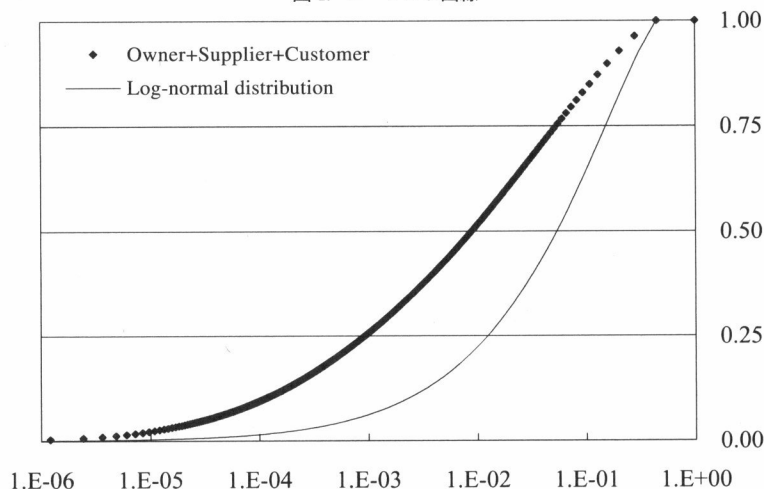


示である。

図2に示した資本関係と比較すると、入次数ゼロの企業数が少ないが(入次数ゼロが全企業に占める割合は「仕入先」で72.4%、「販売先」で67.6%)、その点を除けば、CDFは同じく直線であり、分布の裾がべき分布に従っていることを示している。べき指数をlog rank log size regressionにより推計すると、「仕入先」では、上位5%を抽出した場合に1.472(標準誤差0.009)、上位1%を抽出した場合は1.535である(標準誤差0.023)である。どちらも資本関係のべき指数を上回っているが、その差は僅かである。一方、「販売先」についてべき指数を推計すると、上位5%を抽出した場合で1.293(標準誤差0.008)、上位1%を抽出した場合で1.308(標準誤差0.020)であり、「仕入先」や「資本関係先」に比べ、べき指数が小さいことがわかる。各企業が主要な販売先と認識している先が一部企業に偏っており、そのため集中度が高くなっていると解釈できる¹¹⁾。

図4では、関係数が一部の企業に集中している様子をより詳しく見るために、間係数でランキングしたときに上位に位置する企業が関係数全体のどれくらいを占有しているかを示している。富の不平等などを計測する際に用いられるローレンツ曲線と同じものである。図4では「資本」「仕入」「販売」の3種類の企業間関係を区別せず、全てのデータをプールしている。この図からわかるように、入次数の多い企業の上位1%は全入次数の約50%を占めている。図では比較のために入次数が対数正規分布に従うとした場合の計数も示してあるが、その場合には上位1%の企業が占めるのは全入次数の約20%に過ぎない。入次数の分布の裾が非常に長いことを示している。

図4. ローレンツ曲線



4.3 サブサンプル分析

4.1節と4.2節の分析は全サンプルを対象としたものである。本節では、(1)入次数の多寡、(2)企業の年齢を切り口としてサブサンプルを作成しその性質をみていく。

ハブ企業 企業間関係数がべき分布に従い、ごく少数の企業が多くの入次数を占有していることは、「ハブ企業」の存在を示唆している。それではこれらハブ企業間の関係はどのようになっているのだろうか。ハブ企業だけから成るサブサンプルを作成し、ハブ企業間の関係を調べてみよう。

図5では、入次数の多い企業をハブ企業と定義し、それらハブ企業間での入次数の分布を調

べている。具体的には、入次数が N 以上の企業をハブ企業とし、 N の値を50, 100, 200, 500, 1000と変化させる。ハブ企業だけを取り出すと十分な社数が確保できないので、ここでは「資本」「仕入」「販売」の3種類の関係を区別せず、全てのデータをプールする。ちなみに、プールした場合、 N が50以上の企業は10139社、100以上は4619社、200以上は1984社、500以上は635社、1000以上は265社である。図5でみているのは、例えば $N \geq 10000$ の場合、抽出された265社の間の入次数であり、それ以外の企業との関係は除外している。

図5をみると、ハブ企業間での入次数の分布はほぼ直線になっており、分布の裾がべき分布に従っていることを示している。これは、ハブ企

図5. ハブ企業間の関係

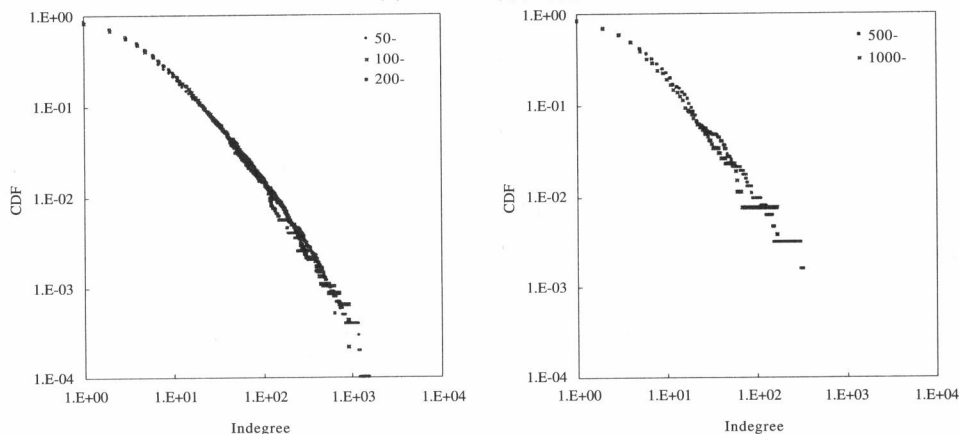
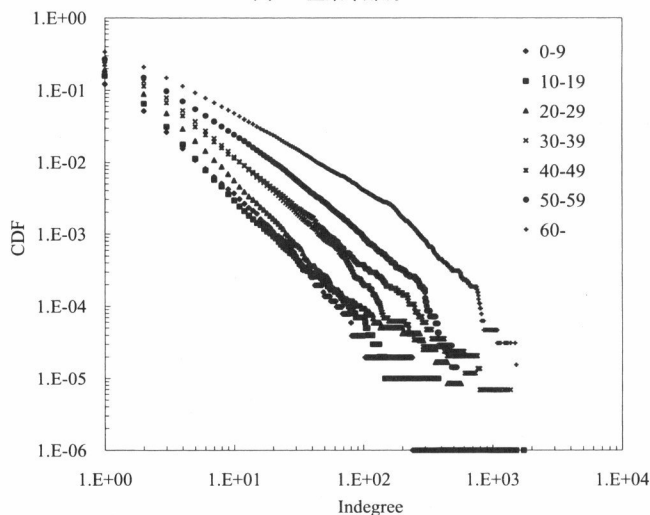


図6. 企業年齢別



業間でも、入次数の多寡に大きなばらつきがあるということであり、ハブ企業の中のハブ企業、つまり超ハブ企業が存在することを示唆している。また、べき指数を log rank log size regression により推計すると(それぞれのサンプルの上位5%を抽出)、 $N \geq 50$ のときには1.728, $N \geq 100$ で1.537, $N \geq 200$ で1.452, $N \geq 500$ で1.379, $N \geq 1000$ で1.073と単調に減少している。これは、例えば、入次数50以上の企業群と入次数1000以上の企業群では後者の方が集中度が高いことを意味している。入次数の大きい企業だけを取り出せばそれら企業間では入次数の格差はなく「平等」になっているとも考えられるが、ここでの結果はそうした認識が正しくなく、大きな企業にいけばいくほど格差が拡大することを示している¹²⁾。

図5の結果は、一般企業の上位層がハブ企業を構成し、それらハブ企業の上位層が超ハブ企業を構成し、さらにその上位層がさらに高次のハブ企業を構成し…というように、企業間関係が重層的な構造をもっていることを示唆している。

企業年齢別 本稿のデータセットは2003年における企業間関係を描写したものであり、クロスセクションデータである。データが一時点し

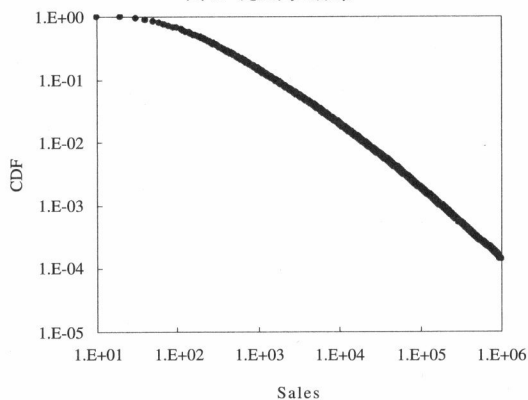
か存在しないため、企業間の関係が時間とともにどのように変化するかを見ることはできない。しかし各企業が法人格を取得した時点から企業年齢がわかるので、企業年齢ごとに企業間関係を見ることができる。若年企業がもっている企業間関係と老年企業の企業間関係を比べることにより、企業の一生の中で企業間関係がどのように変遷するかを見ることができる。

図6は企業を年齢に応じて7つのグループに分け、それぞれのグループ内での入次数のCDFを見たものである。例えば、「0-9歳」というのは、生まれたばかりの企業が同じ年代の企業とどのような関係をもっているかを表している。0-9歳の企業が上の世代の企業と関係をもつことも当然あるがここではそれは無視している。つまり、ここで見ているのは、人間にたとえれば、幼稚園→小学校→中学校というように成長するにつれて自分と同世代の人間とのかかわりがどのように変化していくかである。

図6から以下のことを読み取ることができる。第1に、企業は加齢とともに他企業との関係を充実させる傾向がある。入次数0の企業の割合をみると、0-9歳では87.7%, 10-19歳では84.1%, 20-29歳では80.7%, 30-39歳では73.9%, 40-49歳では76.9%, 50-59歳では73.1%, 60歳以上では66.1%となっており、年齢とともに低下する傾向がある。ただし、入次数0の割合は加齢とともに単調に下がるわけではなく、40歳まで下がるがその後は目立った低下は見られない。他企業との関係性の構築は40歳ごろまでにほぼ完了すると解釈できる。

第2に、年齢ごとのべき指数を log rank log size regression により推計すると(それぞれのサンプルの上位5%を抽出)、0-9歳では1.760, 10-19歳では1.717, 20-29歳では1.412, 30-39歳では1.541, 40-49歳では1.648, 50-59歳では1.862, 60歳以上では1.582となっており、0-29歳の若年期にはべき指数が単調に低下する(集中度が高まる)もののそれ以降は低下の傾向は

図7. 売上高の分布



見られない。この結果も、企業間関係の構築が若年期に集中してなされ、ある程度の年齢になると完了することを示唆している。

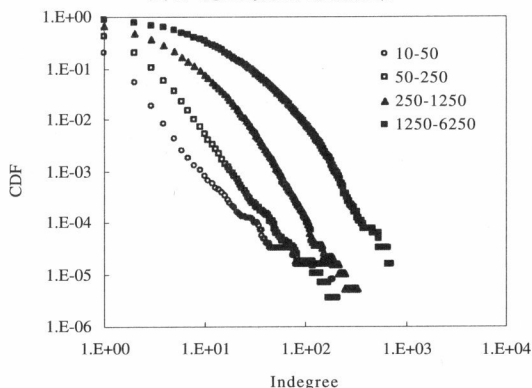
4.4 企業規模と企業間関係

ここまでの分析では、企業間関係を入次数で測り、入次数という一変数を分析対象としてきた。本節では分析対象を企業間関係と企業規模の二変数に拡張し、両者の関係を見ていくことにしよう。

まず企業規模の性質を調べることから始めよう。図7では、企業規模を表す変数として売上高をとり、そのCDFを示している。横軸は売上高(金額)であり、縦軸は累積密度である。これまで見てきたCDFと同じく両軸とも対数表示である。図8からわかるように売上高の分布の裾は直線であり、べき分布に従っている¹³⁾。べき指数をlog rank log size regressionにより推計すると(サンプルの上位5%を抽出)、1.036であり、関係数のべき指数より小さい¹⁴⁾。つまり、一部の企業が売上高の多くの部分を占めるという意味での集中度は関係数よりも高い。

次に、企業間関係数と企業規模の関係を見てみよう。図8では、企業を売上高に応じて分類し、売上高で条件づけた入次数の分布を示してある。具体的には、企業を、売上高10-50百万円、50-250百万円、250-1250百万円、1250-6250百万円の4グループに分割し、グループごとの入次数分布を示している。ここでは、各グループ内の入次数分布を見ているのではなく、

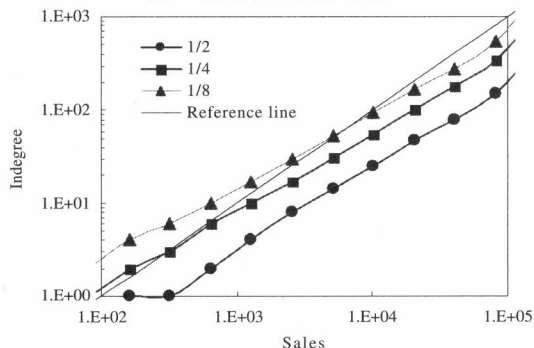
図8. 売上高別の入次数分布



各グループに属する企業が全サンプルの企業との間にもつ関係数を見ている。まず入次数ゼロの企業の割合をみると、売上高10-50百万円の企業群では80.0%であるが、50-250百万円の企業群では58.6%、250-1250百万円の企業群では30.9%、1250-6250百万円の企業群では12.1%と、企業規模が大きくなるにつれて低下していることがわかる。企業規模が大きくなるにつれて入次数も大きくなることを示している。次に各企業群のべき指数をlog rank log size regressionにより推計すると(サンプルの上位5%を抽出)、売上高10-50百万円の企業群で1.485、50-250百万円の企業群で2.387、250-1250百万円の企業群で3.064、1250-6250百万円の企業群で3.037と、企業規模が大きくなるにつれてべき指数が大きくなる(集中度が低下する)傾向がある。

図9では、売上高と入次数の関係を詳しく調べるために、企業を売上高に応じて10分割し、それぞれのグループごとに、入次数のCDFが1/2になる入次数の値、1/4になる入次数の値、1/8になる入次数の値を計算し、プロットしている。横軸は売上高、縦軸は入次数である。まず、入次数のCDFが1/2になる点を見ると(この点は入次数のメディアンに相当する)、売上高の小さい企業群では入次数のメディアンはゼロであるが売上高の大きな企業群ではそれにつれて入次数も大きくなっている。しかもメディアンが正の企業群だけに注目すると、図中に示した参照線とほぼ平行になっている。この参照

図9. 売上高と入次数の関係



線は売上高と入次数が比例関係にある場合のものである。ここでの結果は、メディアンでみると、売上高と入次数は比例関係にあることを示している。

次に、累積密度関数が1/4または1/8になる点、つまり各企業群について入次数の多い順で上から1/4または1/8に位置する企業の入次数をみると、メディアンと同様、売上高とともに増加してはいるものの、その傾きは参照線よりも小さく、比例関係に届いていないことがわかる。つまり、入次数は常に売上高と比例関係にあるわけではなく、入次数の大きい企業については売上高が増えてもそれほどには入次数が増えていない。

この結果は次のように解釈できる。企業と銀行の関係に関する研究によれば、企業が何行の銀行と取引するかは2つの要因のバランスで決まる(例えば Detragiache(2000)を参照)。1つの銀行と取引すると、取引銀行が大きな損失を被り資本が毀損した場合にその銀行から融資を受けられなくなってしまう。しかもその際に別な銀行への乗り換えは容易でない。このような状況に陥るのを回避するには複数の銀行と取引を行い、そのうちのひとつに問題が生じても自らの資金繰りに影響が及ばないようにしておくことが望ましい。また、いわゆるホールドアップ問題を軽くするためにも複数の銀行と取引することが望ましい。しかし、それでは企業は無数の銀行と取引するかという決してそうではない。多くの銀行と取引を行おうとすると、企業はそれぞれの銀行に対して自らのプロジェク

トや業績について説明しなければならず、そのために必要な書類作成の手間も嵩む。つまり、銀行との間で関係を構築し維持していくにはコストがかかるのであり、このコストと銀行が破綻したときの不利益とを比較秤量して最適取引銀行数を決定する¹⁵⁾。

銀行と企業の間と同様に、企業と企業の間関係を構築・維持するにはコストがかかり、企業はそのコストを節約する誘因をもつと考えられる。既に多くの関係を有している企業にとって、さらに多くの関係をもつことから得られる利益は小さく、コストに見合わないため、企業規模が大きくなって関係数はそれほどには増やさないのであろう。

5. おわりに

本稿では、企業間関係の基本的な性質を調べるとともに、企業間関係と企業規模がどのように関係しているかを検討した。具体的には、日本の法人企業約80万社(全法人企業の約3分の1)について他の法人企業との資本・取引関係を記述するユニークなデータセットを用いて分析を行い、以下のファインディングを得た。

第1に、企業の資本・取引関係数の分布はロングテールを持ち、その分布はべき分布に近い。資本・取引関係数の多い企業の上位1%がもつ関係数は全関係数の約50%であり、偏在が顕著である。

第2に、取引関係数が多い企業だけを取り出し、それらの企業間の関係をみると、再びべき分布が観察される。つまり取引関係数の多いハブ企業同士の関係も偏在しており、超ハブ企業に集中している。

第3に、企業規模が大きいほどその企業の関係数は多く、全体としては両者の関係は線形である。ただし既に多くの関係をもつ企業に注目すると、企業規模が拡大するほどには関係数を増やしていない。これは資本・取引関係を維持するためのコストを企業が節約しているためと解釈できる。

(富士通総研・一橋大学経済研究所)

注

* 本稿の作成に際しては岩村充氏、植杉威一郎氏との議論が有益であった。また、独立行政法人経済産業研究所(RIETI)からは本稿で使用するデータセットの提供を受けた。記して感謝したい。

1) 持続的な関係の有り難さがわかるのはそれが損なわれたときである。銀行の資本が毀損するとその銀行と取引をしてきた企業が資金繰り難に陥るというクレジットクランチはその例である。

2) ネットワーク分析の経済現象への応用や経済理論との関係については Jackson(2005, 2007)を参照。

3) 企業数が大きいということは単にカバレッジが高いという利点だけでなく、様々なサブグループで構成されるネットワークの性質を見ることができるといふ利点も持つ。例えばある特定地域、ある特定の企業年齢、ある特定の業種の企業群を抽出し、それらの企業間で構成されるネットワークの性質を見ることが可能である。

4) あるいは、企業3にとっての重要な販売先が本稿で用いるデータセットの対象企業でないという可能性もある。

5) 表1(または図1)の例について出次数を数え、と、「仕入先」では企業1が2、企業2が1、企業3が1、企業4は0である。「販売先」では企業1が1、企業2が1、企業3が0、企業4が1である。一方、入次数を数え、と、「仕入先」では企業1が0、企業2が1、企業3が1、企業4が2である。「販売先」では企業1が1、企業2が1、企業3が0、企業4が1である。

6) ただし、本稿で用いるデータセットはカバレッジが高いとは言っても全企業をカバーしているわけではない。本稿で用いるデータセットに含まれる企業がサンプル外の企業から指名を受けているとしても、それはカウントされない。その意味では入次数にもバイアスが存在する。

7) 以下のアイディアは Barabasi and Albert (1999)によるものであり、preferential attachment とよばれている。このアイディアは Simon(1955)まで遡る。なお、ここでの説明は Mitzenmacher(2004)による。

8) (5)式は(4)式を累積させたものであるが通常の累積密度関数とは反対に累積させているため「逆累積密度関数(“counter-cumulative” distribution function)」とよばれている。

9) (6)式の左辺は $\log i/n$ であり、企業のランキングを表していると解釈できる。

10) この結果は本稿で使用するデータに固有のものではない。Gabaix and Ioannides(2004)などで指摘されているように、ヒル推計量の実務的な問題は、べき分布がどこから始まるのか分からないことであり、しかも推計値が選択した閾値に依存することである。

11) ただし、「販売先」のべき指数が低いという結果はべき指数の推計手法に依存しており必ずしも頑健でない。表3に示したヒル推計量をみると、上位5%を抽出した場合、「販売先」のべき指数は2.375であり、「大株主先」や「仕入先」より大きく、集中度が低

いという結果になっている。なお、「販売先」のヒル推計量は、上位5%を抽出した場合と上位1%を抽出した場合とで大きく異なっており、推計結果が不安定である。

12) この結果は、大企業だけを取り出してべき指数を計測した場合と、中小零細企業も含むサンプルでべき指数を計測した場合で、前者の方がべき指数が小さくなる(集中度が高まる)ことを示唆している。実際、4.1節で紹介したように、上場企業を対象とした既存研究で得られているべき指数は本稿の全サンプルを用いて得られたものより小さくなっている。

13) 企業規模の分布の裾がべき分布であることはこれまで多くの研究によって指摘されている。例えば Axtell(2001), Fujiwara *et al.*(2004)を参照。

14) べき指数が1のべき分布に従うことはジフ則(Zipf's Law)とよばれている。べき指数が1に近いというところの結果は、売上高の分布がジフ則に従うことを示している。同様の結果は Axtell(2001)や Fujiwara *et al.*(2004)など先行研究でも確認されている。なお、ジフ則は企業の規模だけでなく都市の規模(さらには国の規模)についても成立することが知られている(Gabaix and Ioannides(2004)を参照)。

15) 日本の企業では1つの銀行とだけ取引するのは少数派である。一方、10行以上と取引する企業も稀である。多くの企業が3-5行の銀行と取引している。

参考文献

- Axtell, Robert (2001) "Zipf Distribution of U. S. Firm Sizes," *Science*, Vol. 293, No. 5536, pp. 1818-1820.
- Barabasi, A. and R. Albert (1999) "Emergence of Scaling in Random Networks," *Science*, Vol. 286, No. 5439, pp. 509-512.
- Caldarelli G., Garlaschelli D. and Battiston S. (2006) "The Skelton of the Shareholders Networks," H. Takayasu (ed.) *Practical Fruits of Econophysics*, Tokyo: Springer-Verlag, pp. 297-300.
- Detragiache, E., P. G. Garella, and L. Guiso (2000) "Multiple versus Single Banking Relationships: Theory and Evidence," *Journal of Finance*, Vol. 55, No. 3, pp. 1133-1161.
- Fujiwara, Yoshi, Corrado Di Guilmi, Hideki Aoyama, Mauro Gallegati, and Wataru Souma (2004) "Do Pareto-Zipf and Gibrat Laws Hold True? An Analysis with European Firms," *Physica A*, Vol. 335, Issues 1-2, pp. 197-216.
- Gabaix, Xavier (2007) "Power Laws," in *The New Palgrave Dictionary of Economics*, 2nd Edition.
- Gabaix, X., P. Gopikrishnan, V. Plerou, and H. E. Stanley (2006) "Institutional Investors and Stock Market Volatility," *Quarterly Journal of Economics*, Vol. 21, No. 2, pp. 461-504.
- Gabaix, Xavier and Rustam Ibragimov (2006) "Rank-1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents," MIT.
- Gabaix, Xavier, and Yannis M. Ioannides (2004)

- "The Evolution of City Size Distributions," in J. Vernon Henderson and Jacques-Francois Thisse, (eds.), *Handbook of Regional and Urban Economics*, Volume IV, Amsterdam: Elsevier North-Holland, pp. 2341-2378.
- Garlaschelli D., Battiston S., Castri M., Servedio V. D. P. and Caldarelli G. (2005) "The Scale Free Nature of Market Investment Network," *Physica A*, Vol. 350, Issues 2-4, pp. 491-499.
- Jackson, Matthew O. (2007) "Network Formation," in *The New Palgrave Dictionary of Economics and the Law*, MacMillan Press, forthcoming.
- Jackson, Matthew O. (2005) "The Economics of Social Networks," Lecture prepared for the 9-th World Congress of the Econometric Society, August 2005, forthcoming in the *Proceedings of the 9-th World Congress of the Econometric Society*, edited by Richard Blundell, Whitney Newey, and Torsten Persson, Cambridge University Press.
- Mitzenmacher, Michael (2004) "A Brief History of Generative Models for Power Law and Lognormal Distributions," *Internet Mathematics*, Vol. 1, No. 2, pp. 226-251.
- Saito, Yukiko, Tsutomu Watanabe, and Mitsuru Iwamura (2007) "Do Larger Firms Have More Interfirm Relationships?" Forthcoming in *Physica A*.
- Simon, Herbert (1955) "On a Class of Skew Distribution Functions," *Biometrika*, Vol. 42, No. 3-4, pp. 425-440.
- Souma, W., Y. Fujiwara, and H. Aoyama (2005) "Shareholding Networks in Japan," <http://arxiv.org/abs/physics/0503177>.
- Souma, W., Y. Fujiwara, and H. Aoyama (2006) "Change of Ownership Networks in Japan," in H. Takayasu (ed.), *Practical Fruits of Econophysics*, Tokyo: Springer-Verlag, pp. 307-311.