Chapter 5 The distribution of the rate of profit

In this chapter we test the ability of the techniques developed in Chapter Four to identify the distributional laws followed by different measures of the rate of return. It is thus a test of the hypotheses of Gibrat (1931) and of Farjoun and Machover (1983).

Since our method involves inspecting clouds of points in *L*-skewness, *L*-kurtosis space which represent the profit rate in different years, it is simultaneously an exploration of how the precise forms of the distributions vary over the business cycle (but a necessarily limited one, in view of the small number of years covered by our data).

Gibrat's hypothesis of a log normal distribution of the profit rate is simply one of many illustrative examples intended to demonstrate the wide applicability of his Law of Proportionate Effect (LPE).[42] As such, it is not backed by arguments specific to the rate of profit; indeed, we saw in Chapter Two that Gibrat declines to define profit himself (1931:180), although his discussion draws on previous studies of dividends. Moreover, it appears that no one has followed up his suggestion about profit rates.

Likewise, there has been limited work on Farjoun and Machover's notion that the profit rate is a random variable, even though their small following is extremely respectful of their work.[43]

---

[42] Other economic quantities for which Gibrat provides illustrations include not only the celebrated study of firm size, but also income, wealth, and size of towns. Note that despite the indelible association of this law with Gibrat in the economics literature, similar processes were described by Kapteyn (1903) in relation to biological statistics, while the log normal itself was first described by McAlister (1879).

[43] A number of workers cite Farjoun and Machover's approach with respect, but do not test their hypotheses directly (Cockshott and Cottrell, 1994, 1998, Cockshott et al., 1995, Julius, 2005, Puty, n.d., Sheppard and Barnes, 1986, 1990, Webber, Michael, 1987, Webber, Michael J. and Rigby, 1986, Webber, Michael J. and Rigby, 1996). A partial exception is Zachariah (2005); but although he estimates empirical probability density functions he does so for *industries*, not firms, and does not attempt to identify any functional form for them, still less estimate parameters. Wright's work (2005) cited above is part of a project to build an agent-based simulation model of a capitalist economy in which the outcomes of transactions are random variables and the limiting distributions of their outcomes (including the profit rate) conform to accepted stylised facts.

However, Farjoun and Machover's stipulations about the definition of the profit rate, and about the population to be considered, are much more specific than Gibrat's. Moreover they provide an elaborate theoretical justification for their hypothesis. Indeed, that hypothesis is offered as just one confirming instance of the validity of a wide-ranging methodological outlook. Since the key objective of the current chapter is to test the Farjoun and Machover hypothesis, we necessarily devote some space here to elaborating the outline given in previous chapters.

Since Gibrat's case for the log normal distribution of profit rates is at best lightly argued in his book, we can conveniently state the LPE here before continuing: suppose some attribute of the members of a population is described by a variable whose value, for each particular member of the population, changes from one period to another according to some rule which is independent of the initial value for each member, but subject to random variation between members. It can be shown that the limiting form of the variable's distribution, after *n* rounds of change as *n* grows large, is the log normal distribution. For example, if over the interval *t* the size of each of *i* firms varies by $ak_i + e_i$, where *a* is some constant, *k* represents the size of each firm and *e* is a random variable, then in the limit the firm size distribution becomes log normal.

Since his illustrations take firms as the unit of observation, one might conclude that a strict confirmation of Gibrat would entail finding a log normal distribution at the firm level. We think that this would be unduly restrictive, and will take the detection of a log normal distribution at either firm or capital level, for any profit rate measure, as constituting support for Gibrat (in this study we only investigate functional forms for the capital level distribution).

Farjoun and Machover (1983) was not widely reviewed, but the notices which did appear included some by very significant figures (Champernowne, 1985, Steedman, 1983). The most penetratingly critical account was also the most sympathetic (Webber, M.J., 1986); a further review was by Jalali-Naini (1986).

The structure of the chapter is as follows: section 1 reviews Farjoun and Machover's probabilistic political economy and consider the implications of possible outcomes of our tests of their profit rate hypothesis. In section 2 we present charts of the annual estimated *L*-moment ratios for each profit rate measure, and consider the properties of our technique of randomly-sized random samples (RS2) in respect of these; section 3 renews our investigation of the tails of both unweighted and size-weighted versions of our profit rate measures.

5.1 Farjoun and Machover's probabilistic political economy

In this section we review Farjoun and Machover's case for adapting the statistical outlook of ideal gas mechanics to Marxian economics, and consider the possible outcomes of a test of their hypothesis about the profit rate distribution.

In one respect we diverge from their account: where they see their approach as one of 'restoring' the insights of *Capital* Volume I, lost through an unfortunate detour in Volume III as a result of the needless assumption of uniform profit rates, we suggest a different interpretation of Marx, arguing that he adopts a fundamentally probabilistic standpoint from the very beginning of his work in political economy through to its mature expression in *Capital*.

Finally, we discuss some practical considerations in testing Farjoun and Machover's hypothesis of a gamma distribution for the rate of profit, and the implications of the various possible outcomes.

5.1.1 Statistical mechanics, the transformation problem, and the gamma distribution

Our account of the statistical mechanics outlook, and of its applicability to issues to do with the rate of profit, substantially follows Farjoun and Machover (1983). A modern industrial economy contains a very large number of agents: in the United Kingdom, for example, there are tens of thousands of capitalist enterprises employing tens of millions of workers, and the commodities produced are the subject of billions of transactions every week.

Because the U.K. is a competitive capitalist economy, it is intrinsic that all this activity is unplanned, in the sense of not being directed by a central authority.

Since at least the time of Adam Smith, it has been a commonplace that all these activities, although unplanned, are nonetheless co-ordinated, and that this is a fact requiring explanation. Traditionally this explanation has been provided by reasoning about so-called representative agents[44] in terms of how their interaction results in the formation of 'the' price of a commodity, 'the' rate of profit, and so on, where the definite article implies a single uniform value for the variable concerned.

Of course, everyone knows that in practice economic variables do not have uniform values. The price of identical tomatoes in a street market may vary from stall to stall, while there would be little reason for the existence of a stock market if all companies achieved the same rate of profit. But it is assumed that the forces of competition will tend to smooth out such differences, given enough time and an absence of external disturbances, until uniform values are achieved in equilibrium. Indeed, uniformity is taken to be part of the meaning of equilibrium.

This notion of equilibrium is what Farjoun and Machover contest. As they point out, in Marx's economics the forces that tend to bring about uniformity are opposed by other forces that tend to disrupt it. In the case of the rate of profit, movement of capital from one sector to another tends to equalise it, while the search for relative surplus value through technical innovation tends to differentiate it.[45] A firm which can reduce necessary labour will increase its profit rate, at least until such time as its competitors are able to copy the new techniques (or devise even better ones).

---

[44] The agents may be producers and consumers, or workers and capitalists, according to the theoretical tradition in question. But the fundamental method is widespread in all traditions.

[45] Relative surplus value is Marx's term for the extra surplus value gained when capital is able to shorten the time needed for labour to reproduce its wage (necessary labour) within a *fixed* working day; absolute surplus value can be increased by lengthening the total working day while necessary labour time stays constant.

Both these forces are results of capitalist competition, hence internal to the system since competition is essential to the notion of capitalism. An adequate concept of capitalist equilibrium might therefore be one which accords equal status to both equalising and differentiating forces.

Taking uniformity of profit rates as the sole definition of equilibrium in effect treats competition as a force internal to the economic system insofar as it tends to equalise the rate of profit, and as an external, perturbing, force when it has the opposite tendency. As Farjoun and Machover point out (1983:36), this is like regarding the force of gravity as a force internal to a pendulum's system as it swings to the right (and gravity tends to pull it back to the left), but as an external force as it swings to the left (and gravity pulls it to the right).

A concept of equilibrium might instead be consistent in treating competition as an inherent feature of capitalism by incorporating both equalisation and differentiation, hence in treating it as an internal force. It should allow a range of profit rates to exist at any given moment, and be dynamic, in the sense of allowing individual firms' profit rates to vary through time. The required concept would be an equilibrium one in the sense that the proportion of the total social capital that achieves any particular rate of profit is roughly constant through time.

Here is the point of application of (classical) statistical mechanics: in an ideal gas, the velocities of each of the individual particles which compose it are taken to be both widely differentiated and rapidly changing as a result of constant interaction between them, in the form of collisions. The velocities, hence speeds, of the particles determine the heat energy of the gas, which is the sum total of the kinetic energy. This macroscopic variable is stable (assuming no input from outside the system). The law of the conservation of energy stipulates that if one particle gains energy from a particular collision the other particle must lose it. In a gas at macroscopic equilibrium (that is, neither heating up nor cooling down) the particles are in a sense competing for a share of a fixed pool of energy.

There is a is clear analogy here with Marx's discussion in Volume III of *Capital*, where capitalists compete for profits in the form of a equal 'aliquot' share[46] in the total surplus value created in production. The overall rate of profit (compare: total heat energy) is fixed; the individual rates of profit are the result of individual competitive effort – in the form of both shifting capital from one industry to another, and investments in new technology – but the relative success of one firm is at the expense of the relative failure of another.

Farjoun and Machover note that the 'most chaotic' partition of kinetic energy among particles results in a gamma distribution (page 68), but that a formal proof would require a notion of entropy in the firm space.[47,48]

The point of the analogy, for Farjoun and Machover (1983), is that assuming uniform speeds for the particles will give the wrong value for the total kinetic energy, essentially because the energy of each particle is proportional to the square of its speed. Thus the total energy is given by $nk\,\mathbf{E}\,(V^2)$, where n is the number of particles, $k = 0.5$ mass of each particle, V is the velocity of each particle, and $\mathbf{E}\,(V^2)$ denotes the mean of the squared velocities; assuming equal speeds would suggest that $nk\,(\mathbf{E}V)^2$ would also compute the total energy. But in general these expressions do not have the same value: $\mathbf{E}\,(V^2) \quad (\mathbf{E}V)^2$ is the standard definition of the variance of V.

Statistical mechanics uses probabilistic assumptions to derive the equilibrium distributions (and hence the expected values) of position, speed and energy among the particles at a given moment (Farjoun and Machover 1983:55). These are *space* averages (in the sense of sample space). But one can also consider the distribution of a given particle's speed, *etc*., in *time*. To say that a system is in (dynamic) equilibrium is to say that, provided

46 That is, an equal share for each unit of capital they have advanced.

47 In fact Maxwell's distribution is a particular member of the ^{2} family, itself a special case of the gamma distribution (see Appendix).

48 The notion of entropy is first introduced in a note to page 26, where Farjoun and Machover draw attention to another, independent, speculation on the possibility of a probabilistic economics (Jaynes, 1991).

that one considers a sufficiently long time interval, the time average approaches the space average as a limit.[49] An implication is that the time averages of two distinct particles will be approximately equal (Farjoun and Machover 1983:49).

As discussed in Chapter One, Farjoun and Machover (1983) is an intervention in the debate on the transformation problem: how to find a way of reconciling values determined by labour time with the prices needed to equalise the rate of profit between firms employing different proportions of capital and labour. Their '*dis*solution' of the transformation problem is performed by following what might be called a binocular approach: to gain perspective by examining two different economic spaces (the firm space[50] and the market space) and show that each contains analogous variables which can be taken to be identical with high probability.

In the firm space, value and surplus value are created in production; in the market space, value and surplus value are not only realised in exchange, but also – because of the purely probabilistic relations between (money) price and (labour) value – transferred among those who organise its production.

First consider the firm space: after arguing as described above for the gamma distribution of the rate of profit random variable R, Farjoun and Machover define a further random variable Z, the 'rate of wage bill', which is simply the annual gross wage bill of the firm divided by its capital. They appeal to the same heuristic argument to establish that Z is also gamma distributed with (at least) the same beta parameter as R. The random variable R/Z will then correspond to Marx's notion of the rate of surplus value s/v, taken by him to be uniform. Farjoun and Machover appeal to a unique property of the gamma distribution to

[49] Note that the time span over which this so-called ergodic principle operates is long in relation to the micro events, but – in the physical systems considered here – is short by the standards of human observers.

[50] Or, as we prefer, the capital space.

show that if the random variable R/Z is single-valued then its only possible value is 1 (page 72).[51,52]

Now consider the market space, the space of all market exchanges in the same time period as used to think about the firm space. Farjoun and Machover define further random variables

W : the money wage paid per unit of labour-time

: 'specific price', or the money price paid per unit of labour content, where the monetary unit is *defined* to be that which makes the average unit wage $\mathrm{E}(W) = 1$, itself a ratio of two random variables, and .

: money paid in transaction i

: labour content of commodity traded in transaction i

Also, $(i) = \mathrm{V}(i) + \mathrm{S}(i)$, where

V : labour cost of commodity traded

S : profit realised on commodity traded

Farjoun and Machover show that $\mathrm{E}(\ ) = \mathrm{E}(W) + [\ \mathrm{S}(i)/\ \mathrm{V}(i)]$. By their choice of units $\mathrm{E}(W) = 1$, and they argue that $\mathrm{S}(i)/\ \mathrm{V}(i)$ can be identified with R/Z (claimed also to be approximately 1, recall).

---

[51] A point which Farjoun and Machover, surprisingly, do *not* mention, let alone stress, is that Marx always takes $s/v=1$; no doubt this makes the arithmetic of his examples simpler, but other values could give equal simplicity (0.5? 2? 100?), and perhaps Marx might have used one of these if he thought it more reasonable as a stylised fact.

[52] See Appendix for details.

Thus, by considering the division of the price realised in each commodity transaction into payments to workers, and profits for capitalists, a market space analogue of the (statistical) equality of R and Z can be identified.

Although the first - $S(i)/V(i)$ - represents the division between capital and labour of the price of commodities traded, while the second - R/Z - is the division of value created, at least some of it in periods before that in which commodities are exchanged, Farjoun and Machover argue that 'if the economy is at or near dynamic equilibrium, the two ratios must be extremely close to each other, because the ratio between total profits and total wages cannot change rapidly' (1983:118), and indeed has already been shown to be very stable in the region of 1.

> If a competitive market economy has a state of equilibrium, it must be a state in which a whole range of profit rates coexist; it must be a *dynamic* state, in the sense that the profit rate of each firm keeps changing all the time; it can only be a state of equilibrium in the sense that the *proportion* of capital (out of the total social capital) that yields any particular rate of profit remains approximately constant.
>
> (Farjoun and Machover 1983:36)

Similarly they work in the market space (page 135) to show that the average rate of profit for the whole economy can be identified (once again, in a probabilistic way) with $E(R)$.

5.1.2 Farjoun and Machover on Marx

In relation to Marx's political economy, Farjoun and Machover's project can be described as one of both restoration and reconstruction. On the one hand, they see themselves as restoring the essence of Marx's thought, as expressed in *Capital* Volume I, and on the other as reconstructing his theory so as to rescue him from the *impasse* into which they believe he led himself in Volume III – that is, the transformation problem.

Our position is that Marx's political economy is fundamentally probabilistic; we shall suggest an alternative reading to Farjoun and Machover's which reveals Marx to have anticipated their perspective, albeit without the modern technical apparatus.

As Farjoun and Machover point out, there are two versions of the uniform profit rate hypothesis, a hard one and a soft one.

They begin (pages 14-15) with two quotations from Marx:

> ' … there is no doubt, however, that in actual fact, ignoring inessential, accidental circumstances that cancel each other out, no … variation in the average rate of profit exists between different branches of industry, and it could not exist without abolishing the entire system of capitalist production.'
>
> (Marx, 1981: 252)

> ' … capital withdraws from a sphere with a low rate of profit and wends its way to others that yield higher profit. This constant migration, the distribution of capital between the different spheres according to where the profit rate is rising and where it is falling, is what produces a relationship between supply and demand such that the average profit is the same in the different spheres, and values are therefore transformed into prices of production.
>
> (Marx, 1981: 297)

These appear to support the hard version, formalised by Farjoun and Machover as the system of n input-output equations

$$P_0 = \mathbf{p}_i + R\ \ k$$

for each of n industries in economy, where P_o is the price of a unit of output, $\mathbf{p}_i$ is a vector of costs of inputs used up per unit of output, R is the uniform rate of profit and k represents capital per unit of output.

The soft version, in their words, says that 'in a competitive economy, the values of [the rate of profit] for all different branches [of the economy] must be very close to each other, and for theoretical purposes can be taken as uniform across the whole economy' (page 20) and since this doesn't specify what is meant by a branch, if one takes large enough sectors giving, say, a dozen branches (and a long enough time period too) then this becomes realistic and plausible.

In fact, as they point out immediately after this, their quotations from Marx, read in context, support the soft and not the hard version of profit uniformity. We have also seen (Chapter Two) that they have no objections to this version. Hence it is difficult to accept their later statement (page 131) about the 'irony' of Marx's acceptance of 'the economists' story of the alleged tendency of the rate of profit toward uniformity'. As we have seen they – and according to them, Marx – accept that this tendency is real.

Indeed, the real irony is twofold: not merely was Marx's view of profit rate equalisation the one Farjoun and Machover prefer: that of an equilibrium distribution, but this view is part of a consistently probabilistic and statistical outlook throughout his career as a whole, and in *Capital* in particular. Indeed, Farjoun and Machover themselves recognise this elsewhere when they quote Marx's claim that 'the true law' of economics is '*chance*' (Farjoun and Machover, 1985). We now examine Marx's views in more detail.

5.1.3 Marx's probabilistic political economy

We claim that Farjoun and Machover's approach is faithful to Marx in so far as the latter's approach was clearly and consciously statistical from first to last. Here we indicate some evidence for this, beginning with that which they themselves bring forward. The claim that 'the true law of economics is chance' is made very early in Marx's career, in the 'Notes on James Mill' written in 1844:

> [I]n his demonstration that the cost of production is the sole factor in the determination of value Mill succumbs to the error … of defining an *abstract law* without mentioning the fluctuations or the continual suspension through which it comes into being. If *e.g.* it is an

> *invariable* law that in the last analysis - or rather in the sporadic (accidental) coincidence of supply and demand – the cost of production determines price (value), then it is no less an *invariable law* that these relations do not obtain, *i.e.* that value and the cost of production do not stand in any necessary relation. Indeed, supply and demand only ever coincide momentarily thanks to a previous fluctuation in supply and demand, to the disparity between the cost of production and the exchange value. This is the *real* movement, then, and the above-mentioned law is no more than an abstract, contingent as one-sided movement in it. Yet recent economists dismiss it as accident, as inessential. Why? Because if the economists were to attempt to fix this movement in the sharp and precise terms to which they reduce the whole of economics this would produce the following basic formula: laws in economics are determined by their opposite, lawlessness. The true law of economics is *chance*, and we learned people arbitrarily seize on a few moments and establish them as laws.' (Marx, 1975a: 259–260; emphases in original)

Farjoun and Machover (1985) quote only the final two sentences of this,[53] in support of the mild claim that '[e]conomists and economic philosophers have often pointed out the essentially indeterminate and statistical nature of economic categories such as price and rate of profit'. Taken in isolation this might be understood as little more than the claim that good-sized confidence intervals should be installed around the claimed values of any economic data.

But it can be argued that Marx is not talking about the accuracy of data, but about the nature of the process which causes it to have one value rather than another. In fact Marx's interest in the role of chance is manifest in his very first serious intellectual production, his doctoral dissertation (Marx, 1975b). In this work, Epicurus' physics is praised over that of Democritus on the specific ground that Epicurus introduces chance as a way of making room for human free will (McLellan, 1980).

---

[53] In slightly different wording, because they cite the *Collected Works*.

Moreover, the importance to Marx of a probabilistic methodology of economics can be seen from the fact that the ideas of the notes on Mill are repeated in his first published comments on political economy. In the pamphlet *Wage labour and capital* – originally written as a lecture to the German Workingmen's Club in Brussels in 1847 – he summarises his thought thus: 'the total movement of this disorder is its order' (Marx, 1952). Still clearer is the formula in *The poverty of philosophy* (Chapter One, section 2), written in the winter of 1846-47 against Proudhon:

> If M. Proudhon admits that the value of products is determined by labour time, he should equally admit that it is the fluctuating movement alone that in a society founded on individual exchanges makes labour the measure of value. *There is* no ready-made constituted 'proportional relation', but *only a constituting movement.*
>
> (Marx, n.d.: 71; emphases added)

With this in mind we turn to Marx's comments in *Capital* Volume III:

> Competition distributes the social capital between the various spheres of production in such a way that the prices of production in each of these spheres are formed after the model of the prices of production in the spheres of mean composition, *i.e.* k + kp′ (cost price plus the product of the average rate of profit and the cost price) … the rate of profit is thus the same in all spheres of production, because it is adjusted to that of these average spheres, where the average composition of capital prevails.
>
> (Marx, 1981: 273)

On the face of it this passage – so far – appears to support the 'hard' version of profit uniformity discussed by Farjoun and Machover. But Marx continues:

> … Between those spheres that approximate more or less to the social average, there is again a tendency to equalization, which seeks the 'ideal' mean position, *i.e.* a mean position which does not exist in reality. In other words, it tends to shape itself around this ideal as a norm.
>
> (Marx, 1981: 273)

Two points: first, Marx says that the 'ideal' mean position *does not exist*: this is an echo of the distinction drawn by Quetelet between a real average (*moyenne*) and an arithmetic average (*moyenne arithmetique*) (Mosselmans, 2005). The former denotes, for example, finding the position of a star as an average of several observations (there really is just one star there, and at some particular place rather than another). The latter denotes such activities as finding the height of the 'average man', as Quetelet dubbed him, by averaging the heights of a number of different men; but there may be no actual man whose height is equal to the average.[54] Marx read Quetelet (the International Institute for Social History in Amsterdam has a notebook containing his reading notes on Quetelet's *Treatise* (International Institute for Social History, 2006)) and cited him elsewhere, as we will see below. In this light it is hard to read Marx as asserting an actual equalisation, even within those spheres that 'approximate more or less to the social average'.

Second, Marx describes the tendency to profit rate equalisation as a tendency to s*hape itself around* this ideal [mean] as a norm (our emphasis); profit rate equalisation *is* the 'shape around the ideal' towards which the tendency is directed – in other words, it is the formation of a profit-rate distribution ('shape itself around' the mean).

This might be thought a fanciful interpretation were it not that a few pages later, when Marx discusses intra-industry variations in productivity (the difference between the individual value of a commodity and its social value), he not only describes a probability density but also discusses how variations in its shape – symmetric or not, light- or heavy-tailed – will affect the relation of the mean to the whole; he even considers the effect of censoring some of the data (1981: 283-284).

---

[54] The status of the 'average man' is controversial: Quetelet thought that the individual who most nearly attained to the average in every faculty would represent the genius of the age: 'A man can have no real influence on masses – he cannot comprehend them and put them in action – except in proportion as he is infused with the spirit which animates them, and shares their passions, sentiments and necessities, and finally sympathises completely with them. It is in this manner that he is a great man, a great poet, a great artist. It is because he is the best representative of his age, that he is proclaimed to be the greatest genius.' (Quetelet, 1842 [1969]).

In thus considering distributions that might be either non-symmetric or heavy-tailed or both Marx was well ahead of the ideas of many professional statisticians of his day. The Gaussian distribution acquired its misleading alternative designation as 'normal' precisely because many thought it was the distribution to be expected in most social and biometric data; it is of course symmetrical, and in the usual Pearsonian system of moments the description of a distribution as having 'light' or 'heavy' tails is relative to the Gaussian kurtosis. Although the gamma distribution was first derived by Laplace as far back as 1836 (Johnson *et al.*, 1994: 343), it became better known after it was revived by Pearson as part of his system of distributions (in which it is Type III) in 1893. Even then there was considerable resistance to idea that asymmetric distributions might be a useful tool (Stigler, Stephen M., 1986: 333–341).

Finally we come to Marx's concept of social labour. This is the only point where Marx refers to Quetelet in connection with definitely economic ideas: in the first footnote to Chapter 13 of *Capital* Volume 1 Marx cites Quetelet (and Edmund Burke) in support of his explanation of his concept of average social labour:

> Edmund Burke, that famous sophist and sycophant, goes so far as to make the following assertion, based on his practical observations as a farmer: that 'in so small a platoon' as that of five farm labourers, all individual differences in the labour vanish … For example, let the working-day of each individual be 12 hours. ... From the point of view … of the capitalist who employs these 12 men, the working-day is that of the whole dozen. … *But if the 12 men are employed in six pairs, by six different 'small masters', it will be entirely a matter of chance whether each of these masters produces the same value*, and consequently whether he secure the general rate of surplus-value. … The inequalities would cancel out for the society as a whole, but not for the individual masters.
>
> (Marx, 1976: 440-441)

For Marx, and the capitalist, the product of the working day is the total labour of the workers employed; this is a random variable found by summing the random variables constituted by the individual labours of each worker, which are social labour only in so far as they are employed by capitalists employing other social labours.

Here too we see evidence that Marx's probabilistic approach put him, implicitly, in the vanguard of contemporary statistical work; his description of the output of the collective working day as a sum of random variables is an informal version of the central limit theorem (CLT); although this was moderately advanced for the mid-nineteenth century it was hardly novel, a version of the CLT having first been described by de Moivre in the seventeenth century and revived by Laplace at the beginning of the nineteenth (Stigler, Stephen M., 1986: 136). The ordinary version of the CLT covers only sums of distributions with finite variance, that is, with light tails. We have seen that Marx contemplated heavy-tailed distributions; moreover the CLT points to the normal distribution as that of the summed variables, which is symmetrical, whereas the distribution of the value output of a 'gang' of workers is surely bounded below by zero. However, the generalised central limit theorem, which relaxes the restriction to finite variance and allows asymmetric, heavy-tailed distributions as the outcome, was not developed until the 1920s (see also Chapter Seven).

5.1.4 Testing Farjoun and Machover

In this section we discuss various considerations relating to practical tests of Farjoun and Machover's system, and in particular their hypothesis that the rate of profit should have a gamma distribution.

We begin by noting that their full hypothesis is a complex one about the rate of profit R and their rate of wage bill variable Z.[55] As discussed in section 5.2.1 above, they argue that the random variable R/Z, the analogue of the rate of surplus value, should be an almost-

[55] In fact, a complete test of their system would also involve their propositions about the market space. But such a test would require information about the prices paid in individual transactions for many commodities. In principle one might do this, since exactly such data is now collected and stored electronically by both manufacturers and distributors (so-called 'data warehousing').

degenerate one with value = 1. In fact, the full argument on this point also involves the hypotheses that R and Z can both be expected to be gamma distributed with a common scale parameter , and that if their shape parameters are $_R$, $_Z$, then the random variable $R + Z$ should have a gamma distribution with parameters $(_R + _Z,)$.

It is clear from their discussion that it is only the wages of productive workers which should be included in Z. Unfortunately our FAME database does not allow us to recover this information directly. It is included in the cost-of-sales variable COST (which is why we are able to calculate Gillman's Marxian measures of the profit rate), but the FAME variable REMU covers wage-related payments to all employees, not just those who are productive, in the Marxist sense. For this reason we confine ourselves to testing the rate of profit aspect of Farjoun and Machover's work.

Given that we are going to test a variety of profit rate measures we have to consider a range of possible outcomes. We argued in Chapter Two that Farjoun and Machover's discussion of what they understand by 'rate of profit' suggested that their conception was most closely matched by the measure we have labelled Gillman 4 (or, less confidently, Gillman 3). However, more than one measure might turn out to have a gamma distribution, and these might or might not include Gillman 3 or 4.

There is also the question of which version of the gamma distribution we consider. Farjoun and Machover (1983) consider only those versions of the gamma with two parameters (shape and scale), but as shown in the Appendix the general form of the gamma includes a location (threshold) parameter and a secondary shape parameter. Farjoun and Machover explicitly discount consideration of gamma distributions with a negative threshold, on the grounds that loss-making firms are not a significant feature of the economy (1983:66-67).[56] This is, firstly, because they say that firms which makes losses are not only few in number, but generally of small size, and thus account for only a small

[56] Implicitly they also rule out positive threshold parameters.

proportion of the total social capital. Second, their profit rate is calculated before distribution of extra-firm claims such as interest, and few firms will make losses on this basis.

This might be thought excessively restrictive; it can be expected that there will always be firms which make losses, either because they are start-ups, or because they are in danger of failing. Either condition can be viewed as a normal part of the competitive process. If the threshold is treated as a parameter to be estimated, then it has an obvious economic interpretation as the maximum rate of loss compatible with survival, where the fact of survival is determined by social capital's assessment of the stewardship of the firm's managers (see our discussion of Bryer (1994) in Chapter Two).

Although Farjoun and Machover discount the idea of the profit rate distribution having a negative lower bound they do not explicitly rule out positive values for the location parameter (although their own small-scale test does assume location at zero).

Because the four-parameter gamma distribution includes a number of other distributions as special or limiting cases (the latter including the log normal), one might consider both the Farjoun and Machover and Gibrat hypotheses to be special cases of a more general hypothesis yet to be formulated.[57] According to Lienhard and Meyer (1967) the second shape parameter is related to subsidiary details in a number of processes known to generate such distributions as limiting outcomes over time.[58]

We regard the question of the location of the distribution (and also its scale to be fundamentally secondary to the question of its shape. If the distribution of a proft rate

---

[57] McDonald introduces the generalised gamma in connection with income distribution. A much earlier proposal in this connection is that of Amoroso (1925), so Gibrat (1931) was arguably out-of-date on publication.

[58] Some other uses of this distribution in economics include Creedy *et al.* (1994), as a model for the distribution of earnings in the labour market, and Creedy and Martin (1994), as a model for the distribution of market prices in general.

measure can be shown to have a shape which *cannot* be described by any member of the generalised gamma family, appropriately scaled and shifted, we will interpret this as *strong* disconfirmation of Farjoun and Machover's hypothesis; if the distribution *can* be shown to be describable by *some* member of the generalised gamma, but not the three-parameter (shape-scale-location) version, we will regard it as *weak* disconfirmation. Table 5.1 summarises the possibilities.

Table 5.1: testing Farjoun and Machover (1983)

Gillman 3/4 is gamma (n° of parameters)	**Alternative measure is gamma**	**Farjoun and Machover confirmed?**	**Comment**
Yes (two)	Not discussed	Yes	Small-scale test in Farjoun and Machover (1983)
Yes (three)	Yes	'Strong' confirmation	Ideal gas metaphor has profound implications for theory of capitalist competition (1)
Yes (three)	No	'Neutral' confirmation	Ideal gas metaphor is useful and profit rate concept supported (2)
No	Yes	'Weak' confirmation	Ideal gas metaphor is useful, but profit rate concept must be re-thought (3)
Yes (four)	Yes/no	'Weak' disconfirmation	Strict interpretation of Farjoun and Machover disconfirmed; their hypothesis a special case of a wider model (4)
No	No	'Strong' disconfirmation	Ideal gas metaphor not useful, but no particular implications for 'econophysics' approach in general (5) No particular implications for 'marxist' concept of the profit rate (6)

What are the implications of these various possible outcomes?

(1) 'Strong' confirmation: if not only Gillman 3/4 but also one or more conceptually very different measures turn out to have gamma distributions then the ideal gas metaphor would seem have quite profound implications for our understanding of capitalist competition and the way in which claims on surplus value arise.

(2) 'Neutral' confirmation: if only Gillman 4 (and/or Gillman 3, possibly) has a gamma distribution, then Farjoun and Machover are clearly vindicated. Subsequent work to explain the distributions displayed by other measures would clearly be of interest.

(3) 'Weak' confirmation: if one or more profit rate measures had gamma distributions, *but not* Gillman 3/4, this would support the ideal gas metaphor as a way of thinking about capitalist competition, but would also suggest that the appropriate definition of the profit rate needed to be rethought. Farjoun and Machover's hypothesis about the profit rate distribution does not depend on any formal model of the distribution generating process, and certainly not on one relying on particular properties of the rate of profit they discuss.

(4) 'Weak' disconfirmation, discussed above; the strict interpretation of Farjoun and Machover cannot be confirmed, but a broader interpretation can be sustained.

(5) 'Strong' disconfirmation would not necessarily discredit the general approach of the 'econophysics' movement discussed in Chapter One (the application of the statistical mechanics paradigm to economics). But the distribution(s) actually found would clearly have to be accounted for in a different way to that advanced by Farjoun and Machover.

(6) Finally, strong disconfirmation would seem to have no particular consequences for one's view of which profit rate measure was appropriate for empirical work in marxist economics.

5.2 Distributions of rates of return

In this section we apply our *L*-moment ratio analysis to profit rate distributions, investigate the properties of the sampling process by which they are derived, and show that for all the profit rate measures investigated both the size-weighted and unweighted versions have distributions with extended tails displaying power law characteristics. We conclude with a discussion of the further lines of research suggested by these results.

5.2.1 *L*-moment ratio analysis of profit rate distributions

Here we apply the methods developed in Chapter Four to investigate the distributional models of different profit rate measures; we use randomly-sized random samples to estimate the *L*-skewness $_3$ and *L*-kurtosis $_4$ for the data from each of the five years 1991–5, plot the resulting clouds of points and assess their relationship to the loci of the distributions given by Hosking and Wallis (1997). As in previous chapters we group the results in four sections: the marxian measures defined by Gillman (1956), the same author's 'capitalist' measures, the four standard accounting ratios, and the eight measures discussed in Glick (1985).

To assist comparison we begin by showing the clouds of points in a common sub-set of the total *L*-skewness, *L*-kurtosis space, namely –0.25 ≤ $_{3}$ ≤ 1 and 0 ≤ $_{4}$≤ 1 (Figures 5.1 to 5.4); for reference purposes we include the loci of distributions previously shown in Figure 4.1, as provided by Hosking and Wallis (1997).[59] The points representing each year are identified by the appropriate final digit.

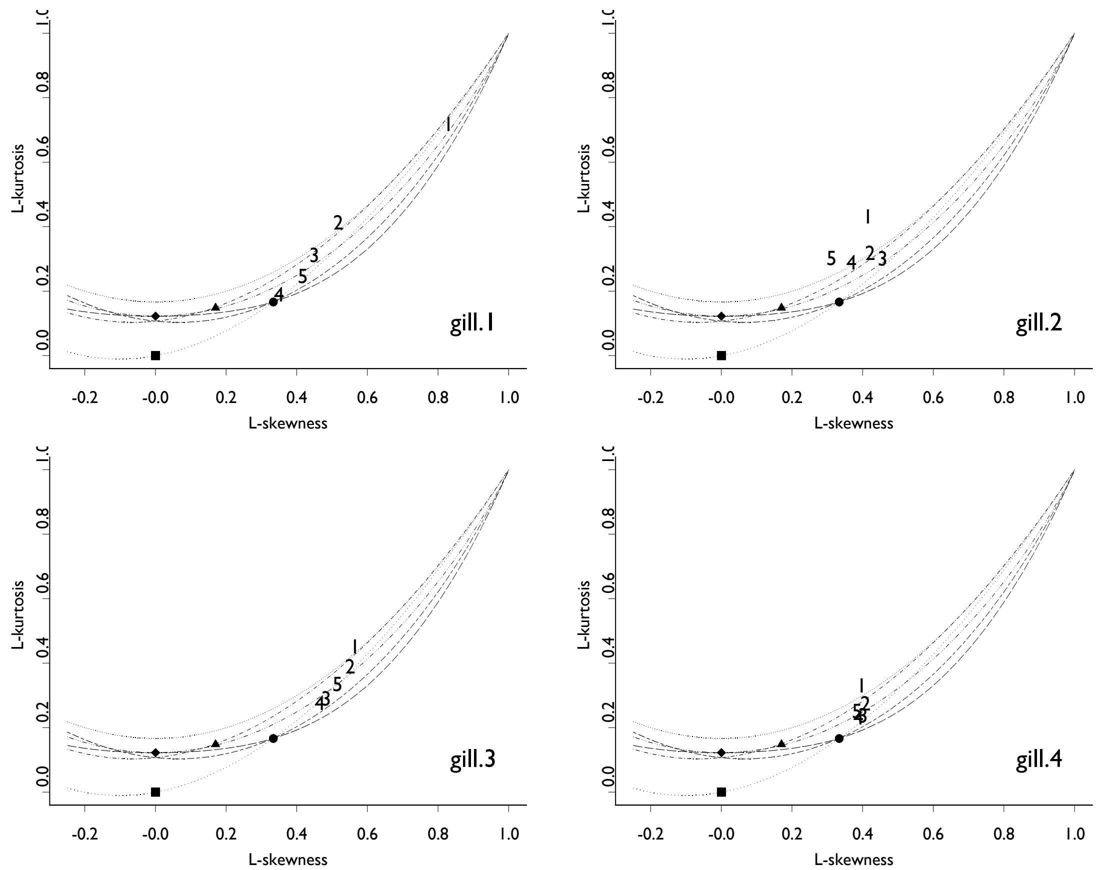


Figure5.1: *L*-skewness and *L*-kurtosis of Gillman marxian measures

[59] To recap, these are the uniform (identified by ■), the Gaussian (◆), the exponential (●), the Gumbel (▲), and the linear loci of six three-parameter distributions: from top to bottom at the =0.25 ordinate they are the generalised logistic, the generalised extreme value, the three-parameter log normal, the three-parameter gamma, the Weibull, and the generalised Pareto.

All four of Gillman's capitalist measures fall within the broad band encompassing Hosking and Wallis's loci (Figure 5.1). However, the point clouds associated with each measure – apart perhaps from Gillman 4 – are too widely-scattered to be readily associated with the locus of any particular one of the distributions shown. Gillman 4 is the measure we identify as the one closest to the profit-rate concept adopted by Farjoun and Machover (1983), so its relatively-tight clustering appears encouraging. In the next section we will investigate this more closely.

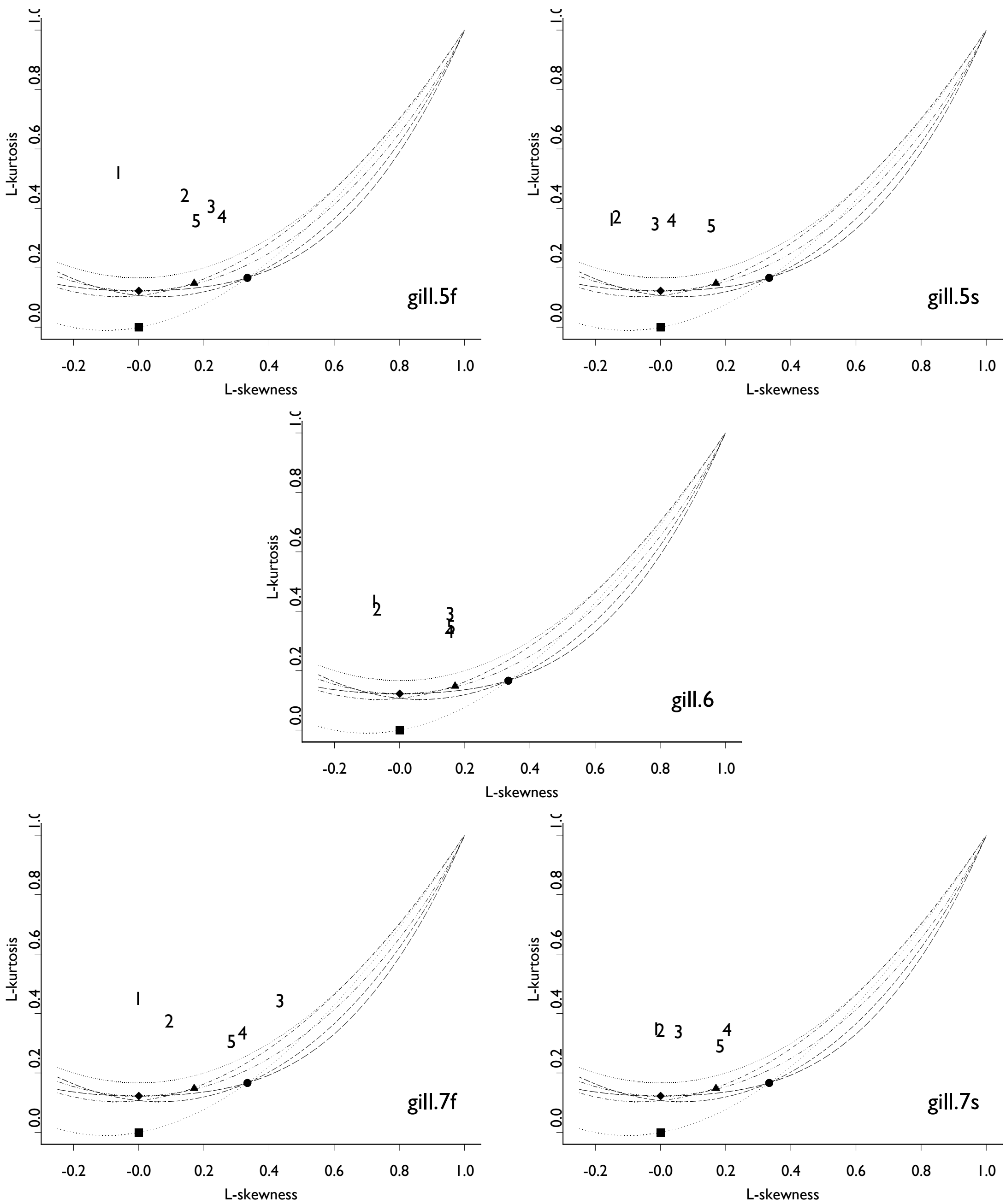


Figure5.2: *L*-skewness and *L*-kurtosis of Gillman capitalist measures:

In contrast, Gillman's 'capitalist' measures all display greater *L*-kurtosis, relative to *L*-skewness, than that attained by any of the Hosking and Wallis distributions (Figure 5.2). Moreover, their kurtosis is, relatively, little-changed from year to year, unlike their skewness, which displays a much wider range than for any of the marxian ratios apart from Gillman 1. A further notable contrast is that they all have years in which the estimated *L*-skewness is either close to zero or actually negative.

An intriguing feature of Gillman 5s is that the skewness is date-related, being at its smallest in 1991, then increasing monotonically through 1995; others of these measures have a partially-similar ordering.

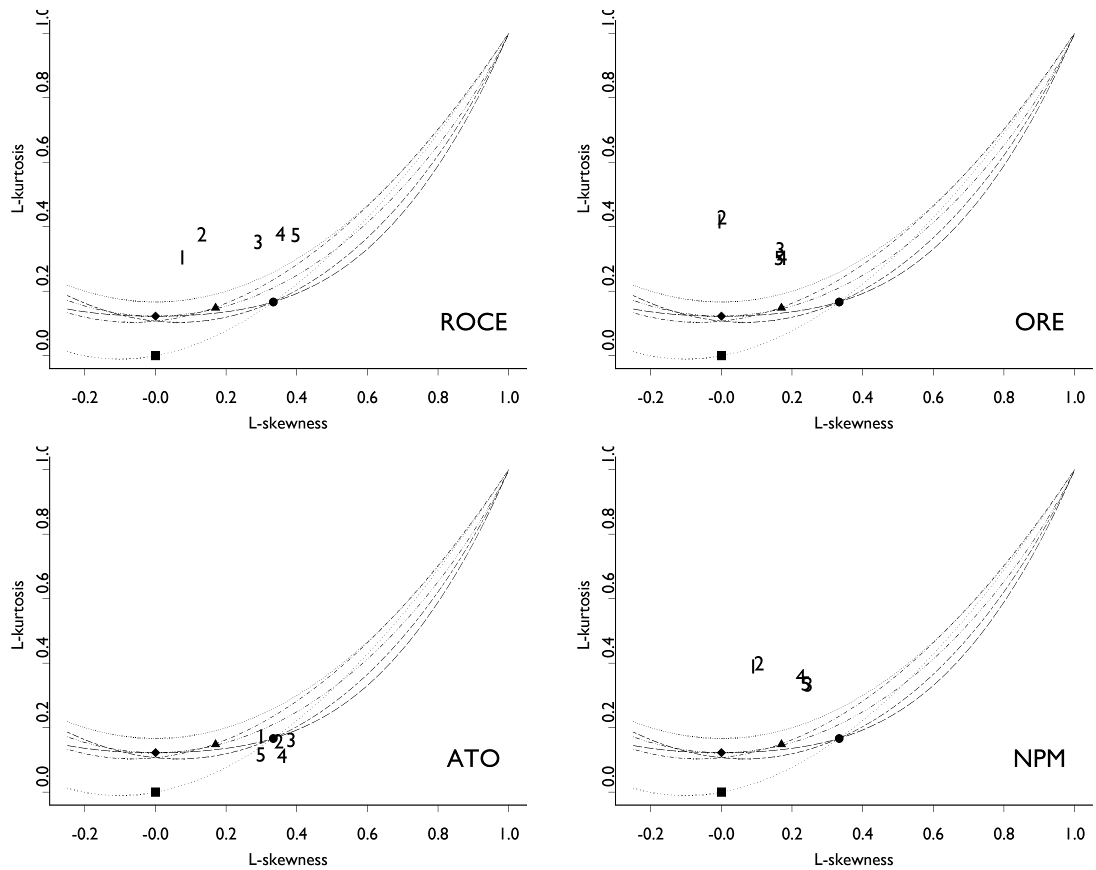


Figure5.3: *L*-skewness and *L*-kurtosis of accounting ratios

Turning to the accounting ratios we find similar results to those for Gillman's 'capitalist' ratios (Figure 5.3): high kurtosis which is relatively unvarying in comparison to skewness, which again shows signs of being date-related.

The sole exception is the asset turnover ratio (ATO). We know from work in the previous two chapters that this measure is unusual: in unweighted form it is clearly bi-modal (Figure 3.5), while the weighted version displays a tail which also hints at the existence of a second mode (Figure 4.5). However, if one neglects this tail the histogram does suggest an exponential form, which is consonant with the clustering around the locus of the exponential distribution (•) seen in Figure 5.3.

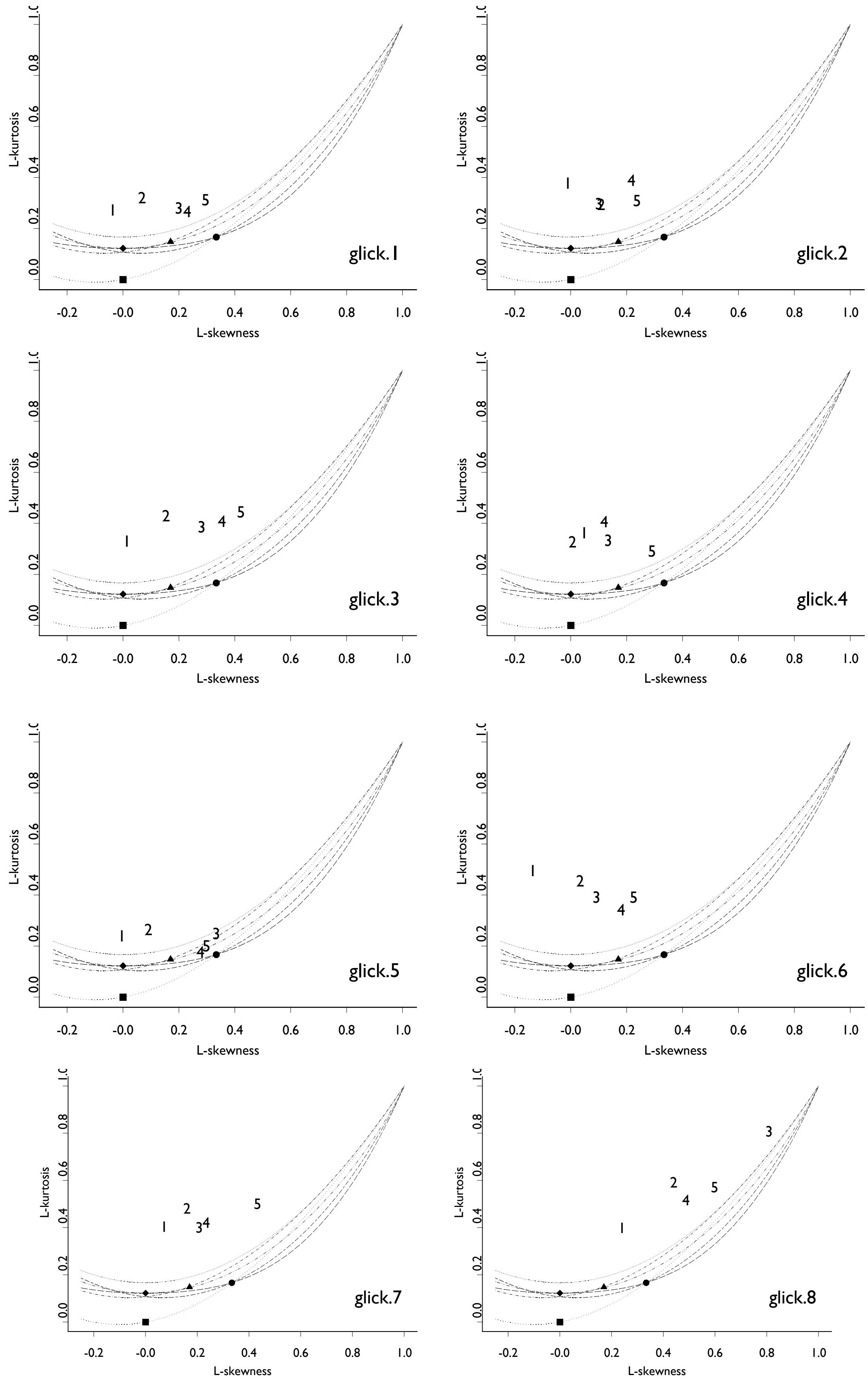


Figure5.4: *L*-skewness and *L*-kurtosis of Glick measures

Finally, the majority of the Glick measures show results similar to those of the accounting ratios, unsurprisingly given their close conceptual relationships. The exceptions are Glick 5 (relatively low levels of *L*-kurtosis) and Glick 8 (wide range of *L*-kurtosis, which is in the main closer to the level corresponding to its *L*-skewness in the Hosking and Wallis distributions). However, even these display the tendency for skewness to be time-related. We will return to this in section 5.4.3 below.

We have seen that of all the 21 profit rate measures tested no less than 19 seem to fall outside the loci of those distributions we can identify using Hosking and Wallis's *L*-moment ratio system. In particular, they do not seem able to be associated with any case of the four-parameter gamma distribution. In the scheme set out in Table 5.1 above we thus have either row 2 ('neutral' confirmation of Farjoun and Machover, with either Gillman 3 or 4 found to be gamma), or row 4, 'weak' disconfirmation (with these two Gillman profit rate measures not gamma distributed). However, we suggested that disconfirmation could also come in strong or weak forms, with the latter consisting of showing that although Gillman 3/4 could not be modelled by a three-parameter gamma distribution, one of the special cases of the four-parameter gamma would provide a model.

As seen in Figure 5.1, Gillman 4 does appear to be a candidate for association with a particular distribution using Hosking and Wallis's *L*-moment ratio system (so does ATO, but this is unlikely by reason of other information about its form). Moreover four of the distributions for which loci are shown in Figure 5.1 are special cases of the four-parameter gamma (the exceptions are the generalised versions of the logistic and Pareto). Hence the locus of the four-parameter gamma is the region bounded by the envelope of these four distributions.

Thus if Gillman 4, the definition most strongly associated with Farjoun and Machover's hypothesis of a gamma model, can be shown to have either a generalised logistic or generalised Pareto distribution we have strong disconfirmation, and if one of the others, weak disconfirmation. We therefore investigate it in more detail.

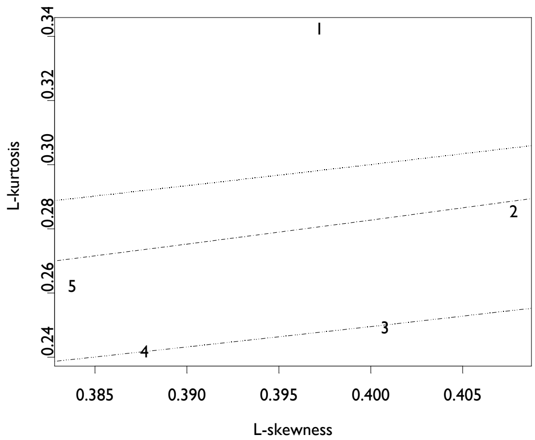


Figure5.5: *L*-skewness and *L*-kurtosis of Gillman 4 in closer focus

Figure 5.5 focuses on the region of *L*-skewness, *L*-kurtosis space containing the observations on Gillman 4. Four years' observations fall in a region bounded above by the locus of the generalised extreme value distribution and below by the locus of the log-normal distribution. Since these are both special cases of the four-parameter gamma this is evidence of weak, as opposed to strong, disconfirmation of Farjoun and Machover (but this is equivalent to confirmation of a broad interpretation of their hypothesis).

However, the absolute variation in *L*-skewness is extremely small compared to that in *L*-kurtosis (masked here by the proportions of Figure 5.5, but more evident in Figure 5.1). Also, the fifth observation, that for 1991, has *L*-kurtosis considerably higher than for the other years, being well above the locus of the generalised logistic distribution.

Clearly we would like to confirm or reject these hypotheses more definitely. There is also the anomalous observation for 1991 to take into account.

5.2.2 Sampling properties of RS2 estimation

As a first step we investigate the sampling properties of our RS2 estimation method in the case of the 1995 data for Gillman 4 (Figure 5.6). Although they will not be reported here, investigations similar to those described below have been made for the other profit rate measures in our study; their results are qualitatively similar and suggest similar conclusions to those that will be drawn below.

It will be recalled that the RS2 sampling procedure involves taking 100 samples from our data, samples in which the probability of any company's profit rate value being included is proportional to the company's size relative to the largest company, as measured by the capital definition involved in the relevant profit rate measure.

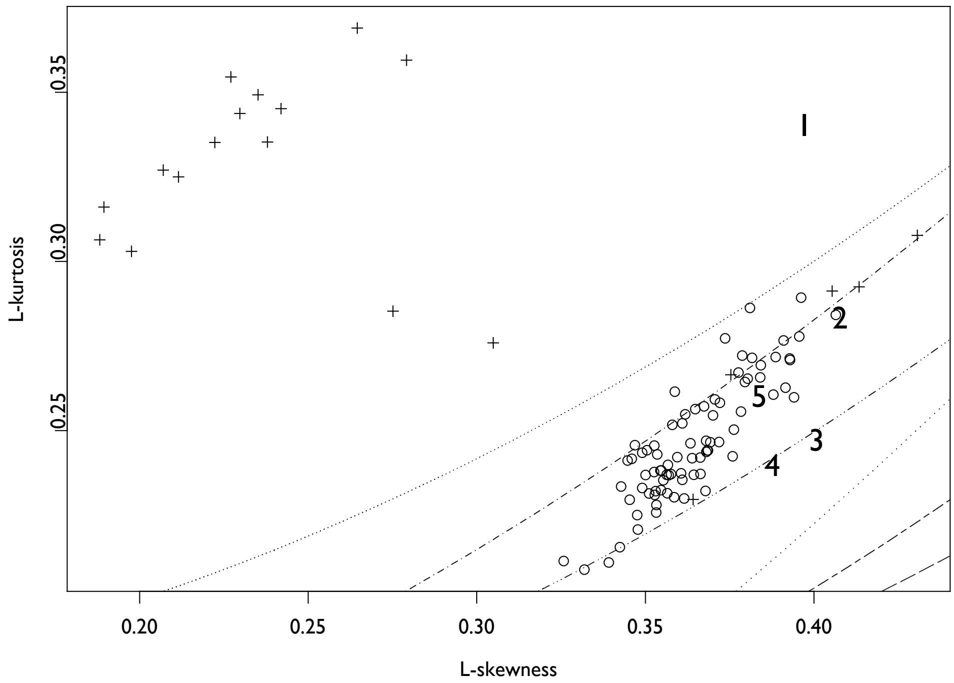


Figure5.6: Gillman 4, 1995; *L*-skewness and *L*-kurtosis of RS2 samples

We begin by plotting the *L*-skewness and *L*-kurtosis of the 100 RS2 samples from 1995. Figure 5.6 shows the majority of samples as open circles; crosses indicate samples identified as discordant by reason of their distance from the main body of samples, using Hosking and Wallis's suggested test (1997:45ff). The estimated *L*-skewness and *L*-kurtosis for each year of our data are plotted by the appropriate final digit.

Several features are notable. First, the samples form two distinct clouds, a main one to the south-east of the plot and a subsidiary cloud of some dozen points to the north-west. The reason for this is unknown but we believe that it is a random outcome, as it did not appear in a re-run of the sampling procedure. More extensive investigation of this was deferred on the grounds of the heavy computational load.

Second, consider the ranges of *L*-skewness and *L*-kurtosis exhibited by the samples: both more than cover the range of annual variation in estimated values, even when the anomalous result for 1991 is included.[60] The difference is especially pronounced in the case of *L*-skewness.

Third, if we consider only the main cloud of samples, the distributions of both *L*-moment ratios have pronounced skewness, with upper tails longer than the lower.

Fourth, the tails produce an impression of strong correlation of the sample *L*-moment ratios.

Fifth, the estimated *L*-moment ratios for 1995 lie outside the core of the main cloud of points (which, recall, represent samples from the 1995 data).

Taking the last point, if one assumed that this divergence between the samples and the estimates to which they give rise results from the skewness of the sample *L*-moment ratios, that this result was also a feature of other years' data, and that some appropriate correction would result in the *L*-moment ratio estimates being translated southwards, then the estimates for 1992–5 might well lie athwart the locus of the log-normal distribution (confirming Gibrat's hypothesis), and even the estimates for 1991 would look less anomalous (a point which would hold even more strongly if one assumed a translation proportional to the starting values).

To see how far such a corrective translation might be justified, and also to see the reason for the skewness of the sample *L*-moment ratios, we now look directly at the sample *L*-moments, as opposed to the *L*-moment ratios.

---

[60] The presence of a sub-group of samples with relatively low skewness and high kurtosis should also be noted, but investigating this was judged to take us too far from the objectives of our present study.

The fact that the estimated skewness and kurtosis both lie outside the main mass of sample values is not (directly) a result of the skewed distribution of the samples. To see this, recall that we estimate *L*-skewness and *L*-kurtosis *not* by averaging the sample values of these ratios, but by averaging the sample moments and taking the ratio of the appropriate results.

This is illustrated in Figure 5.7, where we plot cross-sections of the four-dimensional space containing the first four *L*-moments: (the mean) and $_{2}$, $_{3}$ and $_{4}$.

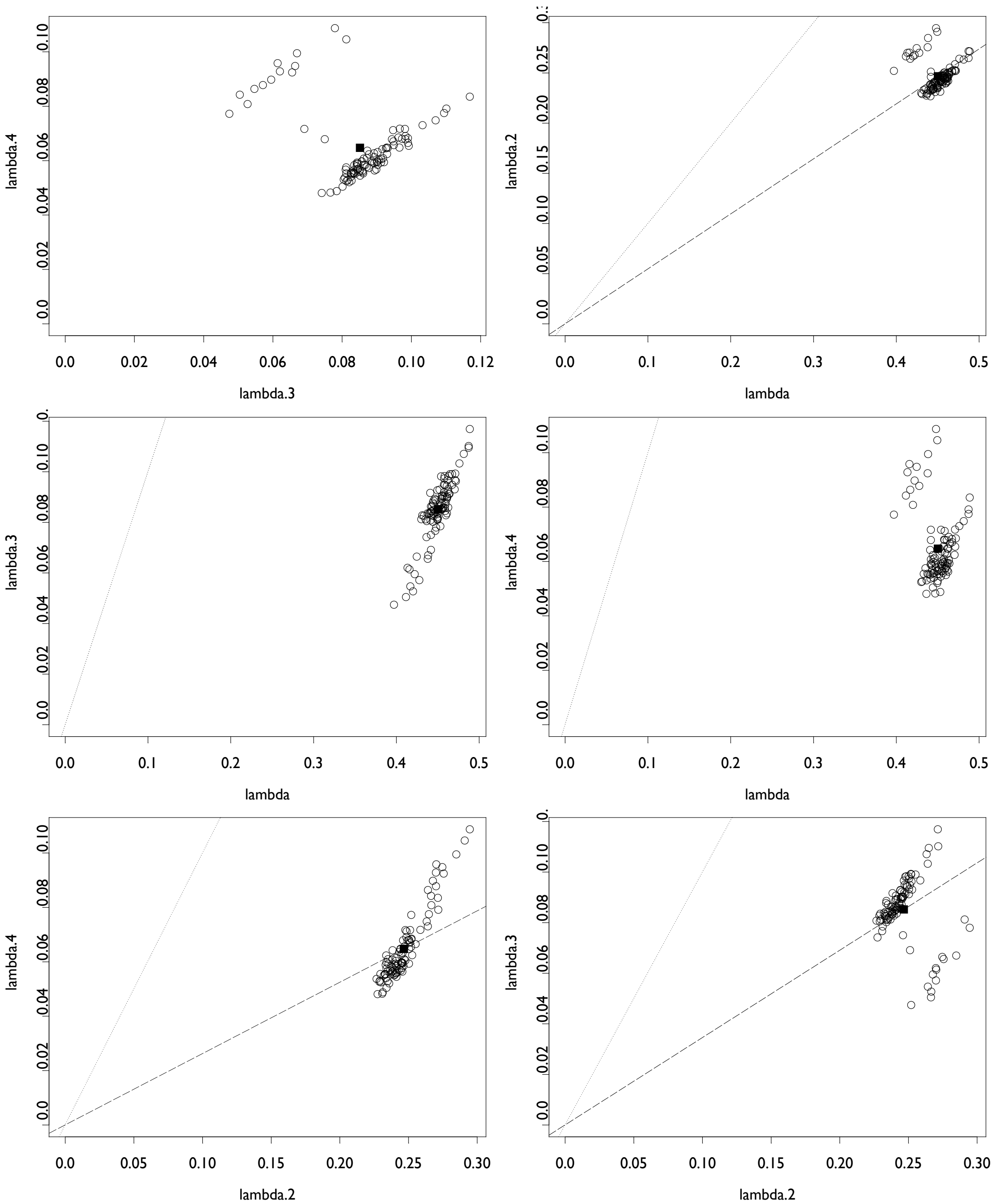


Figure5.7: Gillman 4, 1995; *L*-moment space of RS2 samples

Here the solid squares identify pairs of estimated *L*-moments, and dotted lines indicate the 45° degree line. In the $/_{2}$, $_{3}/_{2}$ and $_{4}/_{2}$ panels, dashed lines indicate a ray from the origin through the estimates of the population *L*-moments; the relevant ratio is thus measured by the gradient of the ray.

The first panel (top row left) is the $_{4}/_{3}$ plane. This *L*-moment cross-section bears a strong resemblance to its *L*-moment ratio counterpart. The reasons for this can be seen from examination of the other panels. First, while in each of these we see a similar pattern to that shown by the *L*-moment ratios – strong correlation between each pair of moments – the apparent degree of association is weakest in the $/_{2}$ plane. Second, in this case the projection of the axis of the four-dimensional point-cloud lies approximately along the 45° line, whereas in the other planes the axis of the cloud is displaced (while still remaining approximately parallel to the 45° line). Third, note the very strong relationship between the third and fourth moments, on the one hand, and the mean on the other (second row of Figure 5.7). Small proportional changes in the mean of the sample imply much larger changes in the third and fourth moments, but little change in the second moment (or in the coefficient of *L*-variation). Since the *L*-skewness and *L*-kurtosis are the ratios of the third and fourth moments, respectively, to the second, the implication is that these ratios are primarily sensitive to differences in the means of the higher moments of the samples.

The third row of Figure 5.7 shows the $_{3}/_{2}$ and $_{4}/_{2}$ cross-sections. Here we note the skewness shown by all three higher moments (but to a lesser degree by the mean).[61] The skewness of the sample moments obviously biases the estimates of each moment, which in turn biases the estimates of the ratios because of the high correlation among the moments. In principle, one might try to mitigate this by omitting samples identified as discordant by Hosking and Wallis's test. Doing so would tend to change the estimated value of each moment – but because the ratios are estimated by the slopes of the rays, it is also clear that this would be unlikely to make a significant difference to the estimates of the ratios (essentially because of the shallow angle between the axis of the four-dimensional cloud of points and the 45° degree line). In the case of Gillman 4 it appears that the *L*-skewness might actually *increase* slightly, rather than decrease: but because of the positive slope of the lognormal locus this would tend to improve the conformity to this locus, for any given decrease in *L*-kurtosis.

[61] We also again note the off-axis samples, which themselves show signs of mutual correlation.

We conclude that the relative lack of success of our RS2 and *L*-moments procedure in clearly identifying distributional models is not due to problems with the procedure.

Instead, recall that our RS2 sampling scheme is designed to attenuate the influence of very small companies, which we know are associated with very wide ranges of profit rate, and which we suspect are often not genuine capitalist entities. However, it is not designed to exclude them entirely, since we have no specific information on which to base the exclusion of any particular company.[62] Thus our sampling procedure will occasionally create samples which include a disproportionately high number of extreme values which give an upward bias to the sample's *L*-moments.

As we have shown, this does not in practice impart undue bias to the skewness and kurtosis estimates. However, the prevalence of samples containing extreme values is shown by the pronounced skewness of the sample of the higher *L*-moments. This directs renewed attention to the tails of the distributions, both weighted and unweighted. The relatively high prevalence of extreme values among the moments of the samples points to the fact that even the size-weighted version of this profit rate measure has a heavy upper tail.

By hypothesis, each year's profit rate observations are drawn from a common distribution (neglecting variations in parameters); *a fortiori* the samples from a given year's data are drawn from a common distribution. The latter evidently have wide variation in skewness and kurtosis, which we have shown to be linked to variation in the mean, itself subject to quite wide fluctuation due to the generation of relatively many samples with many extreme values.

We therefore conjecture that the observed large annual fluctuations in kurtosis for Gillman 4 are principally due to fluctuations occurring in the tails of the distribution, not in

---

[62] If we did, we could exclude it directly.

the main portion. (By extension, we conjecture that variation in the tails drives the annual variations in skewness and kurtosis observed for other profit rate measures.)

One could more confidently assert a four-parameter gamma model, bounded by the log normal and generalised extreme value distributions and thus weakly disconfirming Farjoun and Machover but confirming a more general version of their hypothesis, if one had some justification for omitting the tails from the analysis. We do not know of any such justification on *a priori* grounds, but a more detailed empirical investigation may suggest possible lines of approach. Thus in the next section we examine the tail structure of all our profit rate measures more closely.

5.3 Zipf plot analysis of tails of profit rate distributions

We now investigate the tails of the distributions using Zipf plots, in which the log of the order statistics is plotted on the horizontal axis and the log of their rank on the vertical.[63] Zipf plots are so-named for their ability to illustrate Zipf's Law (Zipf, 1932), a proposal about word frequency in texts; the Zipf distribution is closely related to the Pareto distribution. The plots were, apparently, introduced into economics by Stanley *et al.* (1995) in the analysis of firm-size distributions.

In a Zipf plot distributions from the exponential family (of which the gamma distribution is a member) have upper tails curved at all points. Power-law distributions, such as the Pareto, have straight tails. This is illustrated in Figure 5.8, which compares (a) the standard Gaussian, (b) the standard exponential, (c) a gamma distribution with shape = 2 and scale = 1, (d) a Pareto distribution with location = 10^{-2} and scale = 1.75, and (g) the standard lognormal distribution. For each distribution we plot a random sample with $n = 10^5$.

---

[63] In the interests of speeding both plotting and printing, we modify the full Zipf plot as follows: we show the 100 largest values, and then 1,000 smaller values evenly-spaced as to rank; further, since our empirical data varies in length we shift the plots of the shorter series upwards to give the plot of each series a common starting point as to rank.

Since we plot the log of the values we are of course only considering each distribution's support on the positive real line. The existence of cases where the L-skewness is negative suggests that a fuller analysis should also take into account the left-hand tails (see Tables 3.4, 3.5 and 4.2, and Figures 3.7 to 3.11 and 4.10).

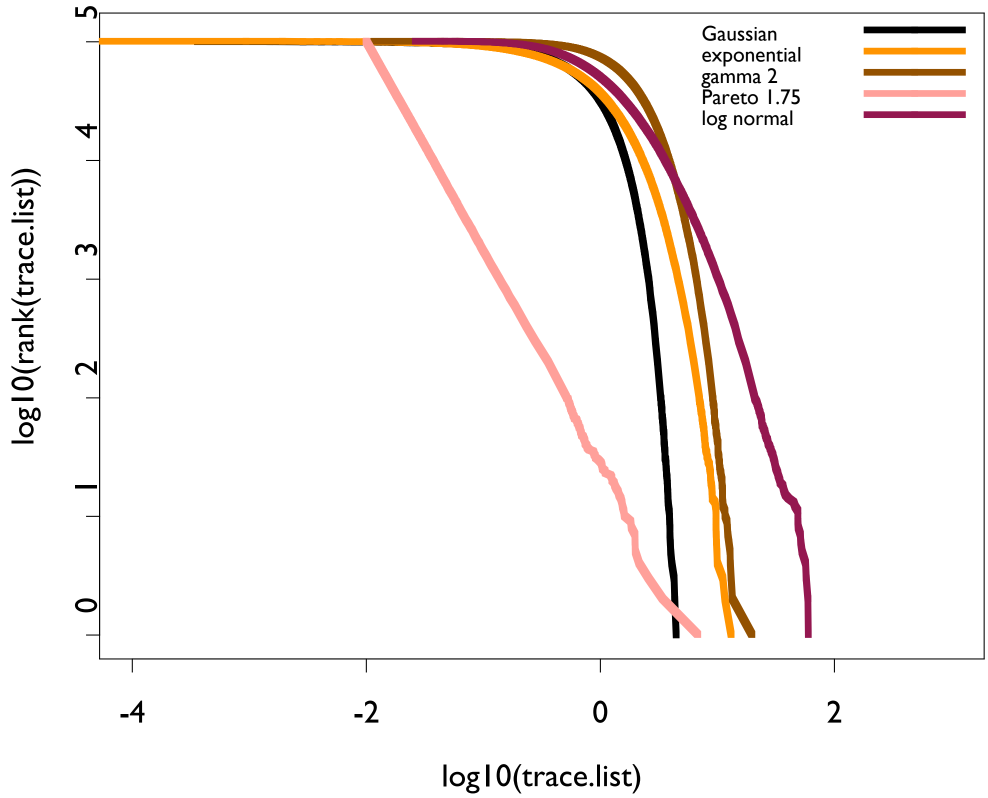


Figure5.8: Zipf plot of samples from exponential samples compared to the Pareto distribution

The difference between the tails of the various distributions of the exponential family and the straight line of the Pareto is obvious. The more subtle but nonetheless definite differences within the exponential family highlight the attraction of the L-moment system to hydrologists, with their need to accurately estimate the likely occurrence of extreme events.

5.3.1 Power-law tails and rates of profit

Given our investigation of the sample L-moments of Gillman 4, and its importance in testing Farjoun and Machover, we begin with that measure. As in Chapter Four, the data we use in discussing the company-size weighted measure is the concatenation of our 100 RS2 samples.

Figure 5.9 compares size-weighted and unweighted versions of Gillman 4; for purposes of comparison we include the plot of the gamma distribution with shape = 2, previously shown in Figure 5.8. Figure 5.10 presents part of the Zipf plot for each of the five annual estimates of the size-weighted distribution.

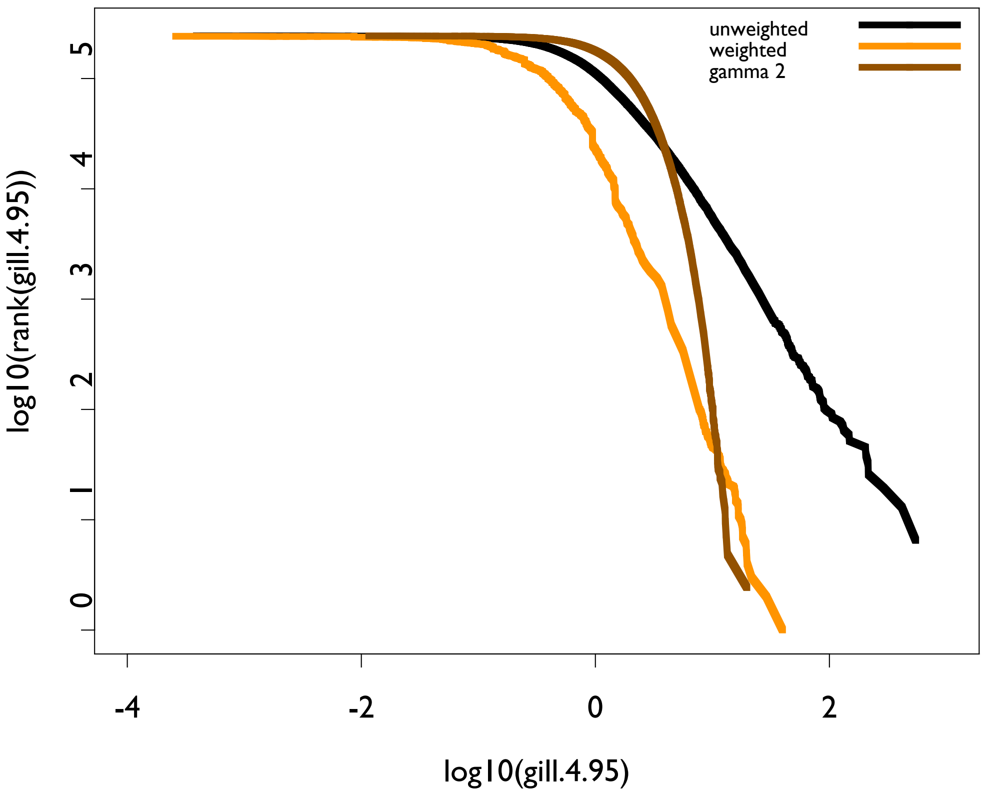


Figure5.9: Zipf plot of unweighted and weighted Gillman 4 measures, with gamma distribution

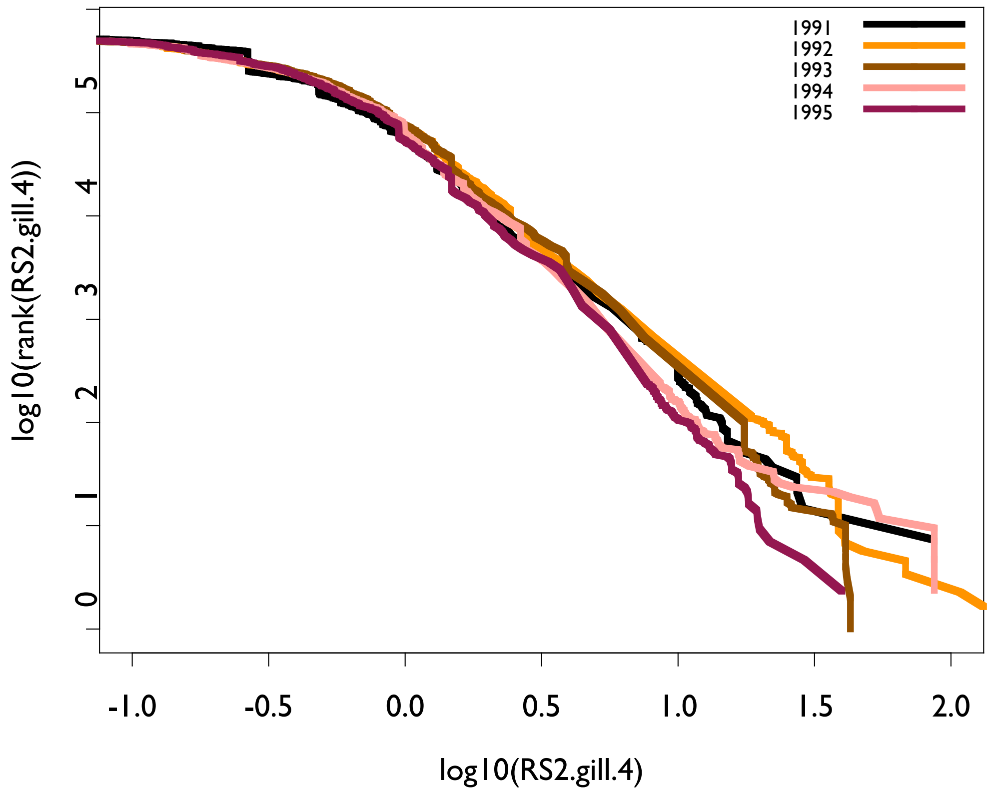


Figure5.10: Zipf plot of annual variation in weighted Gillman 4 measures

There is a good reason for the similarity of the log normal distribution and the power law tails of the empirical distributions, as Mitzenmacher (2003: 229) points out. Consider the logarithm of the density function of the log normal distribution

$$\ln f(x) = \quad \ln x \quad \ln\sqrt{2} \quad \frac{(\ln x \quad)^2}{2 \quad ^2}$$

$$= \quad \frac{(\ln x)^2}{2 \quad ^2} + \frac{}{\ ^2} \quad 1 \quad \ln x \quad \ln\sqrt{2 \quad ^2} \quad \frac{^2}{2 \quad ^2}$$

where and are the mean and standard deviation of the corresponding normal distribution. With sufficiently large, on a log-log plot the logarithm of the density function will appear approximately linear over several orders of magnitude. For comparison, consider perhaps the best-known power law distribution, the Pareto: with shape parameter and scale parameter *k*, the logarithm of the density function is exactly linear:

$$\ln f(x) = (\qquad 1)\ln x + \quad \ln k + \ln$$

As we have seen in Figure 5.8 the Zipf plot of the Pareto is likewise exactly linear (discounting deviations arising from the fact that we plot a sample, not the theoretical curve), and a similar relationship exists between the Pareto and log normal distributions in this case also (although for the log normal the density function is easier to work with).

Both weighted and unweighted versions of the profit rate measure display power law tails. To our knowledge this is the first time this phenomenon has been demonstrated for profit-rate distributions.

The slow rate of change of curvature of the plots makes it difficult to assess the precise start of the power law sections, but it appears that in both cases around 10 per cent of the data has the characteristic power law tail. In the weighted case the approximate range of profit rates included in this tail is from 1 (that is, 100 per cent) to 10 (1,000 per cent), where the estimated mean is 0.45 (45 per cent).

This combination of exponential law for the main mass of data with power law tail(s) is widely accepted as a stylised fact describing not only the firm-size distribution but also the distributions of wealth and income and the returns in financial markets.[64] On the firm-size distribution, see Sutton (1997) for a survey; more recent contributions include Fujiwara (2004) and Russo, Delli Gatti *et al.* (2006); on wealth and income, see Fujiwara (2003), and a very interesting recent paper by Braun (2006); on market returns see Mantegna and Stanley (2000).

Power law tails are a feature of all our profit rate distributions, as the figures below demonstrate. However, before these inter-profit rate measure comparisons we draw

---

[64] Interestingly Stanley *et al.* (1995), who as we have seen above claim to have introduced Zipf plot analysis into economics in connection with this topic, use it to show that with their data the firm-size distribution *does not* exhibit power law characteristics, since the tails are insufficiently heavy.

attention to the right-hand panel of Figure 5.9, comparing the distribution of the size-weighted version of Gillman 4 in different years.

The upper portion of each plot shows that the exponential portion of the data varies from year to year, presumably reflecting annual variation in the shape and scale of the distribution. But it also appears that the slope of (and hence the range covered by) the power law tail undergoes annual change.

Clearly the annual variation in the minimum of the power law tail is comparatively small. Compared to the variation in the maximum value attained it may be taken as zero at a first approximation. That being so, the range covered by the tail varies by about half an order of magnitude – say between $10^{1.6}$ and $10^{2.1}$, or 40 to 125 (profit rates of 4,000 per cent to 12,500 per cent). This supports our conjecture above (section 5.3.2) that it is variation in the tail of the distribution that drives the annual variation in *L*-kurtosis in the Gillman 4 definition of the profit rate.

We now compare the tail behaviour of different profit rate definitions, in both weighted and unweighted forms, and then examine the annual variations in tail behaviour of selected measures.

5.3.2 Comparison of power-law tails between profit rate measures

We begin by comparing the unweighted profit rate measure definitions, grouped as before into Gillman's marxian and 'capitalist' measures, the accounting ratios, and Glick's measures (in this case we present the Glick measures in two panels to aid clarity in plotting).

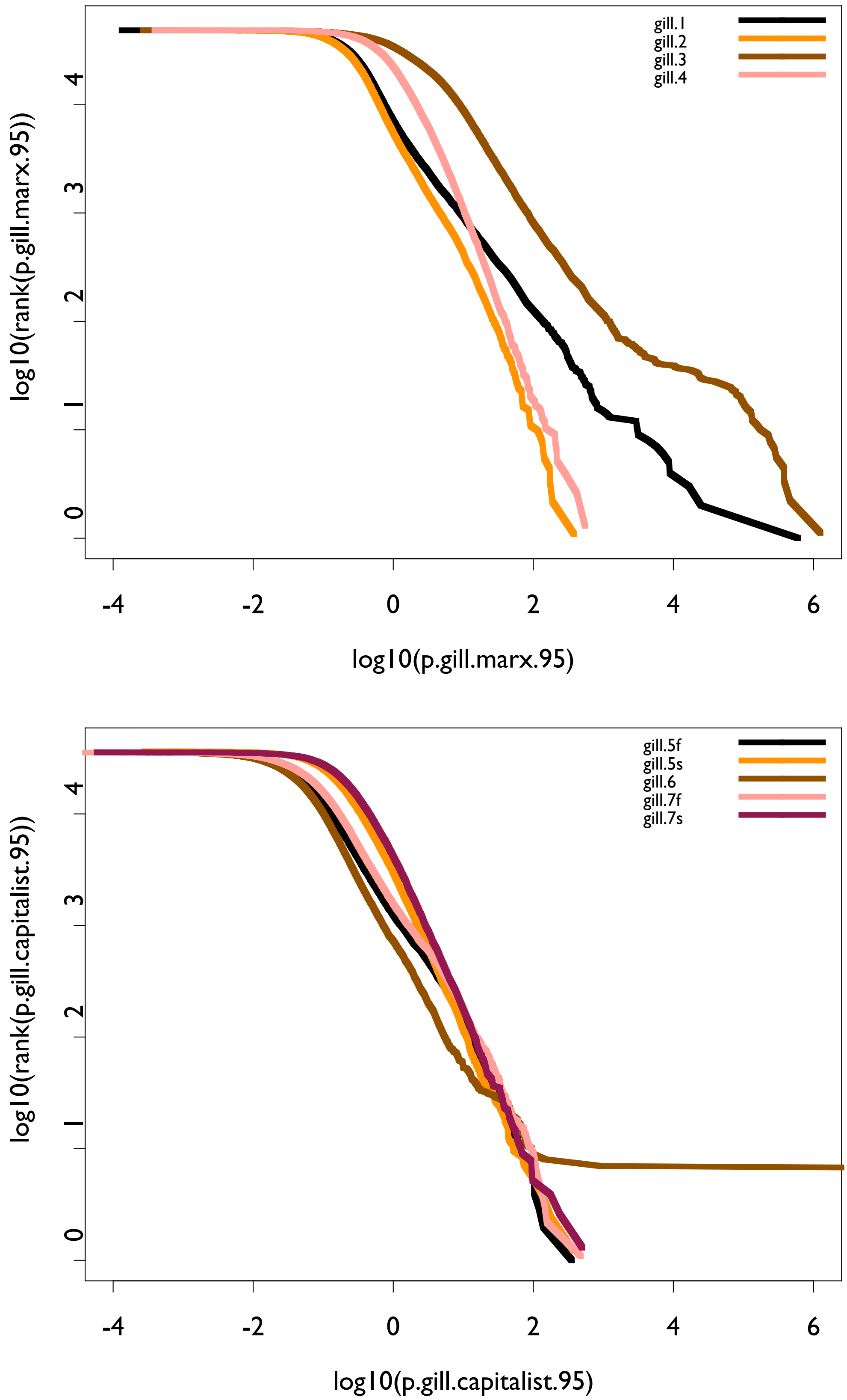


Figure5.11(a): Zipf plot analysis, unweighted Gillman profit rate measures, marxian and capitalist

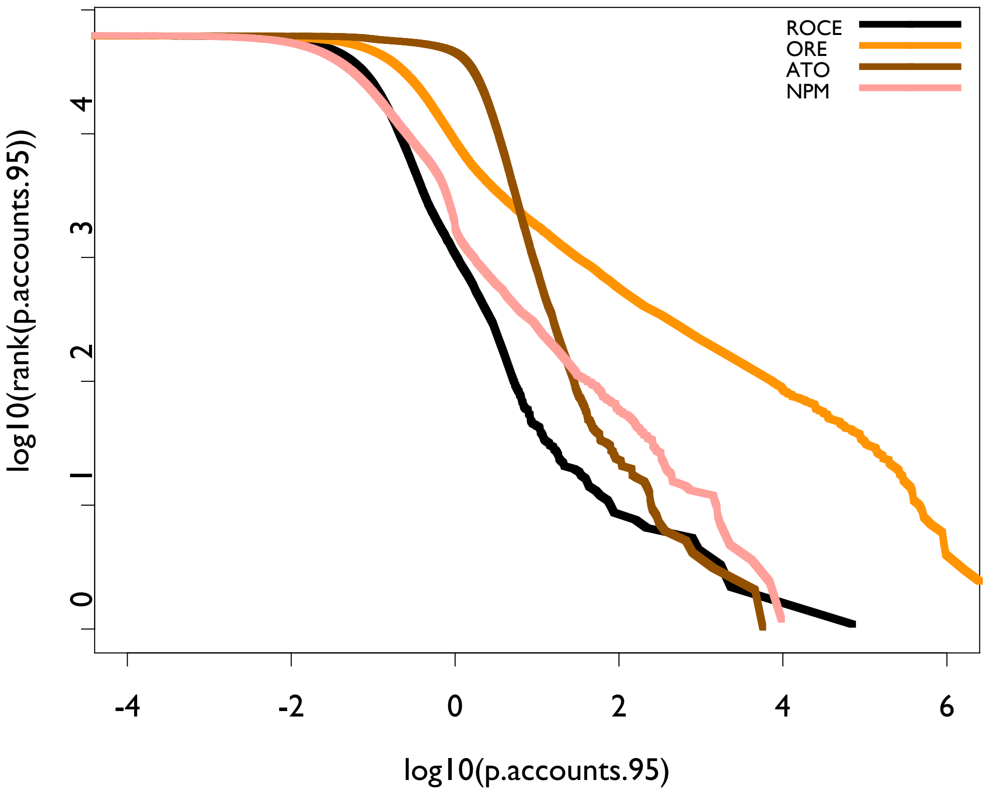
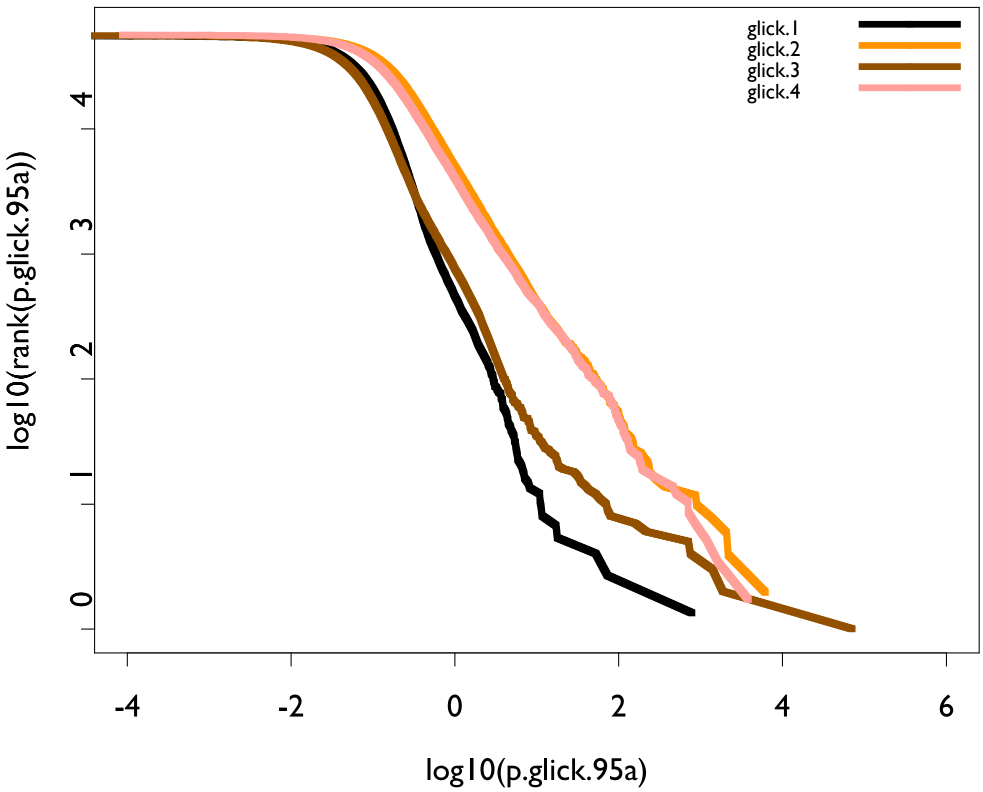


Figure5.11(b): Zipf plot analysis, unweighted accounting ratios



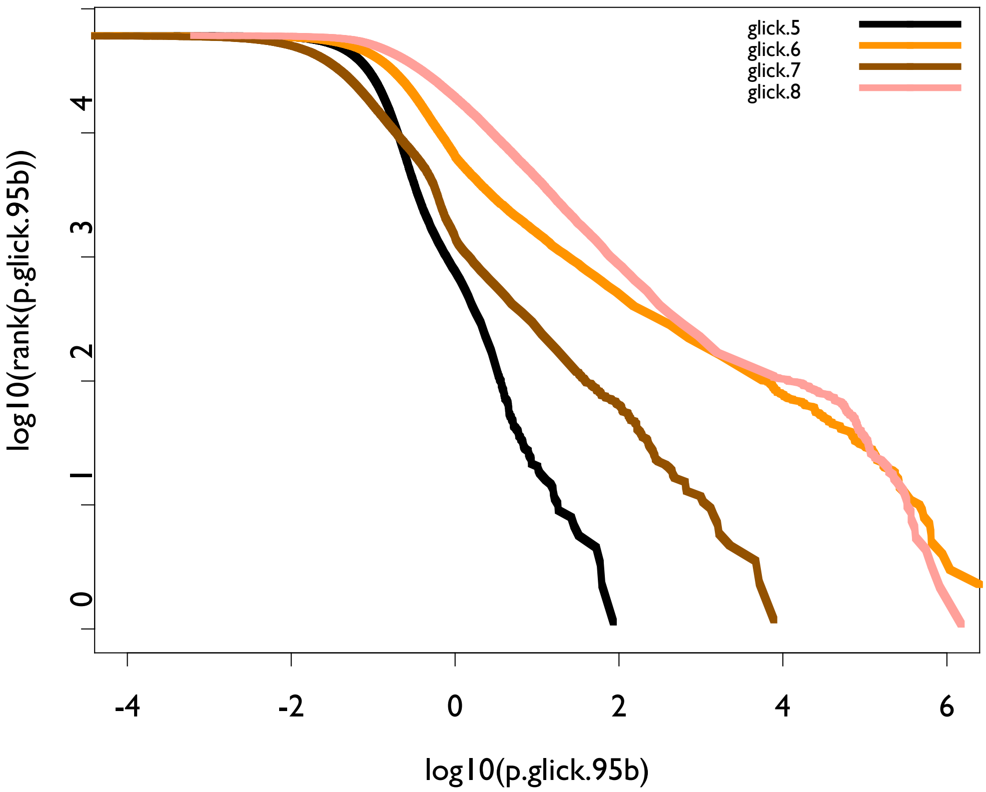


Figure5.11(a): Zipf plot analysis, unweighted Glick profit rate measures

In each plot in Figure 5.11 about $10^{1.6}$ observations, or around 40 per cent of the data, are in the power-law tail, although this proportion appears to be somewhat less for the Glick measures (about $10^{1.5}$ observations, or around 30 per cent of the data) – and much less for Gillman 2 and 4: around 10 per cent, as already seen. These results explain the very large skewness and kurtosis – in both conventional and *L*-moment forms – which we saw in Chapters Thee and Four.

However, even more striking here are the ranges covered by the tails of certain measures – six orders of magnitude in the cases of Gillman 1 and 3, ORE and Glick 7 and 8, and four orders in the cases of the remaining three accounting ratios and Glick 2 to 4 and Glick 6. These may be contrasted with the two orders of magnitude covered by the power law tails of Gillman 2 and 4, Gillman's capitalist measures (including Gillman 6, bar a handful of extreme observations) and Glick 1 and 5.

A number of intriguing patterns are evident in Figure 5.11, but to avoid the scope of the discussion becoming unmanageable we resist the temptation to comment on all but one. This is the fact that the plots of several measures show first a transition from exponential to power law form, and then a second transition to a shallower slope towards the extreme end of the tail. This is discernable among the accounting ratios and Glick 1 and 3.

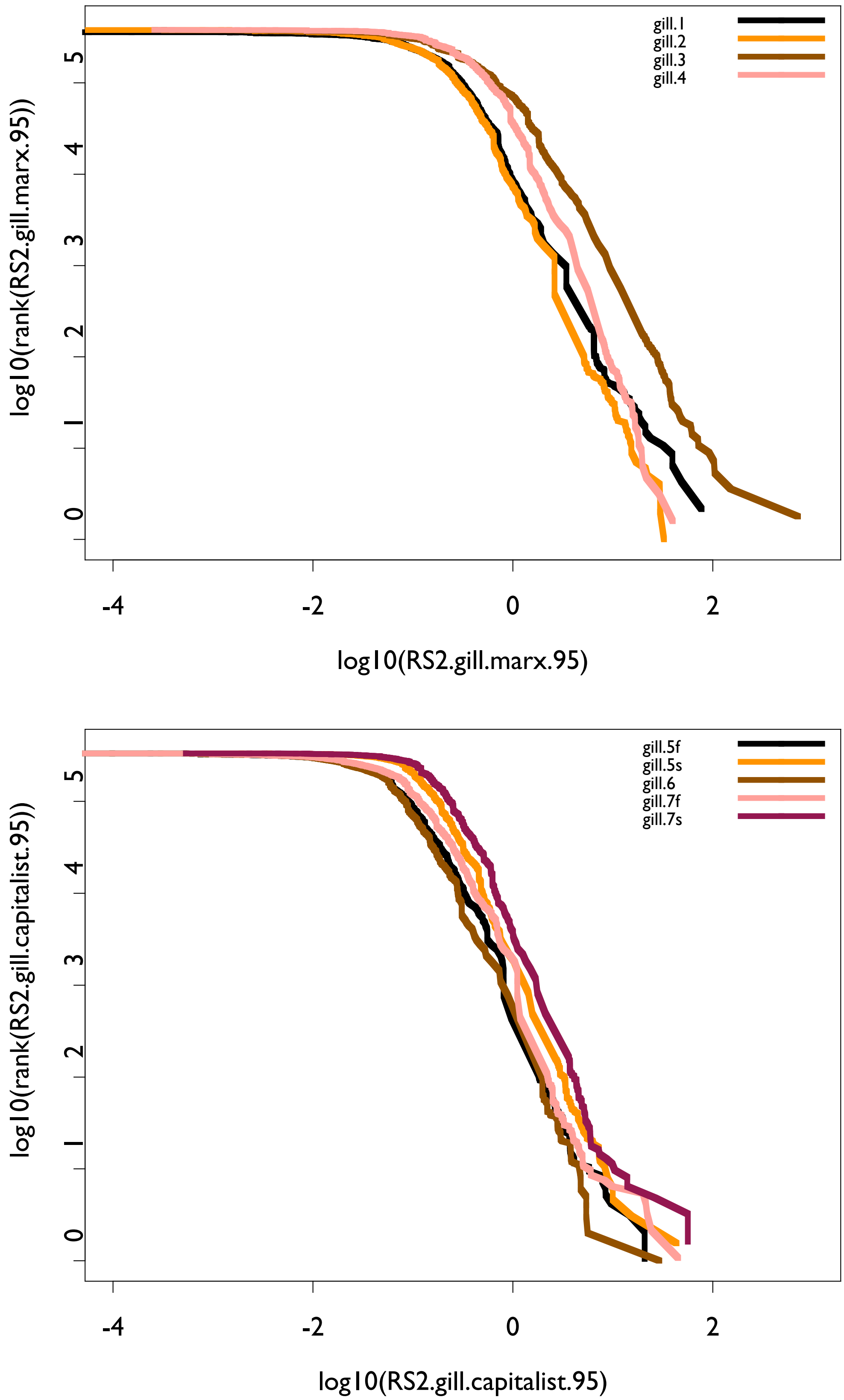


Figure5.12(a): Zipf plot analysis, size-weighted Gillman profit rate measures, marxian and capitalist

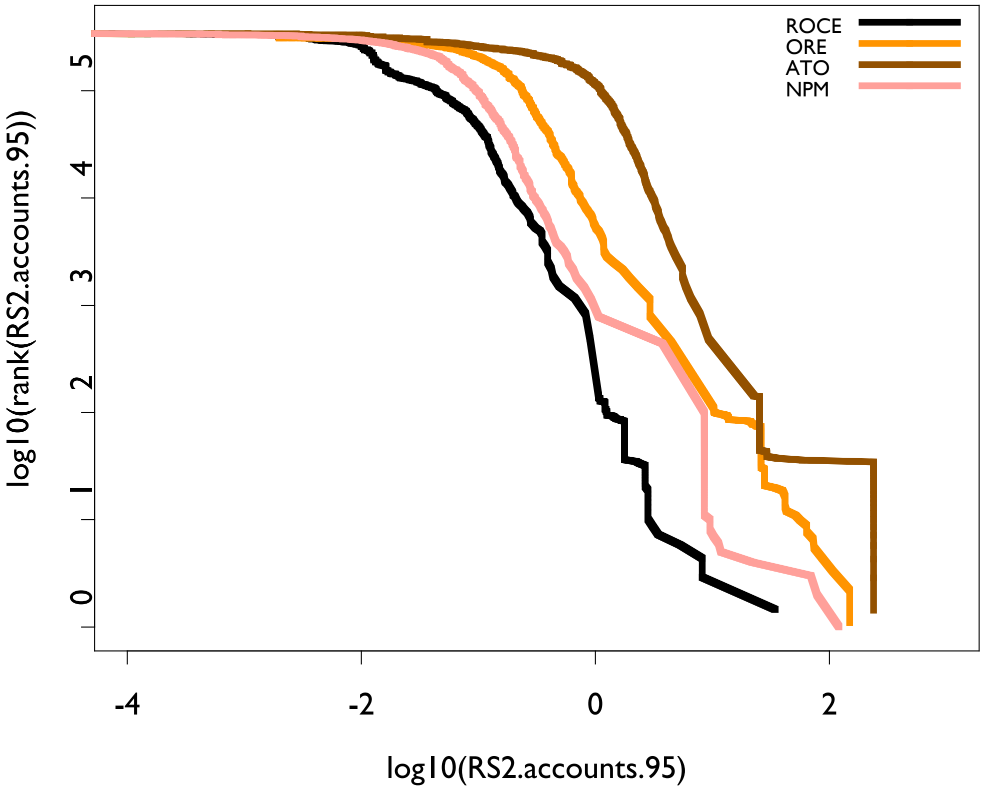


Figure5.12(b): Zipf plot analysis, size-weighted accounting ratios

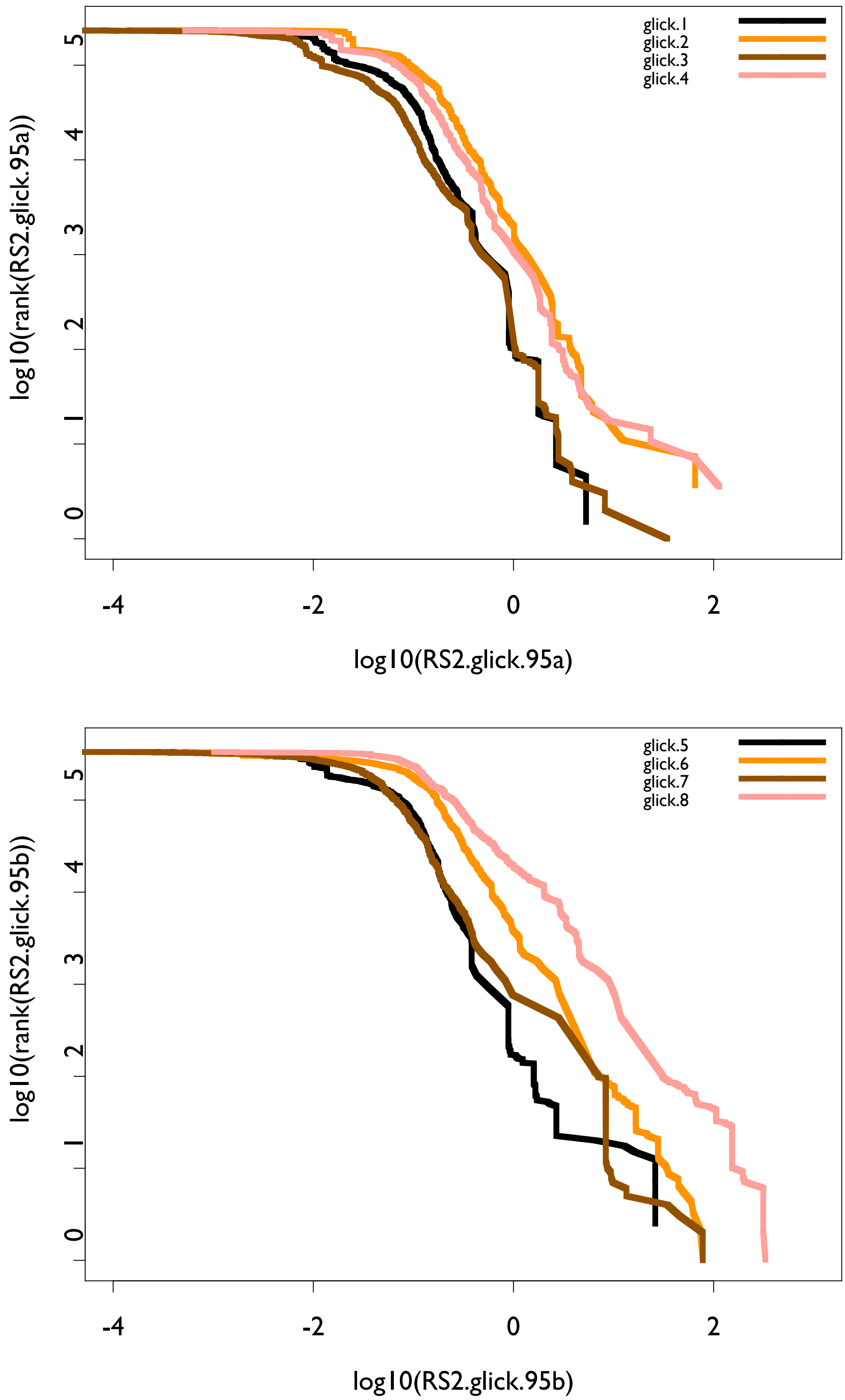


Figure5.12(c): Zipf plot analysis, size-weighted Glick profit rate measures

In Figure 5.12 we compare the weighted versions of our profit rate measures. Three points should be noticed here. Firstly, the RS2 weighting procedure reduces the extent of the tails of all measures and homogenises them, in terms of the number of orders of magnitude covered, as indicated by the steeper and now similar gradients of the power-law portions. (However, the differing lateral displacements of the various tails means that the absolute value of their ranges varies). Second, comparison with Figure 5.11 shows that the exponential part of the distribution has been changed, further undermining our hypothesis that the RS2 procedure might act as a fuzzy rejection scheme for eliminating contaminant observations of non-capitalist entities from the firm-level distribution. Third, the weighting process does not seem to reduce the proportion of the data which they cover: from around 10–17 per cent for the Marxian measures and the accounting ratios, to 40 per cent for Gillman's capitalist measures and for Glick's. The implication is that while there is a clear inverse association between size of firm and range of attainable profit rates, it is nonetheless the case that the proportion of *firms* associated with power-law tails is about the same as the proportion of *capital* which is so linked.

Finally, note that comparison of Figure 5.12 with Figure 5.11 reinforces the point that the weighting procedure has not had the effect of a fuzzy rejection scheme for contaminant observations. Hence we cannot now maintain that estimation of capital-level distributions is equivalent to estimation of company-level ones.

5.3.3 Annual variation in power-law tails

In discussing the tails of the size-weighted Gillman 4 measure, illustrated in Figure 5.10, we suggested that the annual variation in the power law tails, expressed in the varying slopes, explained the high variability in the estimated *L*-kurtosis of that measure's distribution. We now show that similar patterns of annual variation exist in respect of the size-weighted versions of other distributions. We confine our investigation of annual variations in the power-law tails to Gillman's Marxian measures, the four accounting ratios, and the four measures which Glick found to give the best evidence for his notion of gravitation of profit rates (Glick 1, 3, 5 and 6).

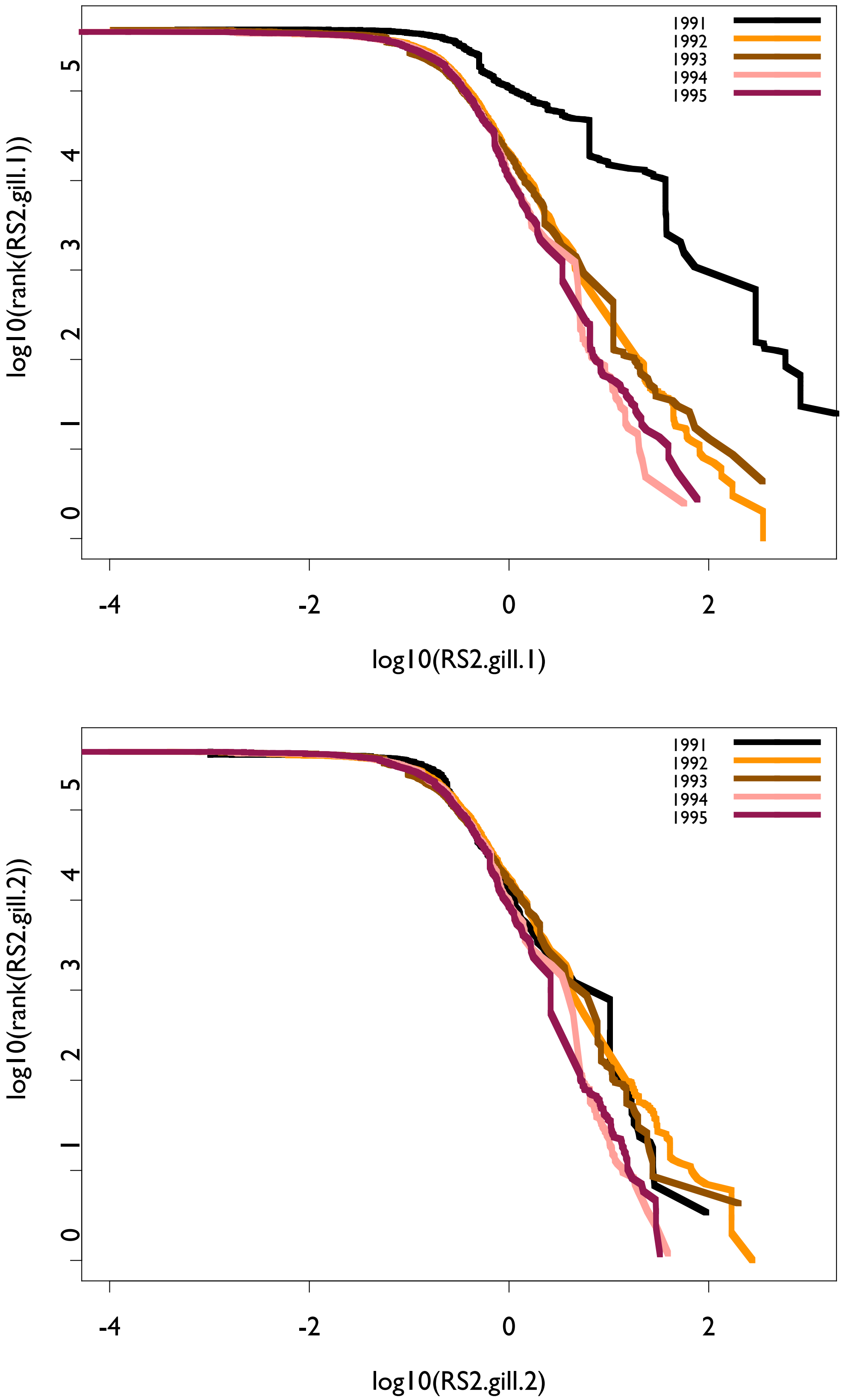


Figure5.13(a): Zipf plot analysis of annual variation; Gillman marxian measures

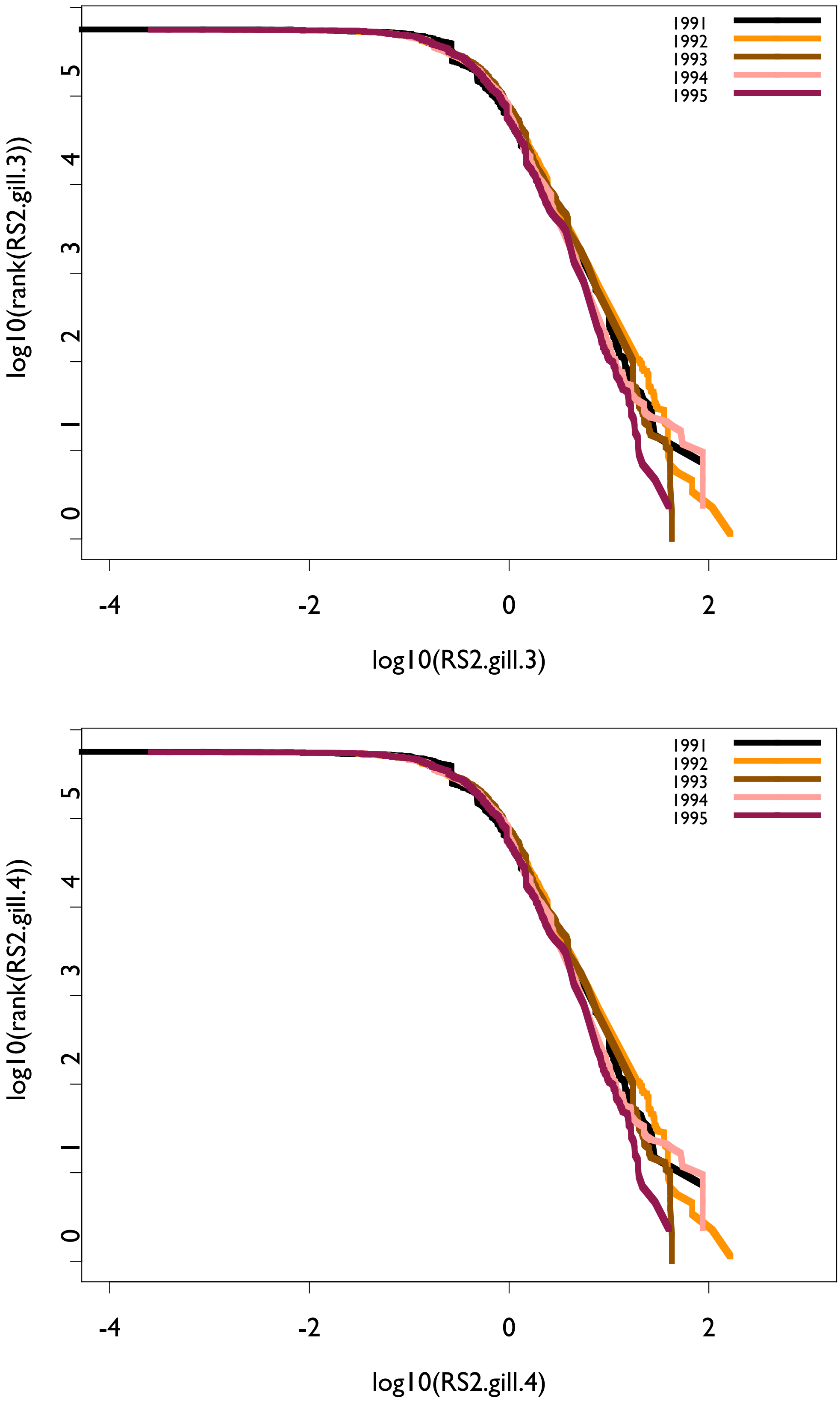


Figure5.13(b): Zipf plot analysis of annual variation; Gillman marxian measures

We begin with Gillman's four Marxian measures, illustrated in Figure 5.13. Very broadly, all these measures have tails whose total range is around two orders of magnitude: a range of profit rates, in percentage terms, from 100 per cent to 10,000 per cent. Variation in the exponential part of each measure's distribution is obscured by the log scaling, but we know from the right-hand panel of Figure 5.9 (which focused on the transition between the exponential and power law parts of the Gillman 4 distribution) that this variation is small in comparison with that of the power law section.

Thus the principal feature of all four measures is the annual variation in tail slope – hence in the value of the power-law exponent (sparsity of available data accounts for the anomalous appearance of the distribution of Gillman 1 for 1991). Thus the maxima of the tails have ranges of the order of about $10^{1.6}$ = 40 to $10^{2.2}$ = 160 for Gillman 4 (in percentage terms, profit rates from 4,000 to 16,000 per cent).

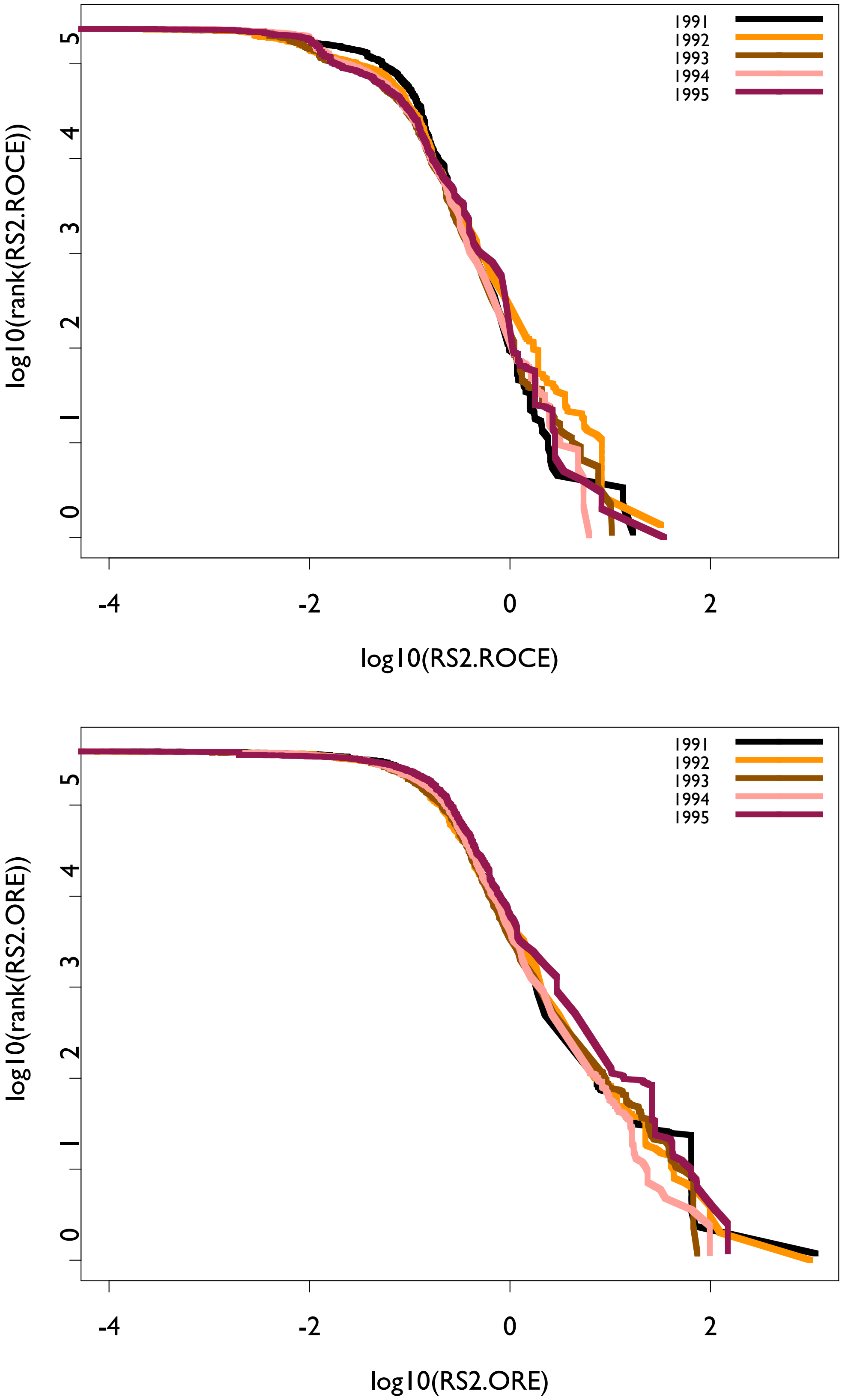
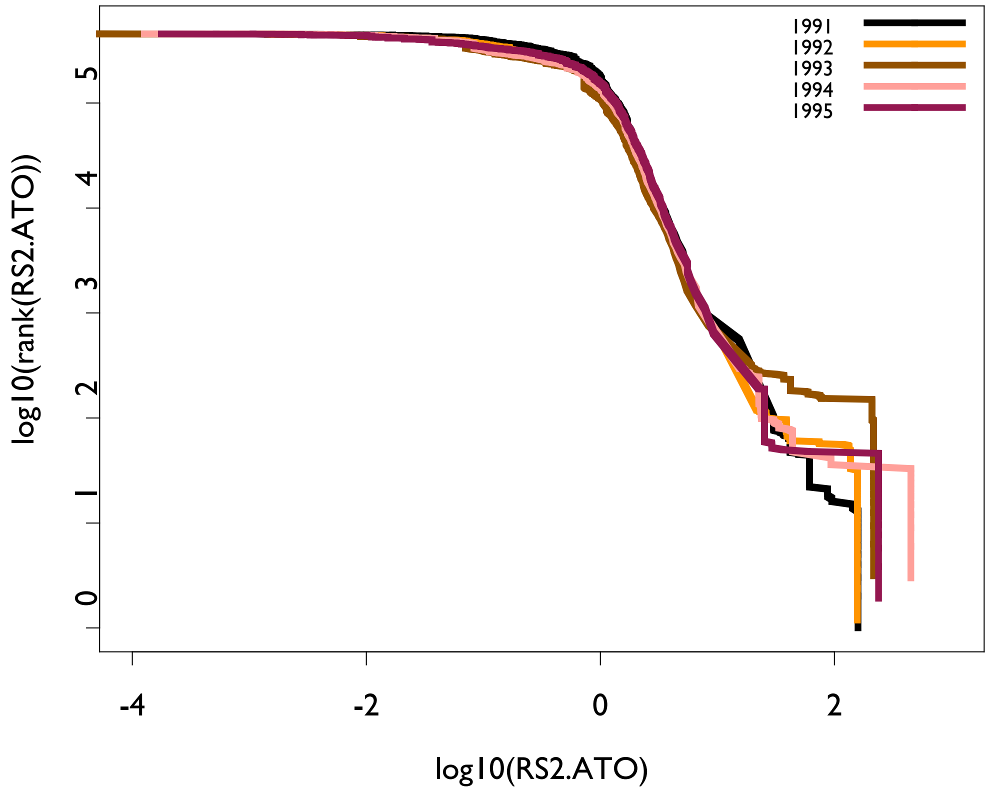


Figure5.14(a): Zipf plot analysis of annual variation; accounting ratios



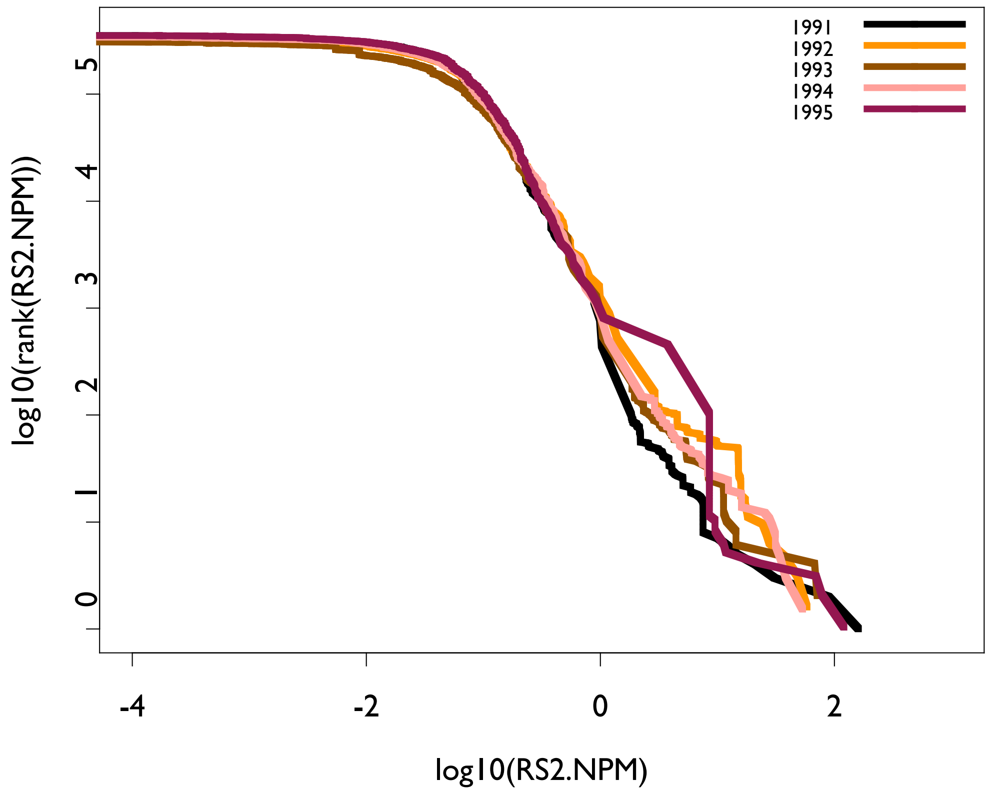


Figure5.14(b): Zipf plot analysis of annual variation; accounting ratios

In Figure 5.14 we show annual variations in the accounting ratios. As with the Gillman measures just examined, the range covered by the tails of these ratios is approximately two orders of magnitude. However, the proportion of the data included in the tails is greater, as already noted.

A distinctive feature here is the way in which the extreme tails of three of the four ratios exhibit either a change of gradient or a marked increase in variability, or both. These are the operating return on equity (ORE), the asset turnover ratio (ATO), and the net profit margin (NPM). (The exception here, ROCE, is also exceptional in showing much greater evidence of annual variation in the exponential part of the plot.)

Not too much importance should be placed on the apparent increase in annual variability in this part of the tails. To see this, note that our weighting procedure effectively produces an estimate of the distribution of the rate of return across units of capital (albeit one with spikes and clumps in the density, due to the clustering of capital in firms, some very large in relation to the total capital). The parts of the tails we are discussing are those between values of 10^0 and 10^3 on the vertical axis – in other words the 1,000 largest values. While these are of course rates of return achieved by companies (since that is the origin of the data), in principle they here represent the rates of return accruing to individual £1 units of capital, hence to the highest-earning £1,000 of the total invested in the corporate sector.[65] In practice; the approximate size of our concatenated RS2 samples lies between 70,000 data points (Glick 2) and 410,000 (Gill 6).

[65] About £929bn, in the case of the capital measure used to calculate Gillman 4.

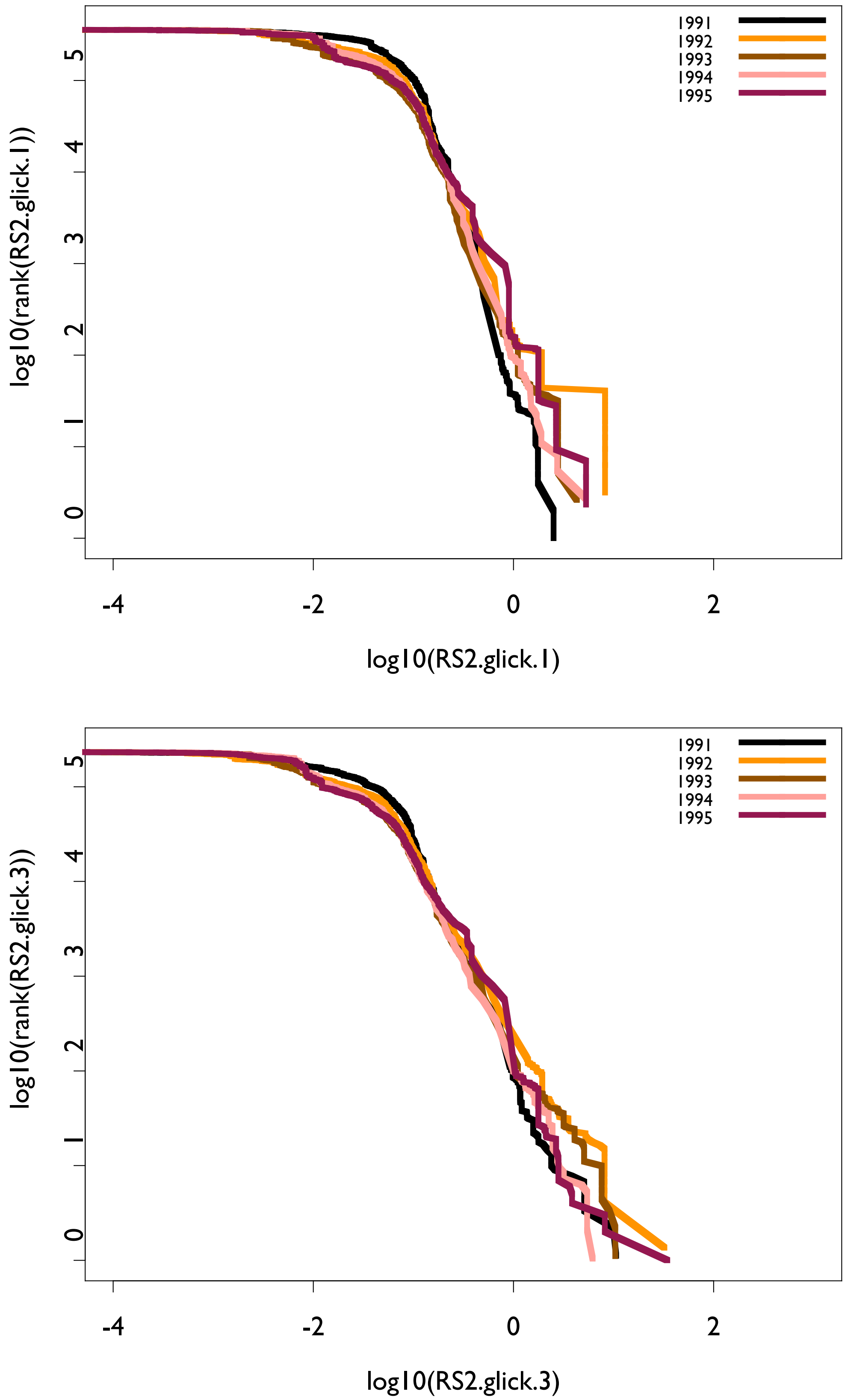


Figure5.15(a): Zipf plot analysis of annual variation; Glick measures

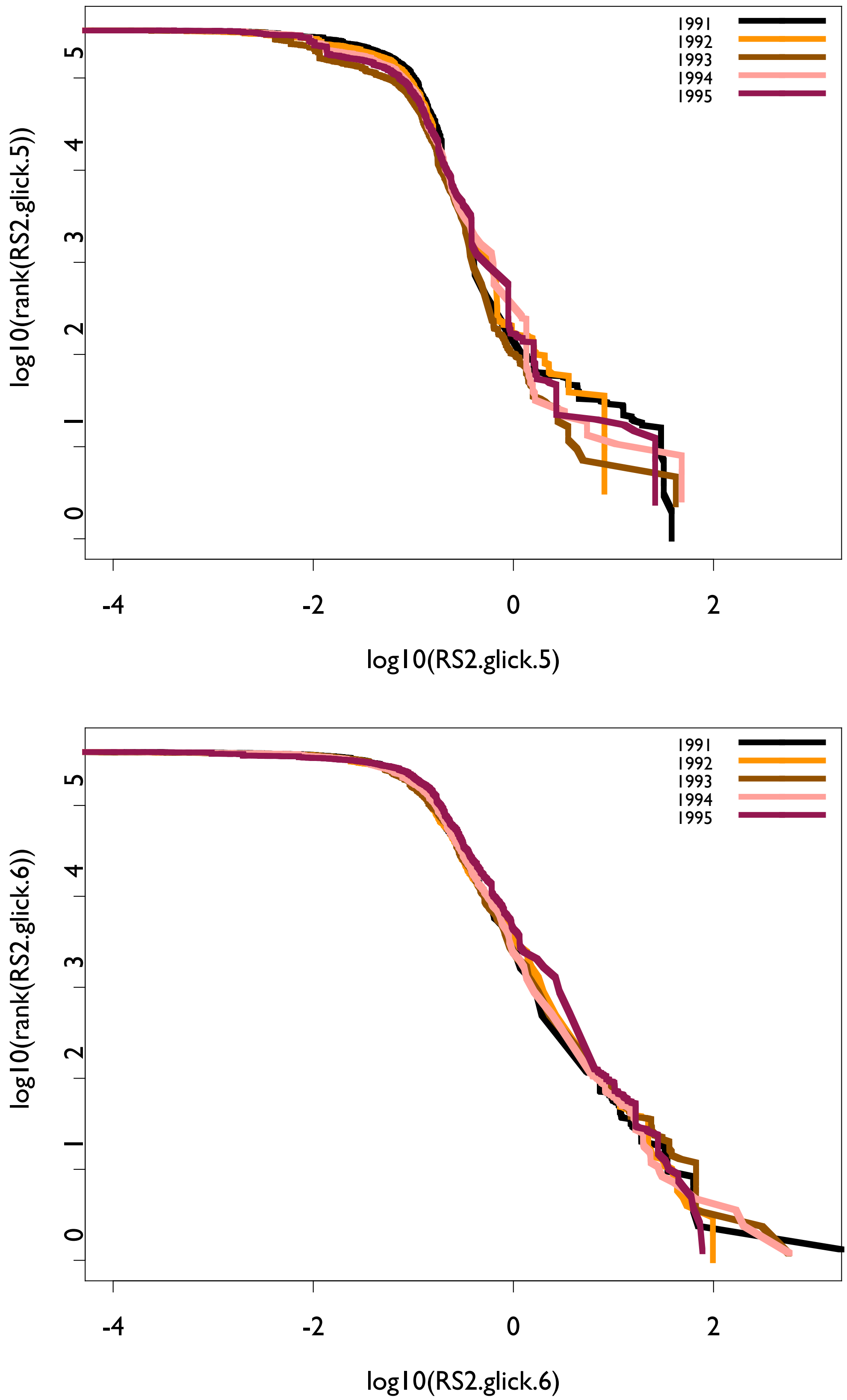


Figure5.15(b): Zipf plot analysis of annual variation; Glick measures

Our final set of Zipf plots, Figure 5.15, shows the tails of our selected Glick profit rate measures. First, we note that the extreme tails of two measures, Glick 5, and to a lesser extent Glick 6, display the same tendency to change of gradient or increased variability, or both, which we saw in the case of the accounting ratios. These three Glick measures (together with ROCE) also share another feature, subjectively greater variation in their exponential sections, compared to Gillman's Marxian measures. Finally, Glick 5 and 6 have tails with ranges somewhat greater than do other measures: perhaps as much as three orders of magnitude in the case of Glick 6 (but note that if one neglects the 100 largest observations in Glick 5, this measure actually has a shorter tail than most other measures, extending only over about one order of magnitude).

One feature of the plots of point clouds in *L*-skewness, *L*-kurtosis space which remains to be explained is the differences in the pattern of annual variation in skewness and kurtosis between Gillman's 'marxian' measures (together with ATO) and the remaining measures (excepting Glick 8). Broadly speaking the former group all fell within the band of Hosking and Wallis loci whereas the others lay well above it; with the exception of Gillman 1 the first group displayed less annual variation than the second, perhaps particularly so in respect of skewness.

In the case of Gillman 4 we suggested that the variation in the power law tails explained annual variation in kurtosis; in the case of the non-'marxian' measures we have instead to explain high – but only lightly-varying – kurtosis, combined with wide variations in skewness.

Our suggestion here is that the latter can be explained by variation in the balance between upper and lower tails, unremarked because our Zipf plot analysis has only looked at the former. Consider Table 5.2: this excerpts from Table 3.5 the information about the lower tail quantiles of the *unweighted* profit rate measures.

As can be seen, only the 'marxian' measures (and ATO) have five per cent quantiles which are non-negative (while Gillman 1 has a zero two per cent quantile and Gillman 4 almost does so). Our suggestion here is that both the large variations in skewness and the greater kurtosis are explained by non-symmetric annual variation in the tails of those profit rate measures where the lower tails contain a significant weight of negative values.

Table 5.2: tail quantiles (excerpted from Table 3.5)

PRM	0.0%	0.5%	1.0%	2.0%	5.0%
p.gill.1	-0.996	-0.176	-0.0626	0	0.0392
p.gill.2	-0.996	-0.281	-0.145	-0.0383	0.0312
p.gill.3	-2.23e5	-1.66	-0.372	-0.0573	0.0769
p.gill.4	-1.33e2	-0.346	-0.121	-4.14e-3	0.0851
p.gill.5s	-2.32e2	-3.78	-2.1	-1.11	-0.406
p.gill.5f	-2.1e2	-1.97	-0.96	-0.473	-0.167
p.gill.6	-7.33e2	-0.664	-0.454	-0.276	-0.113
p.gill.7s	-1.86e2	-3.02	-1.7	-0.861	-0.289
p.gill.7f	-1.58e2	-1.31	-0.702	-0.355	-0.117
p.ORE	-1.47e5	-9.62	-4.26	-1.94	-0.684
p.ROCE	-8.41e1	-0.799	-0.509	-0.305	-0.127
p.ATO	-0.0255	8.96e-3	0.0229	0.0445	0.0884
p.NPM	-4.16e3	-2.04	-0.828	-0.384	-0.119
p.glick.1	-8.41e1	-0.803	-0.52	-0.32	-0.137
p.glick.2	-9.01e21	-5.71	-2.88	-1.41	-0.496
p.glick.3	-8.46e1	-0.768	-0.474	-0.278	-0.115
p.glick.4	-9.01e21	-4.73	-2.46	-1.22	-0.419
p.glick.5	-8.44e1	-0.697	-0.424	-0.236	-0.0827
p.glick.6	-1.6e5	-8.7	-3.8	-1.76	-0.634
p.glick.7	-4.16e3	-1.98	-0.816	-0.364	-0.11
p.glick.8	-4.56e5	-2.42e1	-9.75	-3.81	-0.811

Note: shaded cells indicate non-negative quantiles

5.3.4 Summary of Zipf plot analysis

Although our analysis is broadly qualitative rather than quantitative, several conclusions appear to be well-supported. First, the high levels of L-skewness and L-kurtosis found in both weighted and unweighted versions of all the profit-rate measures examined are accounted for by power-law tails, a sign of which is the much steeper slopes of the weighted versions accompanied by reduction in the estimated L-moment ratios.

Second, the variations in *L*-skewness and *L*-kurtosis are explained by variation in the length and weight of these tails.

Third, the relatively low weight of the power law tails of the four Gillman Marxian measures accounts for the fact that our RS2 procedure for estimating *L*-moment ratios is more nearly successful here than in other cases. This is because these tails are not obviously shorter, in the sense of the range of values covered, than those of the other types of measure (although it should be noted that Glick 5, which on one view has shorter tails than the Marxian measures, is the only other measure apart from ATO for which any years' estimates fall within the band of Hosking and Wallis loci).

In the case of Gillman 4, in particular, the upper tail is light enough that it does not significantly interfere with the *L*-moment ratios' ability to identify a possible log normal distribution.

5.4 Conclusion

In this chapter we began by showing that Farjoun and Machover's probabilistic political economy is entirely in the spirit of Marx's own approach – indeed, we think it little exaggeration to say that their work can best be viewed as a proposal for filling in technical details which the statistical science of Marx's day would have struggled to provide.[66]

We then applied the RS2-L methods developed in Chapter Four in an attempt to test the hypotheses about profit rate distributions put forward by Gibrat and Farjoun and Machover. The results are best described as an encouraging near-miss; Gillman 4, the measure we regard as instantiating the profit rate definition to which Farjoun and Machover's hypothesis relates, was tentatively identified as log normal. As far as it goes, this

---

[66] In particular, Maxwell's kinetic theory of gases, the first contribution to statistical mechanics, is an almost exact contemporary of *Capital*: Maxwell first publication in the area was in 1860, and he produced a fully-worked-out statement in 1865, two years before the publication of Volume I.

is direct support for Gibrat, and also weak confirmation of Farjoun and Machover, in view of the fact that the lognormal is a special case of the four-parameter gamma distribution.

An investigation of the sampling properties of the RS2-L procedure suggested that the unsatisfying aspects of the results for Gillman 4 were likely to be the outcome of extensive tails still remaining after the size-weighted sampling process. Removal of discordant samples could be expected to produce more convincing estimates of *L*-moment ratios.

This was reinforced by the Zipf plot analysis in section 5.4, which showed that both weighted and unweighted versions of this measure have power law tails appended to distributions whose main mass appears to be exponential.

Indeed, all the profit rate measures tested have such power law tails, a fact not previously demonstrated in the literature.

The partial success in the case of Gillman 4 is due to its power law tail being relatively light. Since this profit rate measure is the one we have identified as instantiating Farjoun and Machover's definition of the profit rate relevant to their hypothesis, we have not succeeded in rejecting the hypothesis that other definitions of the rate of profit might also lead to gamma distributions.