Talk at The Foundation for Science and Technology, 10th June, 2009

Did over-reliance on financial models for risk assessment create the financial crisis?[1]

David J. Hand
Imperial College, London

Baron Justus von Liebig said '*We are too much accustomed to attribute to a single cause that which is the product of several.*' It seems to me that this is perfectly reasonable. In order to cope with the awesome complexity of the natural world, and the corresponding complexity of the artificial globalized society and economy we have constructed, in order that our mere human brains can cope with such things, we naturally seek to reduce them to simple components. And then we try to identify the most important of these components and focus our attention on this.

An illustration of this single cause attribution in the current economic crisis is given by the case of Joe Cassano. Cassano was head of the financial products division of AIG, and was responsible for underwriting several hundred billions of dollars worth of debt in the form of credit default swaps. Jackie Speier, a member of the US house of representatives, is quoted as saying that Cassano is "almost single-handedly ... responsible for bringing AIG down and by reference the econonomy of [the US]".

But the fact is that, as Justus von Liebig noted, such single cause attribution is incorrect. The truth is more complex than this. In the present economic crisis, a variety of contributory factors came together, to act synergistically, although many of these have been singled out for special attention. This evening's meeting focuses on just one of them: the role of mathematics and mathematical models in the crisis.

The Turner review (Turner, 2009) gives considerable contextual detail, and enumerates the events building up to the crisis. From our perspective tonight, what is important is that these events produced an environment stimulating financial innovation and a huge growth in securitised credit instruments and derivatives. And this is where the mathematics comes in: the result has been dramatically increased mathematical complexity in the tools for modelling and pricing financial risk.

The question I was asked to address uses the word 'over-reliance'. The Turner report is equally careful in its choice of words. Neither ask if the mathematics was incorrect, or even if it was inappropriate, but merely raise the question of whether people were unwise in the extent to which they relied on it. Noting that, I nevertheless think it will clarify things if we put the question in a broader context. So, we can ask several rather distinct questions.

First, was the mathematics wrong? Whether wrong or not, we can also ask, was the mathematics unrealistic? And if it was unrealistic we can go on to ask if it was based on false premises, or was inappropriate in some way.

[1] The presentation follows the script.

We can also ask if any limitations of the mathematics were not communicated properly. And then, if so, we have to note that communication is a two way process, and ask why the failure in communication occurred. Was it that someone didn't raise appropriate warnings? Or was it that warnings were sounded, but others didn't listen? And then, if people didn't listen, why not? Was it that they couldn't understand the warnings being given, or were they unwilling to listen?

While I don't intend to discuss each of these questions in detail, I will make some comments on some of them.

First, was the mathematics wrong?

The short answer to this is 'no'. Mathematics is concerned with deducing the consequences of a given set of initial premises. While it is possible to make mistakes in the deductive process, in the present case, since so many people were using and had checked the deductions it is essentially inconceivable that the mathematics was wrong.

But perhaps that's a very purist view of mathematics. To a layman, if a piece of mathematics ends up drawing incorrect conclusions it may well be a quibble whether the mathematics per se is wrong or whether the problems lie elsewhere. To take an example from my own work. I have recently demonstrated a fundamental flaw in a particular index used very widely for measuring the performance of risk scorecards in the personal banking sector (Hand, 2009). While the deductive mathematics underlying this measure is fine, what is not so fine are the assumptions on which the mathematics is built. Buried deep in these assumptions is one which means that different risk scorecards are evaluated using different measures. To a layman, who does not distinguish between the underlying assumptions and the deductive process, the consequence is simply that one cannot apply this mathematically derived measure: that the mathematics is 'wrong'.

Although, in this example, there is nothing incorrect with the mathematics per se, the mathematics is somehow missing the point. Missing the point is a consequence of basing the analysis on false premises. The notion that sophisticated securitised credit instruments improve financial stability appears to have been a fundamental premise, a core belief, but it appears to have little or no supporting empirical evidence. Similarly, there appears to have been an assumption that natural selection in the financial markets would mean that innovations, and in particular mathematical innovations, would be beneficial, since those that did not work would be selected out. If something failed, people would not use it. In fact, however, anyone familiar with evolutionary processes will know that evolution can lead in unpredictable directions, especially in complex environments. In any case, it is important to be clear what we mean here. By 'beneficial' we mean beneficial to the economy, and the people in it, in some sense. But evolution is blind. Evolution does not tend towards a particular prespecified objective.

A familiar premise of the mathematical models was that assets could be sold rapidly and easily if necessary. But, as we saw some years ago, with the collapse of Long Term Capital Management, this is not true if everyone tries to do it simultaneously.

We have an unmeasured liquidity risk, so that one of the premises of the mathematical models fails.

At a lower level, there are also criticisms of things such as assumed distributional forms, and independence of events and players. Normality is well-known not to apply here. In fact, as statisticians know very well: *normal distributions do not occur in nature*. A 1989 paper in *Psychological Bulletin* (Micceri, 1989) has the provocative title *The Unicorn, the Normal Curve, and Other Improbable Creatures*. Of course, statisticians also know that certain kinds of statistical techniques are robust to non-normality - such as distributions of the means of samples. But they also know this is not the case for measures describing the size of the tails of distributions, which is often of primary concern in risk evaluation.

The eminent statistician George Box said that 'all models are wrong, but some are useful' (Box, 1979). In particular, this means that one should always have a healthy scepticism about models. Furthermore this scepticism should be greater for some usages: consideration of tail areas of distributions should engender more scepticism than results based on the central limit theorem. To put it bluntly, one should avoid the hubris of assuming that one's models are correct. One should always allow for model uncertainty when deciding what capital reserves to keep, what risks to take. While the model might tell us how risky a venture appears to be, and how much it seems that we need to keep in reserve to allow for that eventuality, we should always ask *what if the model is wrong*?

The late Leo Breiman also had something to say about this. He said 'when a model is fit to data to draw quantitative conclusions ... the conclusions are about the model's mechanism, and not about nature's mechanism ... It follows that if the model is a poor emulation of nature, the conclusions may be wrong.' (Breiman, 2001)

My own view here is that we are really trying to model human behaviour. And humans can be unpredictable. They can even be perverse.

To drive home the dangers and the difficulties of building models which adequately reflect human behaviour, I want to describe a very simple example from my own research. Much of my work is concerned with building statistical models for the retail banking sector: that concerned with credit cards, personal loans, car finance, individual current accounts, mortgages, and so on. The last, of course, is particularly relevant since, the subprime crisis was certainly one of the precipitatory factors. The illustration, however, concerns how people behave when using credit cards in petrol stations, and, in particular, the shape of the distribution of amount spent per transaction, as shown in the card transaction data. Full details of this analysis are given in Hand and Blunt (2001).

One might predict that the shape of the distribution would be roughly normal, but with a slight right skew since the values can only be positive. And, broadly speaking, this is correct. However, what is unlikely to be predicted is a number of very pronounced spikes in the distribution. The data set is very large, so any spikes represent real underlying behavioural aspects - they are not random fluctuations.

Investigation soon reveals an explanation for some of these spikes: they arise as a consequence of marketing initiatives by the petrol companies, incentivising customers to spend more. Others, however, do not have so ready an explanation.

In particular, close examination of the histogram of transaction amounts shows pronounced spikes at values ending in £5 and £10. It seems that many people, even though paying by credit card, choose to spend round numbers of pounds. Supporting evidence for this explanation is given by the existence of slight tails to the right, but not to the left, of these values: it is as if people were aiming at £20 (for example) and occasionally overshot by accident.

Continuing the analysis, even closer examination shows that there are also spikes, albeit lower, at other whole numbers of pounds. For example, the histogram cells at £13 and £16 are substantially higher than cells pennies either side of these values. People clearly prefer to spend whole numbers of pounds.

Hand and Blunt (2001) explore the data in more depth. The deliberate choice of round numbers does not stop at whole numbers of pounds. There are also spikes at 50p, and also at 25p and 75p, and also at other values ending in 5p and 10p, all progressively lower.

The point of this little example is to show that data describing human behaviour, even apparently simple behaviour, can conceal unimagined complexities. When dealing with human beings we are not merely dealing with intrinsic randomness - as when we study the intrinsic randomness in quantum physics. When we study human beings we certainly have intrinsic randomness, but we also have other things to contend with - human motivations, intransigence, greed, and so on. Electrons may have their uncertainties, but they are not greedy.

That brings me to another of the questions mentioned above. Was it in fact *inappropriate* mathematics?

Models such as pricing models are fine, in isolation, and at a low level. But difficulties can arise when they are put together, and embedded in a larger system. In such a situation a larger scale model is needed, a model of the entire complex system, and not of merely a tiny part of it. An econometric model, in fact.

I suppose a very concrete example of this sort of limitation is in automated trading systems, which all react the same way to given market conditions - as we discovered in 1987. The correlation between their behaviour induces a massive swing in one direction or another - a run on a stock or a drying up of liquidity. The famous Prisoner, facing his Dilemma, would doubtless have something to say here.

I think a key issue in all of this is *communication*. If there was an over-reliance on mathematical models it was the reliance placed on them by the higher echelons of management. Perhaps the phrase ‘naive belief’ might be better than over-reliance.

A few weeks ago the *Financial Times* contacted me, asking me to comment on a criticism of the mathematicians who had developed the models, namely the suggestion that it it was the mathematicians’ fault, because they had been unable to

communicate the risks to senior management. Their fault that they had, in fact, developed models which senior management could not understand. This seems an extraordinary position to me. It suggests that, despite not understanding the models and their implications, the managers were happy to act on the basis of the recommendations deriving from them. As I wrote in an article solicited by the London Mathematical Society, if I jumped into the cockpit of a Boeing 747, and crashed it because I didn't know how to fly it, you would hardly blame Joe Sutter, the 747 production chief.

It appears that Joe Cassano was put in charge of the AIG financial markets operation after it was up and running, and that he inherited the mathematical drivers without properly understanding them. If this is true, one might blame him for agreeing to run something he did not understand, and also blame those who appointed him for not grasping what was needed to run such an operation. The only defence I can think of, and poor one at that, is that the mathematical models were developed after the senior bankers had begun their careers and were already in senior posts. Since the half life of material learned at University is five years (a ball-park figure: clearly it varies between disciplines), this at least explains, even though it does not justify, why they didn't understand what they were doing.

With this breakdown in communication in mind, one might ask, were managers in fact given warnings on which they failed to act? Such a scenario is not at all far-fetched. Harry Markopolos had been trying to raise concerns about Bernard Madoff since 1999. In 2005 he sent a report to the SEC, and the SEC made a cursory examination of Madoff and declared his activities legitimate.

More generally, if a given strategy appears to be earning large sums of money for your bank, and in particular large sums of money for you, then you might not be inclined to look too rigorously at criticisms of it. This is, perhaps, a natural human trait: along with greed, an unwillingness to believe bad news, and a tendency to follow the herd.

There have been many financial crises in the past. One has to ask whether the causes of this one are different in kind. It is certainly true that mathematical financial innovations played a role. It is also true that, had people been more aggressively warned about their limitations, arising from the premises on which they were based, and had they listened and taken those warnings on board, then things might well have turned out differently. But in order for that to happen, many other things would also have had to have been done differently. It seems to me that the bottom line is that, unless there is an incentive for people to behave differently, then greed will trump other factors. Fraud illustrates this, but there are also other structural incentive issues. For example:

- the fact of hedge fund managers taking a percentage of profit in good years, but not of loss in bad;
- the fact that if you tried to raise concerns you might be ignored;
- the possibility of the risk rating agencies benefiting if they gave good ratings;

- the fact that employees of regulatory authorities might subsequently want to work for the banks on huge salaries, and so would often be keen to maintain good relations;
- and the risk that moral hazard is built into the system, with a perception that people didn't need to worry about the premises because if things went wrong they would be baled out.

At bottom, can I really argue that one cannot blame the mathematicians? They built and applied the tools, so surely they should share the responsibility? This is, of course, well-worn ethical ground, having been covered in other contexts, such as the relation between nuclear physicists and the atomic bomb. My own view is that it is nonsensical to say that everyone has to share the blame. In a mugging, who bears the responsibility: the man who wields the knife, the owner of the cutlery factory which made it, the receptionist at the entrance of the cutlery factory?

I think it would be a terribly retrograde step if we experienced a backlash against quantitative tools. Analogous quantitative tools have revolutionised so many other aspects of banking, of customer relations, and of life in general, and have had an immense impact for the good. But the mathematical tools have to be used in a proper context. Putting the quants in a back room, instructing them to work their mathematical magic, and then blindly applying the results to the outside world without considering the wider implications is a recipe for disaster. And we are now consuming the product of that recipe.

References

Box G.E.P (1979) *Robustness in the Strategy of Scientific Model Building*. Technical Report, Madison Mathematics Research Center, Wisconsin University.

Breiman L. (2001) Statistical modeling: the two cultures. *Statistical Science*, **16**, 199-215.

Hand D.J. (2009) Measuring classifier performance: a coherent alternative to the area under the ROC curve. To appear in *Machine Learning*.

Hand D.J. and Blunt G. (2001) Prospecting for gems in credit card data. *IMA Journal of Management Mathematics*, **12**, 173-200.

Micceri T. (1989) The Unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, **105**, 156-166.

Turner A. (2009) *The Turner Review: a regulatory response to the global banking crisis*. Pub. ref. 003289, Financial Services Authority.

Did over-reliance on financial models for risk assessment create the financial crisis?

David J. Hand
Imperial College, London

10 June 2009

Was the mathematics wrong?

Was the mathematics unrealistic?

If unrealistic

- was it based on false premises?
- or was it inappropriate ?

Were the limitations of the mathematics not *communicated* properly?

Communication is a two way process

Were appropriate warnings not given?

Or were warnings not heeded?

And if not heeded, why not?

- were the warnings not understood?
- or were people unwilling to listen?

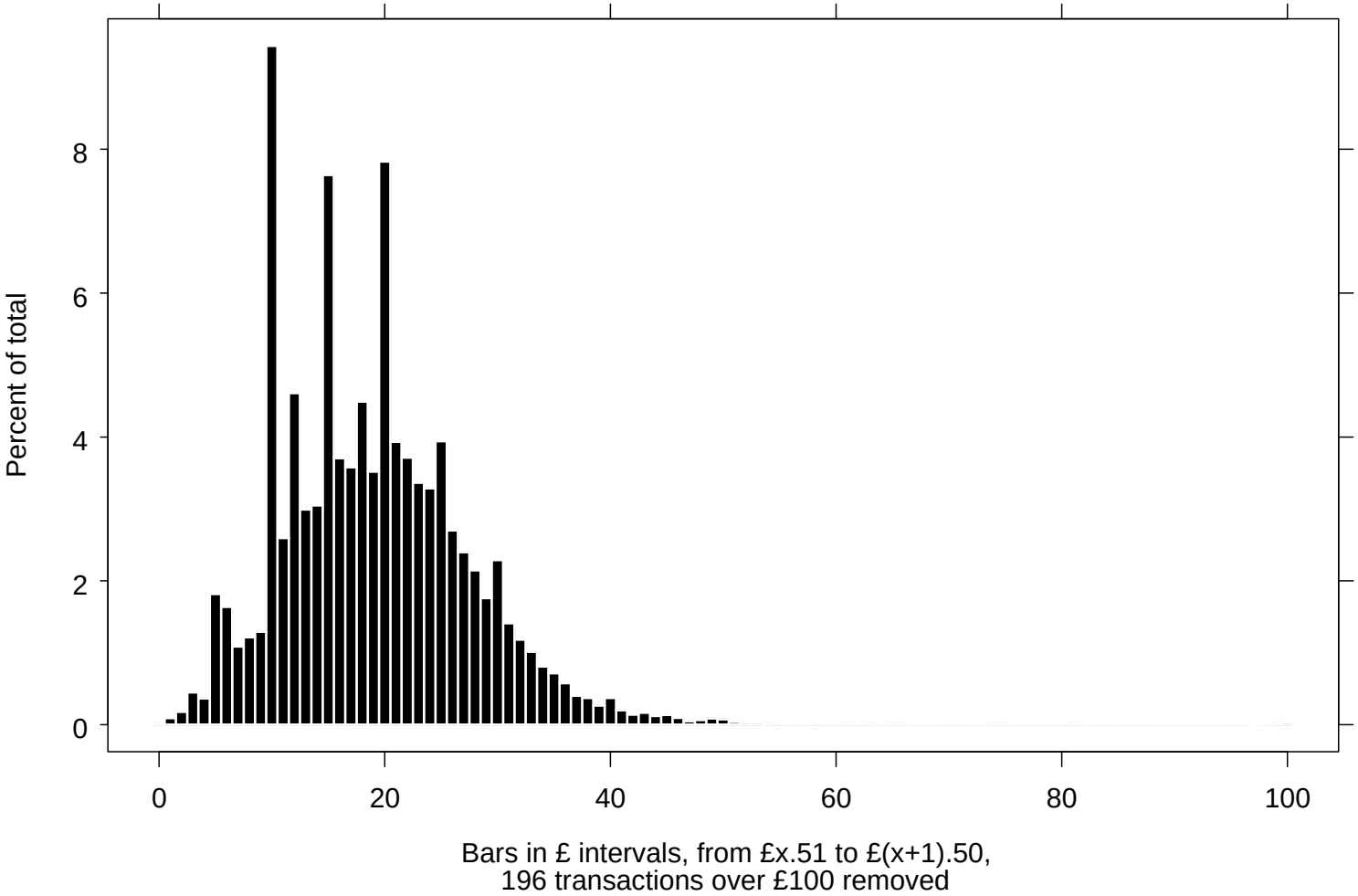
Was the mathematics wrong?

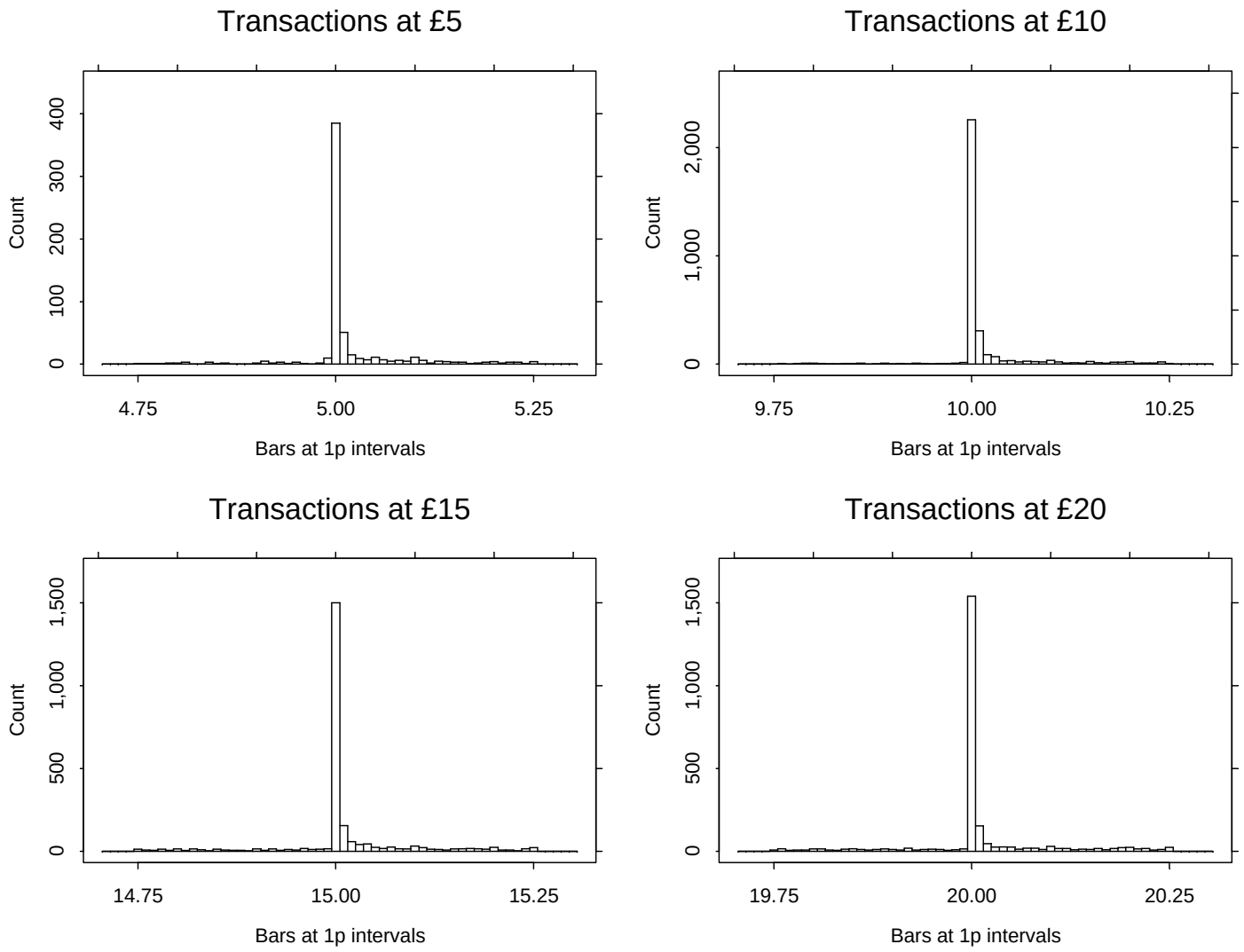
No

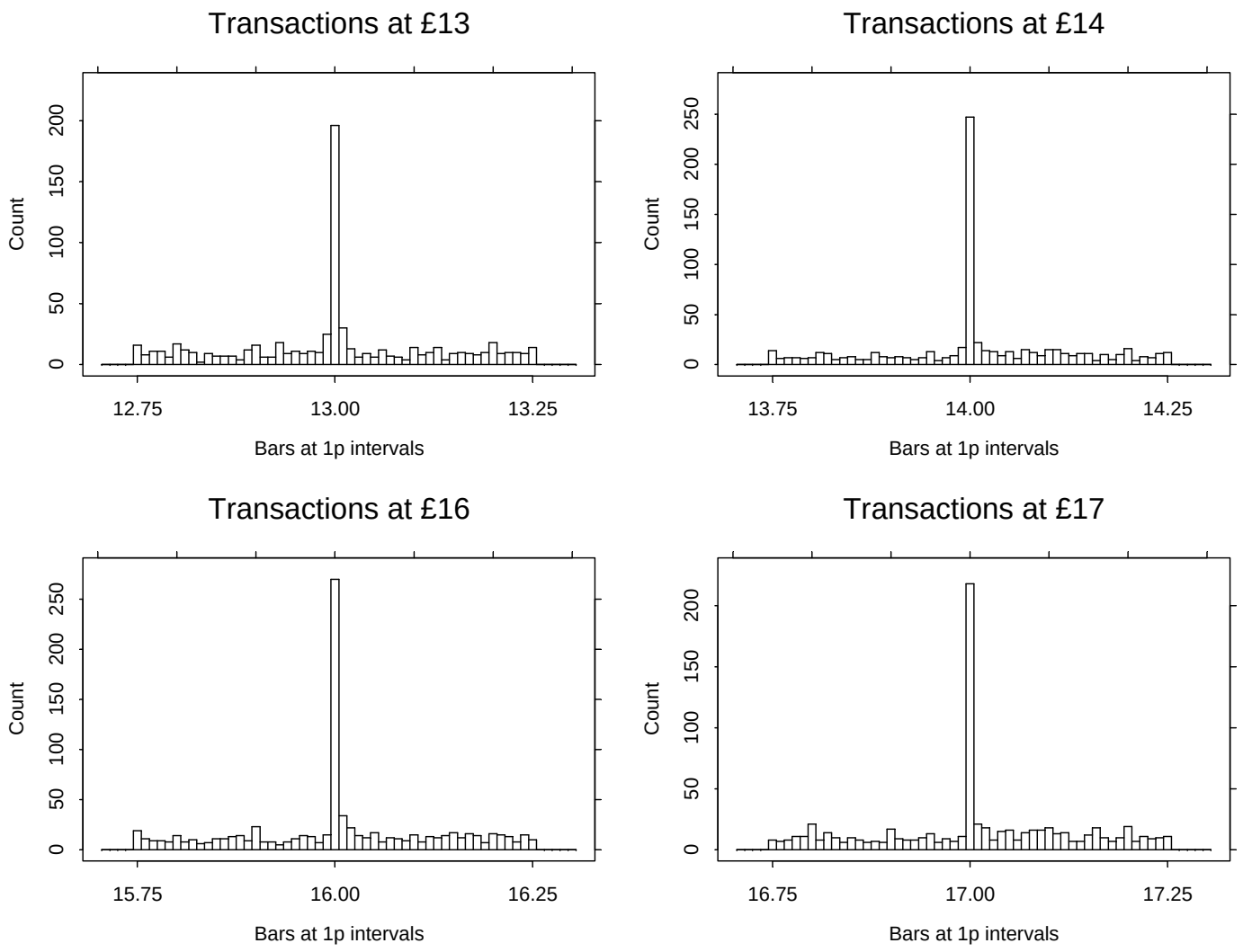
as a piece of deduction, the mathematics is correct

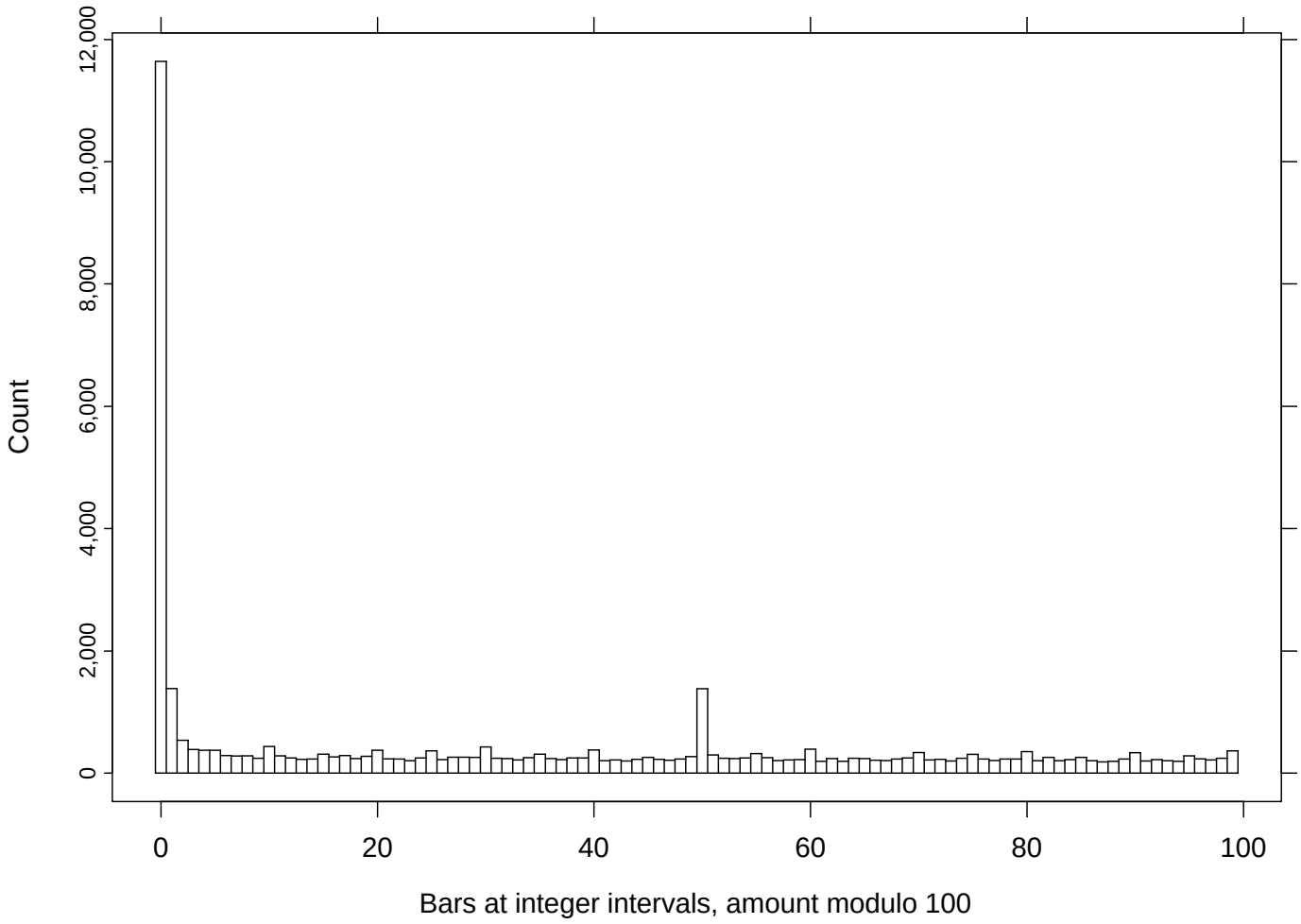
A purist view?

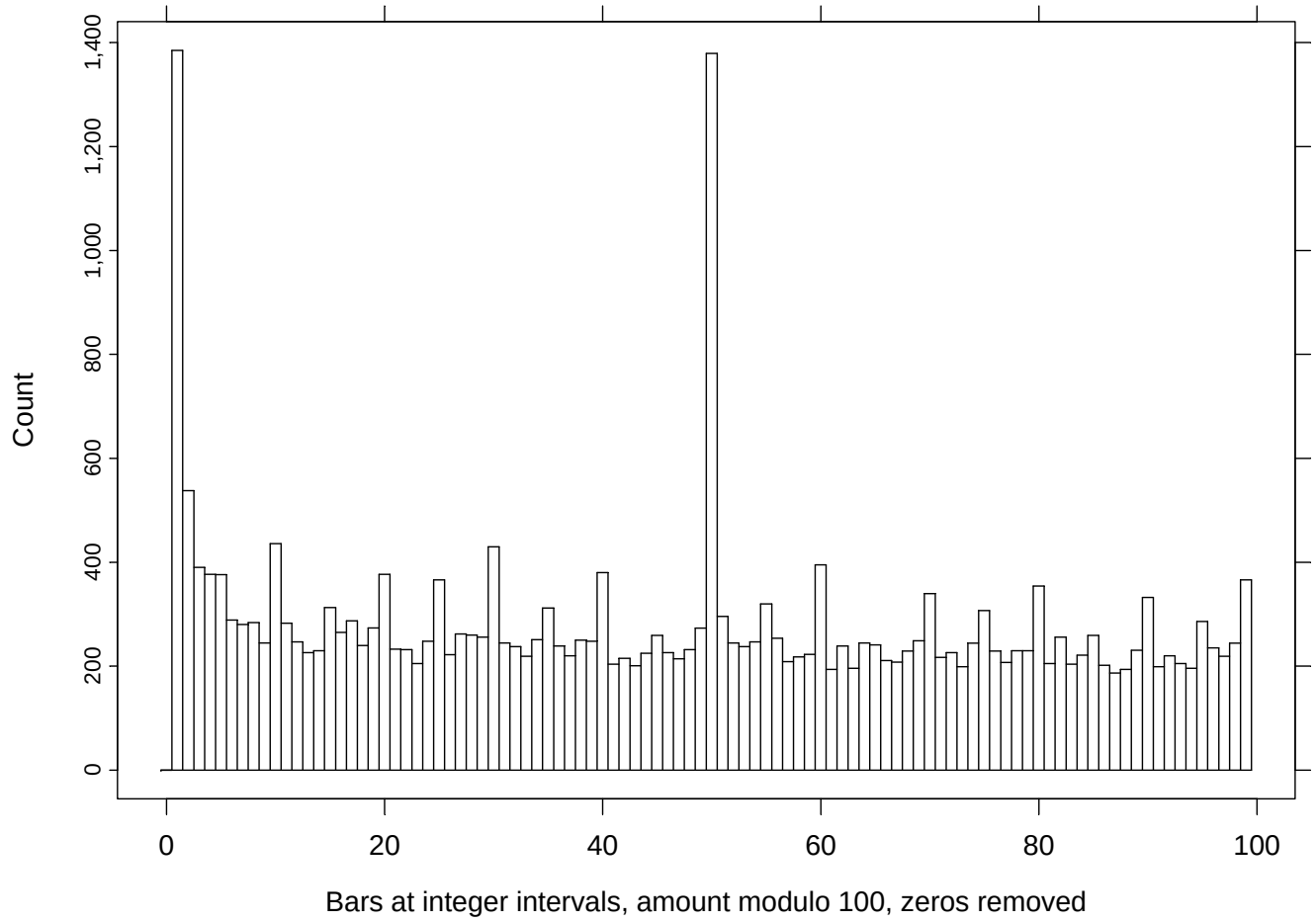
Perhaps the premises were wrong


Percent of total
0
2
4
6
8
0
20
40
60
80
100
Bars in £ intervals, from £x.51 to £(x+1).50,
196 transactions over £100 removed


Transactions at £5
Count
0
100
200
300
400
4.75
5.00
5.25
Bars at 1p intervals
Transactions at £10
Count
0
1,000
2,000
9.75
10.00
10.25
Bars at 1p intervals
Transactions at £15
Count
0
500
1,000
1,500
14.75
15.00
15.25
Bars at 1p intervals
Transactions at £20
Count
0
500
1,000
1,500
19.75
20.00
20.25
Bars at 1p intervals


Transactions at £13
Count
0
50
100
150
200
12.75
13.00
13.25
Bars at 1p intervals
Transactions at £14
Count
0
50
100
150
200
250
13.75
14.00
14.25
Bars at 1p intervals
Transactions at £16
Count
0
50
100
150
200
250
15.75
16.00
16.25
Bars at 1p intervals
Transactions at £17
Count
0
50
100
150
200
16.75
17.00
17.25
Bars at 1p intervals


Count
0
2,000
4,000
6,000
8,000
10,000
12,000
0
20
40
60
80
100
Bars at integer intervals, amount modulo 100


Count
0
200
400
600
800
1,000
1,200
1,400
0
20
40
60
80
100
Bars at integer intervals, amount modulo 100, zeros removed

Was the mathematics *inappropriate*?

Pricing models, credit risk models, etc are fine in isolation

But what happens when they are combined into a complex system?

Was there a breakdown in communication?

‘Over-reliance’ or ‘naive belief’?

Running an operation without understanding it?

But taking action on the basis of warnings requires other things to be in place

Other factors:

- hedge fund incentive structure
- deafness to warnings
- rating agencies reward related to ratings
- staff mobility
- moral hazard

Conclusion:

The fault of the mathematicians?

The fault of the mathematics?

The modern world is built on mathematics