

# VALUATION ACCURACY AND VARIATION: A META ANALYSIS

**Professor R M Skitmore MSc, PhD, FRICS, MCIQB**

School of Urban Development  
Queensland University of Technology  
Gardens Point  
Brisbane Q4001  
Australia  
[rm.skitmore@qut.edu.au](mailto:rm.skitmore@qut.edu.au)

**Ms J J Irons \* BBS (Massey) GradDipPropEc (QUT)**

Lecturer  
Department of Property Studies  
School of Economics and Finance  
Curtin University GPO Box U1987  
Perth WA 6845  
Australia  
[Janine.Irons@cbs.curtin.edu.au](mailto:Janine.Irons@cbs.curtin.edu.au)

**Dr L A Armitage Dip Surv, M Environmental Planning, PhD, FRICS, FAPI (Presenter)**

Senior Lecturer  
Faculty of Architecture, Building and Planning  
The University of Melbourne  
Victoria 3010  
Australia  
[l.armitage@unimelb.edu.au](mailto:l.armitage@unimelb.edu.au)

for *PPRES 2007*

**Corresponding Author:**

Martin Skitmore  
School of Construction Management and Property  
Queensland University of Technology  
Gardens Point  
Brisbane Q4001  
Australia  
[rm.skitmore@qut.edu.au](mailto:rm.skitmore@qut.edu.au)

**Keywords: valuation, accuracy, variation, range, approximation.**

16 February 2007

# **VALUATION ACCURACY AND VARIATION: A META ANALYSIS**

**Professor Martin Skitmore, Ms Janine Irons and Dr Lynne Armitage**

## **ABSTRACT**

This paper provides a first order approximation of the accuracy of commercial property valuations for comparison with the  $\pm 5-10\%$  threshold of tradition, convention and judicial acceptance. The nature of ranges is considered in relation to the uniform and normal probability density functions and the effects of bias considered. Summary statistics are examined for gross differences (differences between property valuations and subsequently realised transaction prices) recorded in the Investment Property Databank (IPD) database and significant yearly changes noted in both the means and standard deviations. A meta analysis of previous work is presented which shows all other results involving gross differences (GDs) to be reasonably consistent with statistics yielded by the IPD database. The results for the two main studies of net differences (GDs adjusted for the lag in time between valuation and transaction dates) also suggest yearly trends of a similar nature to the GDs. The role of intravaluer variability is also examined. The variability associated with the time lag is then estimated and applied to the IPD figures to give the approximation sought, suggesting  $\pm 5-30\%$  to be a realistic range in place of  $\pm 5-10\%$ .

## 1. INTRODUCTION

All property valuations are to one extent or another uncertain (Carsberg, 2002; Mallinson, 1994). Although typically, they are reported as a single figure rather than a range, the uncertainty surrounding them “is a normal market feature [varying] from property to property and from market condition to market condition” (Carsberg, 2002:28). Also, they essentially represent an expression of expert opinion, and as such, valuers may, in considering the same property, rightly and appropriately differ in the value conclusions at which they arrive (Mallinson, 1994; *Fenton Nominees Pty Ltd v Valuer-General*, 1981). However, whilst all value estimates carry with them some degree of uncertainty, it is acknowledged that this must be kept within reasonable bounds of reliability (Carsberg, 2002; Mallinson, 1994). This is well accepted in the market and has led to the widely held (though apparently arbitrarily established) perception that valuers are capable of valuing to “within  $\pm 5$  to 10%” of market value or price (Brown *et al.*, 1998:1). Papers on the perception of valuers in the UK and Australia have suggested 10% to 15% to be a more realistic assumption, and investigations of the legal context of property valuation variation (the margin of error) indicate that judges have allowed up to  $\pm 20\%$  following expert evidence from valuers.

However, a number of authors have noted that the 20% margin is exceeded in about 1 in 10 valuations. This clearly suggests a realistic figure to be higher than 20%, perhaps even 30% being more appropriate. An examination of previous research, or meta analysis, should shed some light on this and is the motivation for the work. Comparison between previous research results, however, is not always straightforward. Often, each is carried out in different time periods, sectors, and locations – even with different methodologies. The effects of these differences are unknown and raise questions such as, for example, is residential property in some way easier to value than commercial property and, if so, does this hold for rented apartments as opposed to owner-occupied residential units? Similarly, to what extent are there differences between rack rented valuations and those of reversionary properties?<sup>1</sup> Are differences to be expected in different countries, where there may be different approaches to valuation and different underlying processes?<sup>2</sup> Also, and of particular concern is the validity of results involving the valuation of hypothetical properties, as different interpretation of real comparables may be a significant cause of variation in real valuations.

One approach to this is to attempt a piecemeal series of analyses to prove the existence of significant influences on valuation variability and accuracy. An alternative, and one which is used in this paper, is the more traditional scientific hypothetico-deductive approach in attempting to falsify a general hypothesis (after Popper, 1959). In this case, the most obvious hypothesis is that there are no significant influences. There are immediate problems with this though, as the Investment Property Databank (IPD) database – one of the largest of its kind in the world – indicates the existence of clear yearly trends. An alternative and better hypothesis, therefore, is that the accuracy of valuations is consistent with the IPD database. The aim of this paper is to test this hypothesis by examining the results of the various empirical studies that have been made of valuation variation and accuracy.

The nature of value ranges is examined in relation to the uniform and normal probability density functions and the effects of bias are considered. Summary statistics for gross differences (differences between property valuations and subsequently realised transaction prices) are analysed drawing on records from within the IPD database. The meta analysis of previous work is then conducted, which shows that all other results for commercial properties involving gross differences (GDs) are consistent with statistics yielded by the IPD database. Further, findings from the two main studies of net differences (GD adjusted for the time lag between valuation and transaction dates) indicate that yearly trends of a similar nature (improving accuracy) to the GDs are present.

The role of intravaluer variability is also examined. This is concerned with the variation of valuations against valuations. Thereafter, the variability associated with the lag is estimated and applied to the IPD figures. This suggests, as expected,  $\pm 5$ -30% to be a realistic range for commercial properties.

---

<sup>1</sup> Whilst it would be expected to include over-rented properties, this is a phenomenon of properties held under long leases is mainly restricted to the UK market and is not likely to occur to the same extent elsewhere.

<sup>2</sup> In some countries valuations may stand outside the process and be genuine independent scorekeepers, in other countries they may be used as benchmarks for deciding which properties are sold and the prices at which they are purchased. In those cases, bias would be expected depending upon the market state, depending on each influence, etc.

## 2. METHOD

The method use is what is termed the meta analysis, described in the Oxford English Dictionary as the analysis of data from a number of independent studies of the same subject (published or unpublished), especially in order to determine overall trends and significance; an instance of this being “the analysis of analyses” and “the analysis of all the trials on a particular subject looking at the results as though they were of one.” In this study the previous analyses comprised those of Adair *et al.* (1996), Blundell & Ward (1997), Brown (1985), Brown *et al.* (1998), Crosby *et al.* (1998), Daniels (1983), Diaz & Wolverton (1998), Drivers Jonas and IPD (2000), Hager & Lord (1985), Matysiak & Wang (1995), Miles *et al.* (1991), Newell & Kishore (1998), Parker (1998) and Webb (1994).

The analysis compared the results of studies using basic three measures of valuation variability: (1) those studies (Blundell and Ward 1997; Brown 1985; Drivers Jonas and IPD 2000; Matysiak and Wang 1995) reporting the results of gross differences (i.e., the difference between valuations and eventual sales prices); (2) those studies (Blundell and Ward 1997; Newell and Kishore 1998; Parker 1998) reporting the results of net differences (i.e., the differences between valuations and eventual sales prices adjusted to a common date); and (3) those studies (Adair *et al.* 1996; Crosby *et al.* 1998; Daniels 1983; Diaz and Wolverton 1998; Hager and Lord 1985) reporting the results of intervaluer variability (i.e., the variability of the value estimates *per se*, irrespective of their relationship to actual values). Due to the idiosyncratic way in which these previous results were reported, often in terms of interquartile-like statistics, these were converted to the more conventional first and second moments to enable comparisons to be made. This was done by computer simulation of random values from both uniform and normal probability distributions - the moments of the simulation distributions being chosen by trial and error until a good match occurred with the reported results.

Having analysed the results of the previous studies in this way, the variance of the lag (gross difference – net difference) was estimated by components of variance. In a similar way, the components of variance model was used to estimate the amount of the variability in net differences that is explained by intravaluer variability and the residual (unpredicted) variability of the actual values themselves. This was then used to make an approximation of the net differences that have occurred over the period of the IPD data.

## 3. ±5-10%, ABSOLUTE, GROSS AND NET DIFFERENCE

### 3.1 ±5-10%

Valuation accuracy is concerned with the relationship between estimated value (the valuation) and actual value (as proxied by the sale price). In this paper the intuitively reasonable approach of having a negative relationship denote an undervalue and a positive relationship denote an overvalue is adopted. The simplest expression of this relationship is the pound or dollar difference between a value determination and the subsequently realised transaction price. More often, though, the percentage difference is used instead. The  $\pm$  sign necessarily indicates a finite range of such possible differences. The most common, as well as the simplest, model for this is the uniform probability density function (pdf). By definition, models with an infinite range, such as the normal pdf, are excluded. But, as shall be seen, they may still provide a reasonable approximation and offer a relatively straightforward means of analysis.

There are several means of modelling differences ‘within  $\pm 5$  to 10%’. However, prior to examining these in detail, it is necessary to first address definitional issues as they relate to the notion of value ranges. Strict literal interpretation of the  $\pm 5$  to 10% concept would result in the range being construed such that all differences are held to lie within the range of  $\pm\{5 \text{ to } 10\}$ . This is however clearly unsound as there is no practical reason to expect no differences in the range  $\pm\{0 \text{ to } 5\}$ . Another more logical and satisfactory definition, and hence the one adopted in this paper, is that the range exists on a continuum encompassing all possible values between the minimum and maximum limits of the range (i.e. an upper boundary of +10% and a lower boundary of -10%).

The  $\pm$  sign also is often taken to indicate that the values are symmetrical about their mean. The range is not given as -5% to +10% for example. If the mean is not stated, then it is assumed that the mean is 0%. This is quite crucial for, in terms of the distribution of value estimates, the mean value is seldom, if ever, exactly 0%. For example, the 1993 IPD figures show the mean difference to be -2.44 (undervalue). Assuming this is the population value then, in order for all value estimates to fall within  $\pm 5.0$  requires that the value range in this case must be  $-2.44 \pm 2.56$  or  $-5.0$  to  $+0.12$ . Table 1 shows this effect in the columns ‘Range to remain within’ on  $\pm 5$

Table 1: Effect of mean difference

| Mean difference<br>( $\bar{x}$ ) | $\pm 5$                                |   |   |                                       |          |         |          |         | $\pm 10$                               |   |   |                                       |          |         |          |         |
|----------------------------------|--|---|---|---------------------------------------|----------|---------|----------|---------|--|---|---|---------------------------------------|----------|---------|----------|---------|
|                                  | Range to remain within<br>( $\alpha$ ) | Variance<br>$\left(v = \frac{\alpha^2}{3}\right)$ | Approx mean square<br>$(m = v + \bar{x}^2)$ | Approx root mean square<br>$\sqrt{m}$ | Uniform  |         | Normal   |         | Range to remain within<br>( $\alpha$ ) | Variance<br>$\left(v = \frac{\alpha^2}{3}\right)$ | Approx mean square<br>$(m = v + \bar{x}^2)$ | Approx root mean square<br>$\sqrt{m}$ | Uniform  |         | Normal   |         |
|                                  |  |   |   |                                       | Mean abs | Var abs | Mean abs | Var abs |  |   |   |                                       | Mean abs | Var abs | Mean abs | Var abs |
| -10.0                            | -                                      | -   | -   | -                                     | -        | -       | -        | -       | $\pm 0.0$                              | 0.0   | 100.0                                       | 10.0                                  | 10.0     | 0.0     | 10.0     | 0.0     |
| -9.0                             | -                                      | -   | -   | -                                     | -        | -       | -        | -       | $\pm 1.0$                              | 0.3   | 81.3  | 9.0                                   | 9.0      | 0.3     | 9.0      | 0.3     |
| -8.0                             | -                                      | -   | -   | -                                     | -        | -       | -        | -       | $\pm 2.0$                              | 1.3   | 64.7  | 8.0                                   | 8.0      | 1.3     | 8.0      | 1.3     |
| -7.0                             | -                                      | -   | -   | -                                     | -        | -       | -        | -       | $\pm 3.0$                              | 3.0   | 52.0  | 7.2                                   | 7.0      | 3.0     | 7.0      | 3.1     |
| -6.0                             | -                                      | -   | -   | -                                     | -        | -       | -        | -       | $\pm 4.0$                              | 5.3   | 41.3  | 6.4                                   | 6.0      | 5.4     | 6.0      | 5.3     |
| -5.0                             | $\pm 0.0$                              | 0.0   | 25.0  | 5.0                                   | 5.0      | 0.0     | 5.0      | 0.0     | $\pm 5.0$                              | 8.3   | 33.3  | 5.8                                   | 5.0      | 8.3     | 5.1      | 7.3     |
| -4.0                             | $\pm 1.0$                              | 0.3   | 16.3  | 4.0                                   | 4.0      | 0.3     | 4.0      | 0.3     | $\pm 6.0$                              | 12.0  | 28.0  | 5.3                                   | 4.3      | 9.2     | 4.4      | 8.5     |
| -3.0                             | $\pm 2.0$                              | 1.3   | 10.3  | 3.2                                   | 3.0      | 1.3     | 3.0      | 1.3     | $\pm 7.0$                              | 16.3  | 25.3  | 5.0                                   | 4.2      | 8.2     | 4.1      | 8.8     |
| -2.0                             | $\pm 3.0$                              | 3.0   | 7.0   | 2.6                                   | 2.2      | 2.3     | 2.2      | 2.1     | $\pm 8.0$                              | 21.3  | 25.3  | 5.0                                   | 4.3      | 7.3     | 4.0      | 9.2     |
| -1.0                             | $\pm 4.0$                              | 5.3   | 6.3   | 2.5                                   | 2.1      | 1.8     | 2.0      | 2.3     | $\pm 9.0$                              | 27.0  | 28.0  | 5.3                                   | 4.5      | 7.3     | 4.2      | 10.0    |
| 0.0                              | $\pm 5.0$                              | 8.3   | 8.3   | 2.9                                   | 2.5      | 2.0     | 2.3      | 3.0     | $\pm 10.0$                             | 33.3  | 33.3  | 5.8                                   | 5.0      | 8.3     | 4.6      | 12.1    |
| 1.0                              | $\pm 4.0$                              | 5.3   | 6.3   | 2.5                                   | 2.1      | 1.8     | 2.0      | 2.3     | $\pm 9.0$                              | 27.0  | 28.0  | 5.3                                   | 4.5      | 7.2     | 4.2      | 10.2    |
| 2.0                              | $\pm 3.0$                              | 3.0   | 7.0   | 2.6                                   | 2.2      | 2.3     | 2.2      | 2.1     | $\pm 8.0$                              | 21.3  | 25.3  | 5.0                                   | 4.3      | 7.3     | 4.0      | 9.1     |
| 3.0                              | $\pm 2.0$                              | 1.3   | 10.3  | 3.2                                   | 3.0      | 1.3     | 3.0      | 1.3     | $\pm 7.0$                              | 16.3  | 25.3  | 5.0                                   | 4.1      | 8.1     | 4.0      | 8.7     |
| 4.0                              | $\pm 1.0$                              | 0.3   | 16.3  | 4.0                                   | 4.0      | 0.3     | 4.0      | 0.3     | $\pm 6.0$                              | 12.0  | 28.0  | 5.3                                   | 4.3      | 9.2     | 4.4      | 8.8     |
| 5.0                              | $\pm 0.0$                              | 0.0   | 25.0  | 5.0                                   | 5.0      | 0.0     | 5.0      | 0.0     | $\pm 5.0$                              | 8.3   | 33.3  | 5.8                                   | 5.0      | 8.3     | 5.1      | 7.4     |

and  $\pm 10$  for a series of mean differences. Therefore, for a given accuracy, such as ‘within 5%’, this always implies a smaller range about the mean unless the mean is exactly 0%.

Accordingly, to model this better it is necessary to combine the mean of the differences with the range around the mean. The most obvious way to do this is by use of the mean square measure. This can be obtained from the mean and variance by the formula:

$$\frac{1}{n} \sum_{i=1}^n x_i^2 = \frac{v(n-1)}{n} + \bar{x}^2 \quad (1)$$

where  $\frac{1}{n} \sum_{i=1}^n x_i^2$  is the mean square,  $\bar{x}$ ,  $v$  and  $n$  are the sample mean, variance and size respectively. For the

uniform distribution  $v = \frac{(2\alpha)^2}{12} = \frac{\alpha^2}{3}$ , where  $\alpha$  denotes the value of the  $\pm$  range. Thus, for example, a range of  $\pm 5$  has a  $v=8.3$  and a range of  $\pm 10$  has a  $v=33.3$ . Table 1 summarises the results obtained for a series of mean differences, showing the mean square and root mean square values needed to remain within  $\pm 5$  and  $\pm 10$ . For example, with a mean difference of  $-4.0$ , the maximum root mean square that will allow all values to be within  $\pm 5$  is  $4.0$  ( $5.3$  in the case of  $\pm 10$ ).

As mentioned earlier, it is often more convenient to work with models with a theoretically infinite range. The normal pdf is a particular example of this. Here, a popular ‘range’ is  $\pm 2$  standard deviations (SDs) from the mean as this covers 95% of the values. Using this measure, instead of  $\pm 5$  with the uniform model, we use a SD of  $2.5$  with the normal model. Likewise, instead of  $\pm 10$  with the uniform model, we use a SD of  $5$  with the normal model.

### 3.2 Absolute difference

A number of studies have used the absolute difference as a measure of accuracy. It is clearly of intuitive appeal, with a mean say, of  $5$ , suggesting that there are approximately as many differences that are less than  $\pm 5$  than there are more than  $\pm 5$ . What this does not reveal however is the proportion that lies below some other figure than the mean absolute difference. In particular, we need to know the mean absolute difference that corresponds to a uniform (or normal equivalent)  $\pm 5$  and  $\pm 10$ . Fortunately, it is easy enough to find this by simulation. Table 1 summarises the results for  $30,000$  simulations using the mean differences and variances already described. This indicates the importance of the using some other measure, such as the variance of the absolute differences, when estimating the parameters of the original pdf, as the mean alone can represent a wide range of possibilities.

### 3.3 Gross and net difference

With a few notable exceptions, the empirical measure of the difference between estimated and actual value is that of the valuation and subsequently realised transaction price of the same property. We term this the **gross difference** (GD). In order to avoid the possibility of the latter influencing the former, and thus becoming a self-fulfilling prophesy, analysis is usually restricted to valuation-transaction events that are separated by a suitable minimum length of time. This minimum length of time, or lag, varies between two and four months. This is only a minimum, however, and the average is rather longer (the average IPD lag is consistently close to  $10$  months). Of course, it is usually reasonable to expect the value to have changed over the period between the valuation and the transaction, so the estimated-actual value comparisons are misaligned.

The **net difference** (ND) should then compensate for this misalignment. Some studies have attempted to do this (eg. Blundell and Ward, 1997; Newell and Kishore, 1998). The intention, therefore, is that the ND reflects the true difference between estimated and actual values, i.e.,  $d_n = d_g - l$ , where  $d_n$ ,  $d_g$  and  $l$  denote the ND, GD and lag ‘effect’ respectively. Of particular interest here is the value of the  $l$  effect as, with most of the data available being concerned with measures of GDs, knowing this enables us to arrive at the NDs.

Fig 1: IDP yearly trends

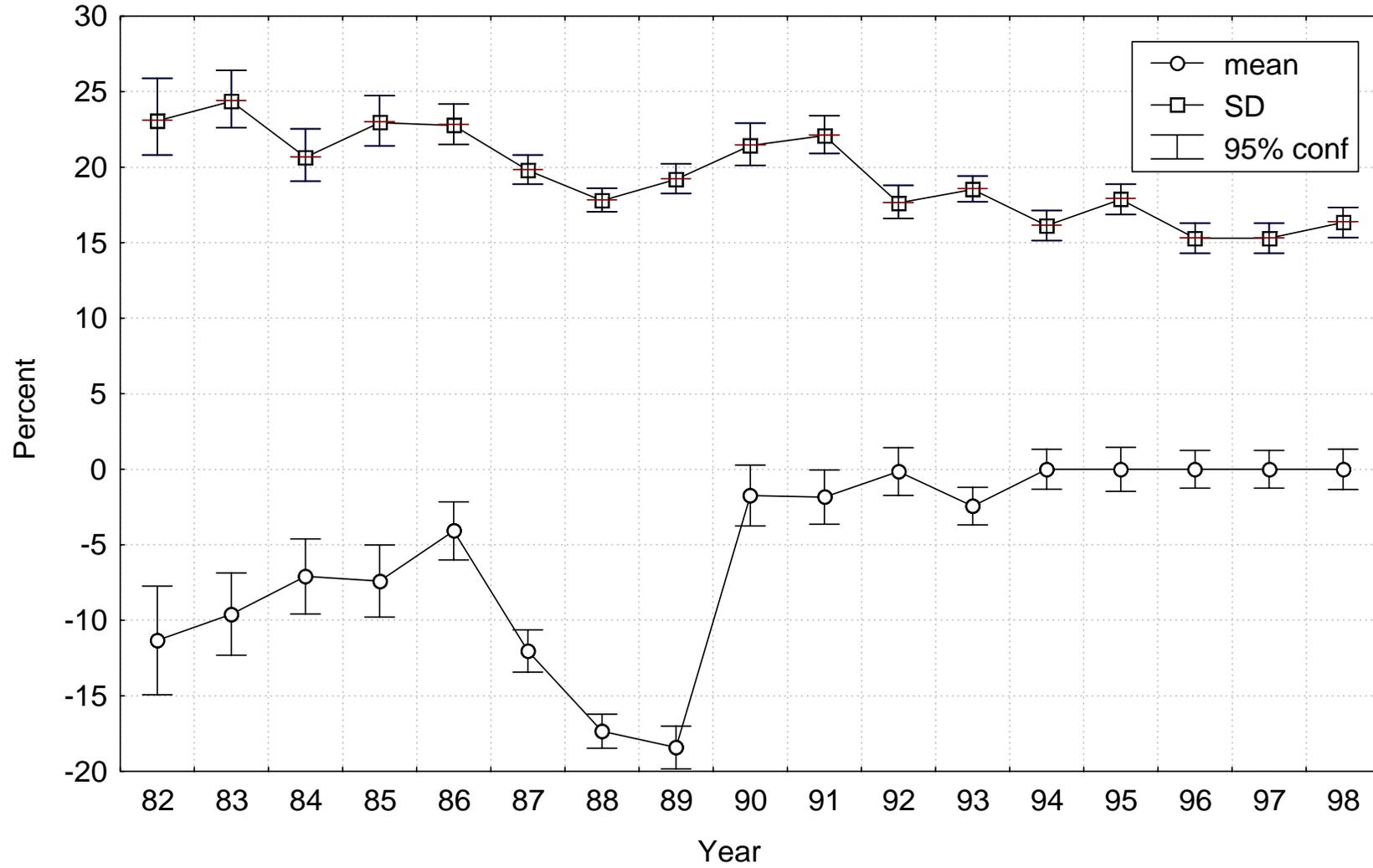
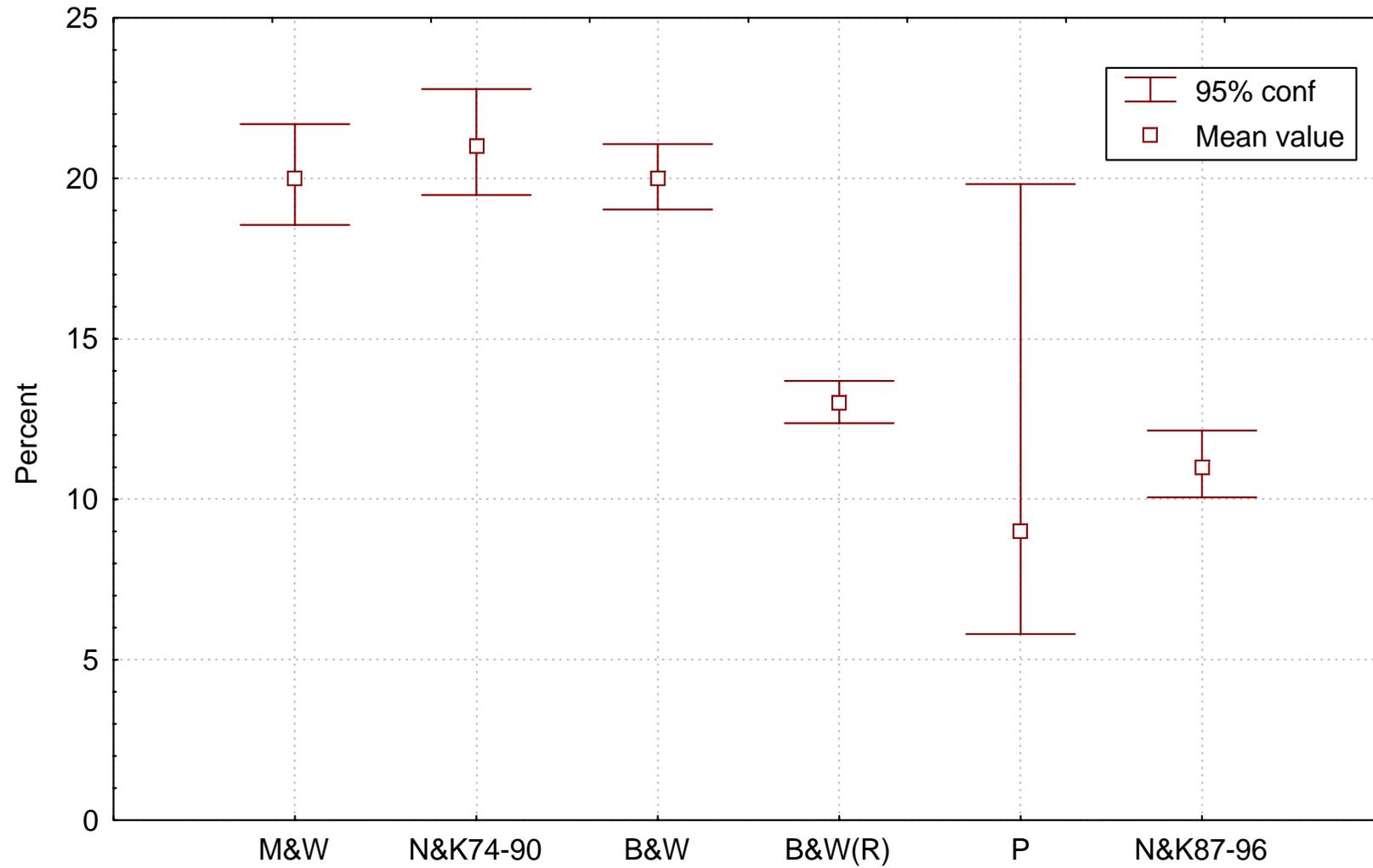


Fig 2: Standard deviation of gross differences



## 4. PREVIOUS STUDIES

### 4.1 Gross difference

By far the largest dataset of GDs is that produced by IPD, said to reflect “the profile and performance of something over 75% of the [UK] institutional market in commercial properties” (Cullen, 1994: 91). The dataset contains records of some 23,000 individual properties of which nearly 11,000 have been sold since 1982. Of those 11,000, nearly 7,000 represent property sales each of which is preceded by at least two full professional valuations that are also recorded in the dataset. The published figures have an imposed minimum lag of four months between the date of the valuation and the date of the sale. Figure 1 provides a yearly summary of these. This shows the mean differences and the SD around the mean. Also shown is the approximate 95% confidence interval for the means ( $\pm 2$  standard errors) and SDs (from the  $\chi^2$  distribution)<sup>3</sup>. Clearly observable is the dip in mean estimated values over 1987-9 showing the extent of under valuations during that boom period in the UK property market – a smoothing effect already noted anecdotally by Matysiak and Wang (1995). The several significant yearly changes in SDs observable also serve to highlight the importance of changes in market conditions. Overall, there appears to be a trend of reducing SDs over the period – reducing from around 23% in the early eighties to a current figure of just over 15%.

Table 2 summarises the majority of previous results from studies concerned with estimated-actual value differences. With the exception of Blundell and Ward (1997) and Newell and Kishore’s (1998) ‘adjusted’ results, all are related to GDs.

In one of the earliest of these, Matysiak and Wang (1995) sought primarily to consider valuer performance under different market conditions. They also analysed a sample of 317 commercial properties from the JLW Property Performance Analysis System that were sold in the period 1973-1991 for which valuations had been conducted in the period 3-6 months prior to sale. The mean difference and mean absolute difference was found to be -6.9% and 16.7% respectively, with 30% of the valuations lying within a range of plus or minus 10% of the sale price, 55% within plus or minus 15%, and 70% within plus or minus 20%. By generating 30,000 computer simulated values from a **normal** pdf, the nearest matching results to these were found with mean and SD values of -7% and 20% respectively. Similarly, by generating 30,000 computer simulated values from a **uniform** pdf, the nearest matching results to these were found with mean and SD values of -7% and 19% respectively. With a sample size of 317, this indicates the true (population) SD for the normal pdf to lie within the 95% confidence interval of 18.55 to 21.69%. As an alternative, a regression model was built via the IPD statistics, using the SD as the dependent variable and the mean difference, mean absolute difference and percentage valuations lying within  $\pm 20\%$  as the independent variables, together with an interaction term. Applying this to Matysiak and Wang’s results produced a predicted SD of 21.34%. The 95% confidence interval in this case is therefore 19.89 to 23.03%. Although the IPD dataset does not include records as far back as 1973, a comparison can be made from 1982 to 1991. The average SD of the IPD data over this period is 21.41, which is not significantly different from Matysiak and Wang’s results, regardless of the manner in which they are calculated.

A further study proceeding along similar lines was undertaken by Blundell and Ward (1997). The database employed by Matysiak and Wang was again utilised, but a larger sample group of 747 properties, for which sale prices were available over the period 1974-1990 and for which valuations had been obtained in the period 3-6 months prior to sale, was identified. The findings indicated approximately 80% of the valuations lay within plus or minus 20% of the sale price, and only 35% were within plus or minus 10% of the sale price. Computer simulations indicate the SD of the nearest normal pdf equivalent to be 20% (19.03 to 21.07% at the 95% confidence interval), which is again close to the IPD figure.

One of the earliest studies into GDs was undertaken by Brown (1985). Using a sample of 29 properties, Brown found that regressing appraised values with the actual sale prices subsequently achieved produced an  $r^2$  value of 0.99. A later study by Drivers Jonas and IPD (2000) employed a similar regression methodology to Brown.

---

<sup>3</sup> The sample estimates of the first two moments follow the well-known standard error and  $\chi^2$  distribution.

Table 2: Valuation/sales deviations

| Source                     | Data years         | Type               | n     | Frequency |      |      |      | Pdf | Range          | Mean abs  | Mean      | SD  |
|----------------------------|--------------------|--------------------|-------|-----------|------|------|------|-----|----------------|-----------|-----------|-----|
|                            |                    |                    |       | <5%       | <10% | <15% | <20% |     |                |           |           |     |
| Matysiak & Wang (1995)     | 1973-91            | Retail & office    | 317   | -         | 30%  | 55%  | 70%  | E   |                | 16.7%     | -6.9%     |     |
|                            |                    |                    |       | 19%       | 36%  | 52%  | 65%  | N   | 16.8%          | 7         | 20        |     |
|                            |                    |                    |       | 15%       | 30%  | 46%  | 61%  | U   | 16.5%          | 7         | 19        |     |
| Blundell & Ward (1997)     | 1974-90            | Retail and office  | 747   | -         | 35%  | -    | 80%  | E   |                | -         | Avg -7%   |     |
|                            |                    |                    |       | 18%       | 36%  | 52%  | 66%  | N   | 16.9%          | -7        | 20        |     |
|                            | Ditto (adjusted)   | Ditto              | Ditto | -         | 55%  | -    | 85%  | E   |                | -         | Avg -3%   |     |
|                            |                    |                    |       | 29%       | 55%  | 74%  | 87%  | N   | 10.7%          | -3        | 13        |     |
|                            |                    |                    |       | 26%       | 52%  | 79%  | 95%  | U   | 9.8%           | -3        | 11        |     |
| Parker (1998)              | Nov 1995           | All                | 7     | 15%       | 85%  | 100% | -    | E   | -14.3 to -1.6% | 7.7%      | Avg -3.2% |     |
|                            |                    |                    |       | 40%       | 71%  | 88%  | 96%  | N   |                | 7.6%      | -3.2      | 9   |
|                            |                    |                    |       | 34%       | 68%  | 90%  | 100% | U   |                | 7.7%      | -3.2      | 8.5 |
|                            |                    | Retail             |       | -         | -    | -    | -    | E   |                | -         | Avg -2.6% |     |
|                            |                    | Commercial         | -     | -         | -    | -    | E    | -   | Avg 4.1%       |           |           |     |
|                            |                    | Industrial         | -     | -         | -    | -    | E    | -   | Avg -8.5%      |           |           |     |
| Newell & Kishore (1998)    | 1987-96 (adjusted) | Offices and retail | 218   | -         | 65%  | -    | 91%  | E   |                | Approx 9% | Avg -2%   |     |
|                            |                    |                    |       | 35%       | 63%  | 82%  | 92%  | N   | 9.0%           | -2        | 11        |     |
|                            |                    |                    |       | 29%       | 58%  | 87%  | 100% | U   | 8.8%           | -2        | 10        |     |
| Webb (1994)                | 1978-90            | ?                  | 469   | -         | -    | -    | -    | E   | 11%            | -         | -         |     |
| Miles <i>et al.</i> (1991) |                    | In NCREIF database |       | -         | -    | -    | -    | E   | 10.7%          | -         | -         |     |

Table 3:  $r^2$  values for simulated data

|                                   | 1982   | 1983   | 1984   | 1985   | 1986   | 1987   | 1988   | 1989   | 1990   | 1991   | 1992   | 1993   |
|-----------------------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| Mean price                        | 47.45  | 65.21  | 86.82  | 89.34  | 84.01  | 92.91  | 122.76 | 160.09 | 174.67 | 164.88 | 142.78 | 137.48 |
| SD                                | 67.35  | 55.26  | 197.92 | 152.96 | 114.81 | 119.44 | 147.20 | 182.18 | 194.62 | 185.34 | 150.48 | 145.55 |
| Mean price<br>(sim)               | 45.29  | 66.77  | 83.92  | 78.74  | 81.11  | 96.49  | 126.41 | 163.54 | 178.79 | 173.95 | 143.18 | 136.87 |
| SD                                | 56.16  | 54.75  | 155.21 | 119.67 | 103.31 | 130.81 | 151.48 | 203.06 | 203.27 | 196.70 | 152.38 | 127.74 |
| Mean error%                       | -11.33 | -9.59  | -7.09  | -7.40  | -4.08  | -12.03 | -17.34 | -18.42 | -1.74  | -1.84  | -0.15  | -2.44  |
| SD                                | 23.07  | 24.36  | 20.66  | 22.95  | 22.76  | 19.79  | 17.79  | 19.19  | 21.43  | 22.09  | 17.63  | 18.52  |
| Mean error%<br>(sim)              | -10.70 | -9.42  | -7.34  | -7.99  | -5.05  | -12.72 | -17.09 | -18.22 | -1.88  | -0.59  | -0.28  | -2.77  |
| SD                                | 23.04  | 24.71  | 20.47  | 23.03  | 22.65  | 20.14  | 18.07  | 19.11  | 21.26  | 22.08  | 17.89  | 18.32  |
| Mean abs<br>error                 | 19.71  | 19.94  | 14.69  | 17.30  | 16.75  | 17.34  | 20.01  | 21.18  | 15.79  | 16.00  | 11.76  | 12.77  |
| SD                                | 16.50  | 16.96  | 16.17  | 16.80  | 15.95  | 15.35  | 14.72  | 16.10  | 14.60  | 15.34  | 13.14  | 13.63  |
| Mean abs<br>error (sim)           | 21.24  | 22.36  | 18.47  | 20.69  | 19.88  | 19.92  | 20.16  | 21.50  | 18.41  | 19.18  | 15.45  | 16.03  |
| SD                                | 13.93  | 14.11  | 11.46  | 12.88  | 11.95  | 13.07  | 14.55  | 15.33  | 10.78  | 10.94  | 9.00   | 9.27   |
| ±20%                              | 55.69  | 58.10  | 73.45  | 63.73  | 67.61  | 64.36  | 55.87  | 51.80  | 69.74  | 70.70  | 80.68  | 79.01  |
| ±20% (sim)                        | 52.50  | 47.8   | 57.30  | 50.60  | 52.30  | 55.50  | 54.10  | 53.90  | 55.00  | 52.70  | 63.90  | 63.50  |
| Mean<br>valuation                 | 40.69  | 87.36  | 77.18  | 79.87  | 76.71  | 77.59  | 97.19  | 129.82 | 173.11 | 164.85 | 142.91 | 132.69 |
| SD                                | 59.98  | 71.04  | 186.62 | 143.18 | 105.05 | 96.21  | 110.57 | 147.27 | 197.28 | 188.99 | 149.39 | 140.18 |
| Mean<br>valuation<br>(sim)        | 40.79  | 60.94  | 77.31  | 73.00  | 76.78  | 84.34  | 105.90 | 133.80 | 176.30 | 172.06 | 143.68 | 131.88 |
| SD                                | 55.15  | 56.96  | 138.71 | 115.11 | 103.10 | 112.44 | 131.08 | 163.30 | 220.78 | 204.35 | 160.86 | 124.42 |
| $r^2$                             | 0.9539 | 0.9370 | 0.9664 | 0.9458 | 0.9649 | 0.9480 | 0.9628 | 0.9598 | 0.9632 | 0.9641 | 0.9729 | 0.9635 |
| $r$                               | 0.9099 | 0.8780 | 0.9340 | 0.8945 | 0.9310 | 0.8987 | 0.9269 | 0.9211 | 0.9278 | 0.9294 | 0.9465 | 0.9284 |
| Prob>0.99<br>for 29<br>valuations | 0.020  | 0.001  | 0.078  | 0.044  | 0.028  | 0.025  | 0.019  | 0.013  | 0.012  | 0.022  | 0.023  | 0.012  |

The 1988 study utilised a sample of nearly 1,450 properties for which a sale price in the period 1982 to 1988 was available and a minimum of two valuations in the 24-month period prior to the sale had been conducted. All valuations undertaken within four months of any sale transaction were disregarded to ensure that the sale price was unknown at the date of valuation. On average valuations were conducted almost 10 months prior to sale date. The analysis produced an  $r^2$  of 0.93.

A simulation study is used to aid in the understanding of this result. Using the IPD statistics for 1982, for example, a random value was generated from a log normal pdf (to avoid the possibility of negative values) with a mean of 47.45 and SD of 67.35 to represent price on a per square foot (psf) basis together with a random value from a normal pdf with mean -11.39 and SD 23.07 to represent the error percentage. The valuation was then represented by  $p(1 + e/100)$ , where  $p$  is the simulated price psf and  $e$  is the simulated error percentage. Repeating this 1,000 times produced a set of simulated valuations. The simulated prices were then regressed with the associated simulated valuations producing a correlation coefficient of  $r=0.9539$  ( $r^2=0.9099$ ). This was repeated for the years 1983 to 1993. The results are summarised in Table 3, which also gives the actual and simulated (italic) statistics for comparative purposes. These results suggest that the simulations, with notably few exceptions, provide a reasonable first approximation in terms of means and SDs of the various summary statistics and the percentage of valuations falling into the  $\pm 20\%$  range as well as the Drivers Jonas and IPD (2000)  $r^2=0.93$  result. Brown's  $r^2=0.99$  result may be due to the small sample size involved. To check this, the simulations were repeated, this time for 29 instead of 1,000 valuations. The  $r^2$  was recorded and the process then repeated 1,000 times, counting the number of times the  $r^2$  value exceeded 0.99. For the 1982 simulation, this occurred 20 times – a probability of  $20/1000=0.020$ . The results for all the years are given also in Table 3, indicating the unlikelihood of Brown's result occurring with the IPD data. All are below the conventional  $p=0.05$  cut-off, with the exception of the 1984 result – coinciding with the year of publication of Brown's result. It is just possible, therefore, that Brown's data coincided with just the one period the same result could have occurred with a similarly small sample of IPD data!

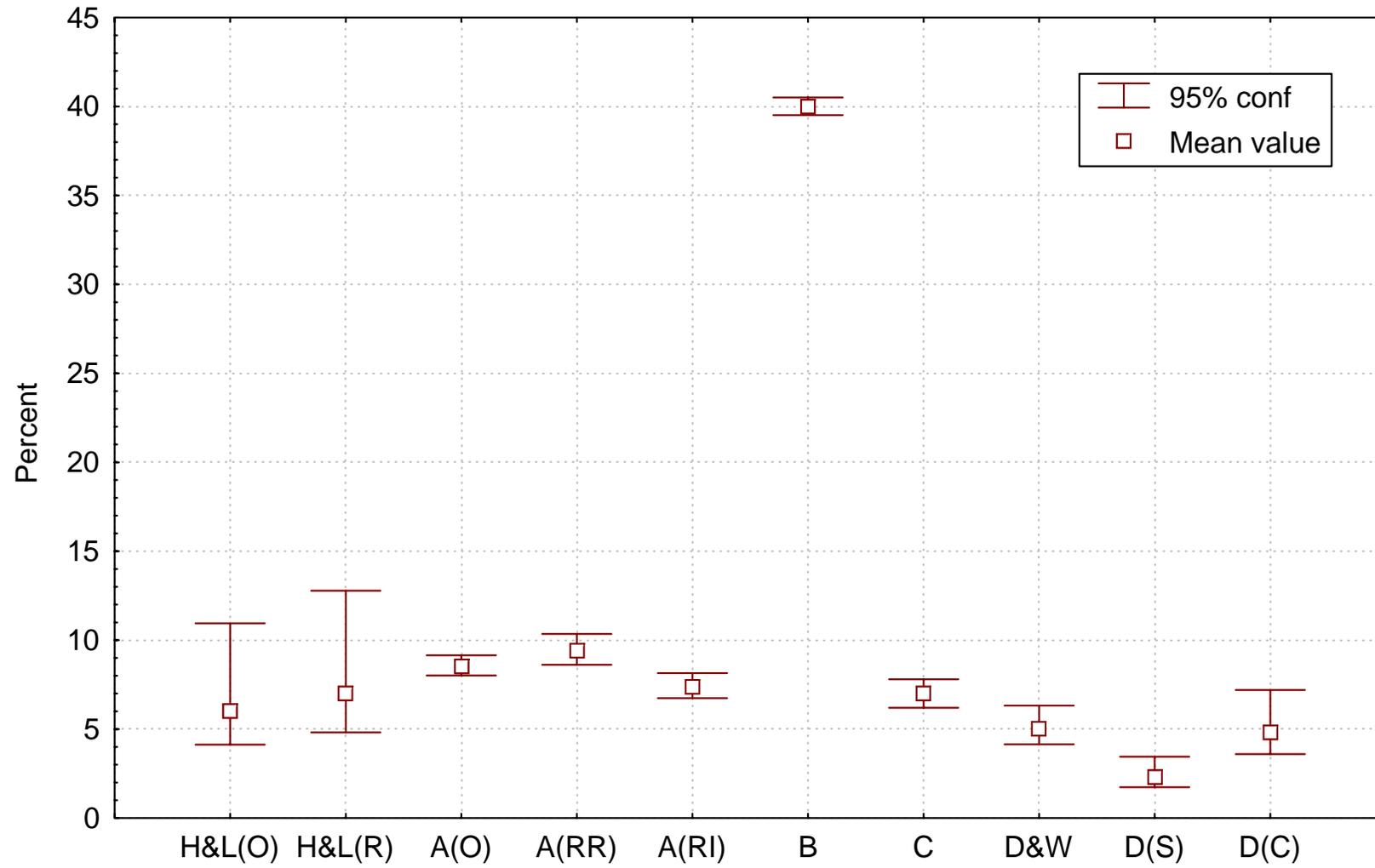
## 4.2 Net difference

Blundell and Ward (1997) attempted to remove the lag effect by adjusting their data to reflect subsequent movements in the market. They acknowledged the deficiencies in the relatively crude adjustment approach applied, noting that it failed to capture differential movements in the market on a geographical basis. Analysis of the adjusted data still suggested an average under valuation in the region of 3%. It was further suggested that 85% of the valuations would lie within a range of plus or minus 20% of the sale price, and that only 55% would lie within a range of 10% of sale price. Our computer simulations indicate the SD of the nearest normal pdf equivalent of this to be 13% (12.37 to 13.69% at the 95% confidence interval).

An Australian based study undertaken by Newell and Kishore (1998) assessed sale prices relative to valuations for 101 office properties and 117 retail properties in the Sydney area over the period 1987 and 1996. The time period covered allowed not only the issue of accuracy to be considered but also permitted some assessment to be made of valuer performance in different market states. Data were drawn from the Commercial Property Monitor database, records of the New South Wales Valuer-General, and information furnished by the Independent Property Trust Review. The study included only those properties for which a valuation within the 12 months prior to sale was undertaken. The time lag seen in the data was accordingly relatively short with an average time period of 4.5 months present between the date of last valuation and sale. The data was also adjusted to reflect market movement in the intervening period between valuation and sale date using the Property Council of Australia Indices. The analysis indicated that 65% of the valuations assessed were within a range of plus or minus 10% of the sale price and 91% lay within plus or minus 20% of the sale price. The mean difference and mean absolute difference was -2% and 9% respectively. Our computer simulations indicate the SD of the nearest normal pdf equivalent of this to be 11% (10.06 to 12.14% at the 95% confidence interval).

Parker (1998) also provides some insight into ND in an Australian context. The sample group utilised in the study consisted of a small group of properties that were offered for open market sale by tender. The subject properties were independently valued by one national valuation company as at the date on which the tender closed, being a day in November 1995. Each valuer was furnished with identical instructions, together with a data set containing full information on the properties, and a normal market fee for undertaking the appraisals was charged.

Fig 3: Intervaluer standard deviations



Further, the portfolio contained seven standard (i.e. not special or unusual) industrial, office and retail investment properties. The methodology adopted was notable in that it served to overcome many of the limitations and problems inherent in previous studies in this area. Not least amongst these was the absence of any time lag, with the concurrent performance of the sale transactions and valuation exercises. The mean difference and mean absolute difference was  $-3.2\%$  and  $7.7\%$  respectively. On a sectoral basis, lower variation, and hence higher accuracy, was exhibited by retail property ( $-2.6\%$  mean difference) followed by commercial ( $4.1\%$ ) and industrial property ( $-8.5\%$ ). On an individual property basis the difference ranged between a low of  $1.6\%$  to a high of  $14.3\%$ . Additionally,  $15\%$  of the subject valuations were found to lie within a range of  $5\%$  either side of sale price,  $85\%$  within a range of  $10\%$ , and all valuations fell within  $\pm 15\%$  of sale price. Our computer simulations indicate the SD of the nearest normal pdf equivalent of this to be  $9\%$  ( $5.80$  to  $19.82\%$  at the  $95\%$  confidence interval).

To explain these three results, attention again turns to the IPD data. It has been observed that the SD of the IPD GDs has improved from around  $23$  to  $17\%$  over the period  $1982-98$ . A similar improvement in NDs appears to have occurred from Blundell and Ward's  $1974-90$  period ( $13\%$  SD) to Newell and Kishore's  $1987-96$  period ( $11\%$  SD). Also, if Parker's mean result is to be given credence, there is a suggestion of an even further improvement to  $9\%$  by  $1995$ .

### 4.3 Intervaluer variability

In addition to gross or net differences, several studies have considered the variability of the value estimates *per se*, irrespective of their relationship to actual values – in other words, intervaluer variability. Of these (Table 4), the two with significantly larger sample sizes are those reported in Crosby *et al.* (1998) and Adair *et al.* (1996).

Crosby *et al.* conducted an analysis on data collected by Morgan (1993) from five companies covering the period  $1983$  to  $1985$ . The data set utilised contained  $120$  retail, office, and industrial investment properties in eight individual portfolios within the United Kingdom, for which two professional valuations had been undertaken. The analysis considered the paired valuations, one provided by one firm, and the second furnished by one of four firms. The difference between the valuations was then determined. The average of these differences was in the order of  $8.6\%$ . Over  $40\%$  of the valuations lay within a range of  $5\%$  either side of each other, over  $65\%$  were within  $10\%$ , and  $90\%$  were within  $20\%$ . The average absolute difference was the lowest for retail property followed closely by office property. By generating random values from two iid (independently and identically distributed) normal pdfs, the closest result to this was obtained with mean zero and SD  $7\%$  ( $6.2$  to  $7.8\%$  at the  $95\%$  confidence interval).

In Adair *et al.*'s work, a sample of valuers was invited to value a number of UK commercial properties, including hypothetical retail, office, and industrial property in  $14$  main centres throughout the UK. For each of the  $14$  centres five valuers from local firms, and five from national firms were to provide valuations on hypothetical subject properties in actual locations, yielding  $446$  valuations. As the valuations were conducted on hypothetical properties (though real locations were used) and the study participants were not paid for completing their valuations, the results reported in the study may, *inter alia*, be a product of these factors. Nevertheless, the overall findings were that  $61\%$  of all valuations conducted on the rack rented properties lie within a range of  $10\%$  of the mean of the valuations, and  $85\%$  within a range of  $20\%$  of the mean. For the reversionary valuations,  $69\%$  lie within a range of  $10\%$  of the mean, and over  $90\%$  within a range of  $20\%$  of the mean. The overall average absolute variation was  $9.53\%$ , with  $10.5\%$  and  $8.48\%$  for the rack rented properties and reversionary interest properties respectively. Our computer simulations indicate the SD of the nearest normal pdf equivalent of this to be  $9.41\%$  ( $8.62$  to  $10.35\%$  at the  $95\%$  confidence interval) for the rack-rented properties and  $7.38\%$  ( $6.74$  to  $8.15\%$  at the  $95\%$  confidence interval) for the reversionary interest properties, making the SD for the rack rented properties significantly **higher** than the Crosby *et al.* study. This is exactly as expected in the UK market given that the problem with rack rented properties is that, although the valuation (traditional approach) only requires two major inputs, market rental value (MRV) and all risks yield (ARY), both are unknown and subject to variation caused by variable information flows and different interpretation. Reversionary properties have four inputs, the extra ones being rent and unexpired term to next rent change, both fixed. Variation in MRV and ARY has greater effect on rack rented valuation than reversionary so that the hypothesis should be that reversionary property valuations will vary less than rack rented due to the certainty of cash flow to the next rent change. Incentives will increase the variation but they cause difficulties for both rack-rented and reversionary properties as both valuation require a MRV estimate.

Table 4: Intervaller valuation/sales deviation

| Source                      | Valuation Date/Data Years | Type                           | n   | Frequency  |       |      |       | Dist      | Mean abs   | Mean  | SD   |
|-----------------------------|---------------------------|--------------------------------|---|------------|-------|------|-------|-----------|------------|-------|------|
|                             |                           |                                |   | <5%        | <10%  | <15% | <20%  |           |            |       |      |
| Hager & Lord (1985)         |                           | Rack rented office             | 10 valuations<br>1 property               | 40%        | 90%   | -    | 100%  | Empirical | -          | -     |      |
|                             |                           |                                |   | 60%        | 90%   | 99%  | 99%   | Normal    | 4.8%       | 0     | 6    |
|                             |                           | Retail - reversionary          | 10 valuations<br>1 property               | 50%        | 80%   | -    | 90%   | E         | -          | -     |      |
|                             |                           |                                |   | 52%        | 85%   | 97%  | 99%   | N         | 6.0%       | 0     | 7    |
| Adair <i>et al.</i> (1996)  | 15 Feb 1995               | Offices, retail & industrial   | 446 valuations<br>hypothetical properties | 34%        | 65%   | -    | 89%   | E         | 9.53%      | -     | 8.55 |
|                             |                           |                                |   | 30         | 57    | 79   | 92    | N         | 9.6        | 8     | 8.55 |
|                             |                           |                                |   | 34         | 57    | 74   | 90    | U         | 9.55       | 8     | 8.55 |
|                             |                           | Rack rented                    | 232                                       | 30%        | 61%   | -    | 85%   | E         | 10.5%      | -     | 9.41 |
|                             |                           |                                |   | 29         | 56    | 76   | 90    | N         | 10.0       | 8     | 9.41 |
|                             |                           |                                |   | 30         | 53    | 68   | 84    | U         | 10.6       | 9     | 9.41 |
| Reversionary interest       | 214                       | 30%                            | 69%                                       | -          | 93%   | E    | 8.48% | -         | 7.38       |       |      |
|                             |                           | 34                             | 64  | 86         | 97    | N    | 8.7   | 7         | 7.38       |       |      |
|                             |                           | 39                             | 62  | 81         | 100   | U    | 8.3   | 7         | 7.38       |       |      |
| Crosby <i>et al.</i> (1998) | 1983-5                    | Retail, offices and industrial | 120 properties                            | 43%        | 69%   | 78%  | 88%   | E         | 8.6%       | -2.81 |      |
|                             |                           |                                | 240 valuations                            | 37%        | 67%   | 85%  | 95%   | N         | 8.3%       | -2.81 | 7    |
| Diaz & Wolverton (1998)     | 1995                      | 1 apartment property           | 45 valuations                             | almost 70% | 93.5% | -    | -     | E         | 3.74-5.31% | -     |      |
|                             |                           |                                |   | 68%        | 95%   | 100% | 100%  | N         | 4.3%       | 0     | 5    |
|                             |                           |                                |   | 58%        | 100%  | 100% | 100%  | U         | 4.3%       | 0     | 5    |
| Daniels (1983)              |                           | 1 simple residential           | 18 valuers                                | 50%        | 95%   | 100% | -     | E         | 5.3%       | -     | 2.3  |
|                             |                           |                                |   | 50         | 98    | 100  | 100   | N         | 5.0        | 5     | 2.3  |
|                             |                           |                                |   | 43         | 100   | 100  | 100   | U         | 5.5        | 5.5   | 2.3  |
|                             |                           | 1 complex residential          | 18 valuers                                | 39%        | 50%   | 95%  | -     | E         | 8.9%       | -     | 4.8  |
|                             |                           |                                |   | 23         | 62    | 91   | 99    | N         | 8.6        | 8.5   | 4.8  |
|                             |                           |                                |   | 26         | 56    | 86   | 100   | U         | 9.0        | 9     | 4.8  |

All other intervaluer studies of this kind have much smaller sample sizes, and therefore larger confidence intervals of the SD. The early paper by Hager and Lord (1985), for example, also considered intervaluer variability. In this study, all the valuers were given identical valuation instructions, and two properties were valued, a rack-rented office property, and a reversionary retail shop. The resulting valuations were then compared with 'control' values determined by an expert valuer. For the office property valuations, 40% lay within 5%, and 90% within 10%, of the control. For the retail premises, 50% lay within 5%, and 80% within 10%, of the control. Of the ten valuations for each of the two properties, all but one lay within a range of 20% of the control. The findings tend to support the later Adair *et al.* study but also, as with Adair *et al.*'s work, the validity of the instructions and information provided to the valuers and the quality of the valuations, as no fee was involved, has been questioned (Reid, 1985). Our computer simulations indicate the SD of the nearest normal pdf equivalent in this case to be 6% (4.13 to 10.94% at the 95% confidence interval) for the offices and 7% (4.81 to 12.78% at the 95% confidence interval) for the retail reversionary property, which encompasses both the Crosby *et al.* and Adair *et al.* results.

Several recent American studies have also considered, though not always as a primary focus, the issue of intervaluer variability. Diaz and Wolverton (1998), for instance, focused on the degree to which valuers are influenced by valuations they have previously conducted; the hypothesis being that valuers would anchor onto their prior value estimates and thus be inappropriately influenced by this figure when conducting a subsequent re-valuation. The authors' arranged for three sets of valuations to be conducted on the same hypothetical Atlanta apartment property within the space of 12 months from April 1995. The sample group of valuers appraised the property and subsequently, eight months later, were requested to revalue the property in light of specified property and market changes. Diaz and Wolverton also obtained at this time 'control' valuations from an independent sample of valuers. All valuers were furnished with identical instructions and a detailed data set on the property. The 46 valuations gathered were then examined to assess the variation seen in the results. They found a very low level of variation, with an absolute average variation of between 3.74% and 5.31% seen in the three sets of valuations. Almost 70% of the valuations obtained lay within a range of 5% of the mean of each set of valuations, and only 6.5% of the valuations were further than 10% from the mean. Our computer simulations indicate the SD of the nearest normal pdf equivalent in this case to be 5% (4.14 to 6.32% at the 95% confidence interval), which is clearly lower than both the Crosby *et al.* and Adair *et al.* results, probably due to the frequently simpler task of valuing residential property.

The issue of intervaluer variability has also been considered in an Australian context by Daniels (1983) in a small-scale study of residential properties in South Australia. In this case, two freehold detached residential houses, one a 'typical' mid-value property located in a 'typical' suburb, and the other a more difficult exercise being a large home in a superior suburb, were valued as at the same date by 18 different valuers. Each valuer inspected both properties on the same day but all inspections were conducted at different times. The mean absolute error was 5.3% for the 'simple' property and 8.9% for the more 'complex' property. Additionally, for the 'simple' property, 50% of the valuations lay within a range of  $\pm 5\%$  from the mean of the valuations, 95% within  $\pm 10\%$ , and 100% within  $\pm 15\%$ . This time, our computer simulations indicate the SD of the nearest normal pdf equivalent to be 2.3% (1.73 to 3.45% at the 95% confidence interval) for the 'simple' property and 4.8% (3.60 to 7.20% at the 95% confidence interval) for the 'complex' property, again significantly lower than the Crosby *et al.* and Adair *et al.* results. Again, this may reflect the relative ease of the task of residential property valuation when compared with the valuation of other property classes. The 'simple' result was even significantly lower than Diaz and Wolverton's result, indicating once more the effect of the frequently more straightforward process of residential valuation on valuation variability.

## **5. COMPONENTS OF VARIANCE FOR COMMERCIAL PROPERTIES IN GENERAL**

The above analysis suggests that:

1. the IPD statistics are representative of the GDs in commercial properties in general.
2. the Blundell and Ward and Newell and Kishore statistics, though significantly different to each other, are representative of the NDs in commercial properties in general.

3. the Adair *et al* rack rented intervaluer results, though significantly different to those of the reversionary valuations, can be discounted due to unusual market conditions.

### 5.1 Lag variance

At the time the Blundell and Ward data was collected, over the period 1974 to 1990, it is assumed, as a first approximation, that the IPD GD statistics for the period also apply. Unfortunately, the IPD series only started in 1982, so the best that can be done is to use the statistics for the period 1982 to 1990 instead. The average GD SD over the 1982 to 1990 period for the IPD data is 21.3 (variance=21.3<sup>2</sup>=454), in contrast with Blundell and Ward's ND estimated SD of 13.0 (variance=169). The variance of the lag is therefore 454-169=285 (SD= $\sqrt{286}$  =16.9). Similarly, the average GD SD over the 1987 to 1996 period for the IPD data is 18.1 (variance=328), in contrast with Newell and Kishore's ND estimated SD of 11.0 (variance=121), giving a variance of the lag of 328-121=207 (SD= $\sqrt{207}$  =14.4). Table 5 summarises this for Blundell and Ward, Newell and Kishore and Parker's results.

| Component | Blundell & Ward |     | Newell & Kishore |     | Parker |     |
|-----------|-----------------|-----|------------------|-----|--------|-----|
|           | SD              | var | SD               | var | SD     | var |
| GD        | 21.3            | 454 | 18.1             | 328 | 17.9   | 320 |
| Lag       | 16.9            | 285 | 14.4             | 207 | 15.5   | 239 |
| ND        | 13.0            | 169 | 11.0             | 121 | 9.0    | 81  |

Table 5: Estimation of lag variance

Now, there is no logical reason why the lag variance should vary over time because every precaution has been taken in the studies to avoid the valuation being affected in any way by the sale price. In other words, at least two of the three estimates shown in Table 5 must be incorrect. Clearly, one of these must be Parker's, as the range given for possible true ND SD values encompasses both Blundell and Ward and Newell and Kishore's ND SDs. The clearer candidate for the second incorrect value is Blundell and Ward's result as, by their own admission, there are 'deficiencies in the relatively crude adjustment approach applied'. This leaves the Newell and Kishore variance of 207 (14.4 SD) as the most accurate. However, applying this lag variance of 207 to the Blundell and Ward result implies either a GD SD of 19.5 instead of 21.3, or ND SD of 15.7 instead of 13.0. Although some of the IPD GD SD values were missing in this series, it was not felt the 21.3 would be an overestimate, particularly in view of the apparent declining GD SD values in general over the years. Neither does it seem likely that Blundell and Ward could have so grossly underestimated the ND SD by 2.7 points. Bearing in mind that the Newell and Kishore data was adjusted by the Property Council of Australia Indices, which in itself is a source of some error, it seems likely that the true lag SD value will be somewhere above Newell and Kishore's 14.5 and below the GD SD of 15.3 in 1996-7.

### 5.2 Residual (unpredicted) actual value variability

In a similar way, the components of variance can be used to estimate the amount of the variability in NDs that is explained by intravaluer variability and the residual (unpredicted) variability of the actual values themselves.

Utilising Crosby *et al.* and Adair *et al.*'s contribution, the intervaluer SD was estimated at 7.0 and 8.6 (overall) respectively. This raises the question as to what might be a suitable intravaluer SD. Clearly, it is to be expected that some valuers are more consistent than others, which means that the intervaluer results of Crosby *et al.* and Adair *et al.* must be regarded as the average of these. In the absence of any information concerning the individual valuers themselves we are forced to assume that all intravaluer SDs are equal and, therefore, the intervaluer SD obtained via these studies is the intravaluer SD.

Table 6: Estimation of residual (unpredicted) variability of actual value

| Component   | Crosby <i>et al.</i> (1983-5) |     | Adair <i>et al.</i> (1995) |     |
|-------------|-------------------------------|-----|----------------------------|-----|
|             | SD                            | Var | SD                         | Var |
| GD          | 22.7                          | 515 | 17.9                       | 320 |
| Lag         | 15.0                          | 225 | 15.0                       | 225 |
| ND          | 17.0                          | 290 | 9.7                        | 95  |
| Intravaluer | 7.0                           | 49  | 8.6                        | 74  |
| Residual    | 15.5                          | 241 | 4.6                        | 21  |

Now, the ND variability comprises two components (1) the variability of the value estimates and (2) the variability of the unpredicted portion of actual values, or residual (unpredicted) actual value variability. For the Crosby *et al.* data, collected over 1983-1985, this can be estimated by taking the average IPD GD SD of 22.7 (515 variance), from which the lag variance of, say, 225 (15.0 SD), is subtracted to give a ND variance of 290 (17.0 SD). From this the Crosby *et al.* intravaluer variance of 49 (7.0 SD) is also subtracted to find a residual (unpredicted) actual value variance of 241 (15.5 SD). The results are summarised in Table 6, together with the Adair *et al.* equivalent. They suggest quite strongly that, although the ND SD has dropped considerably between 1983-5 and 1995, with the intravaluer variability remaining constant, if not increasing, the drop is mostly due to the residual variance over the period between the two studies.

Table 7: Estimated  $\pm$  NDs

| Year | GDs    |       |     | Lag  |     | NDs         |      |           | Approx $\pm$ equivalent<br>( $\sqrt{3m}$ ) |      |
|------|--------|-------|-----|------|-----|-------------|------|-----------|--|------|
|      | Mean   | SD    | Var | SD   | Var | Var ( $v$ ) | SD   | $\bar{x}$ |  |      |
| 1982 | -11.33 | 23.07 | 532 | 15.0 | 225 | 307         | 17.5 | -2.0      | 311  | 30.5 |
| 1983 | -9.59  | 24.36 | 593 | 15.0 | 225 | 368         | 19.2 | -2.0      | 372  | 33.4 |
| 1984 | -7.09  | 20.66 | 427 | 15.0 | 225 | 202         | 14.2 | -2.0      | 206  | 24.9 |
| 1985 | -7.40  | 22.95 | 527 | 15.0 | 225 | 302         | 17.4 | -2.0      | 306  | 30.3 |
| 1986 | -4.08  | 22.76 | 518 | 15.0 | 225 | 293         | 17.1 | -2.0      | 297  | 29.8 |
| 1987 | -12.03 | 19.79 | 392 | 15.0 | 225 | 167         | 12.9 | -2.0      | 171  | 22.6 |
| 1988 | -17.34 | 17.79 | 316 | 15.0 | 225 | 91          | 9.5  | -2.0      | 95   | 16.9 |
| 1989 | -18.42 | 19.19 | 368 | 15.0 | 225 | 143         | 12.0 | -2.0      | 147  | 21.0 |
| 1990 | -1.74  | 21.43 | 459 | 15.0 | 225 | 234         | 15.3 | -2.0      | 238  | 26.7 |
| 1991 | -1.84  | 22.09 | 488 | 15.0 | 225 | 263         | 16.2 | -2.0      | 267  | 28.3 |
| 1992 | -0.15  | 17.63 | 311 | 15.0 | 225 | 86          | 9.3  | -2.0      | 90   | 16.4 |
| 1993 | -2.44  | 18.52 | 343 | 15.0 | 225 | 118         | 10.9 | -2.0      | 122  | 19.1 |
| 1994 | 0.0    | 16.13 | 260 | 15.0 | 225 | 35          | 5.9  | -2.0      | 39   | 10.8 |
| 1995 | 0.0    | 17.87 | 319 | 15.0 | 225 | 94          | 9.7  | -2.0      | 98   | 17.1 |
| 1996 | 0.0    | 15.30 | 234 | 15.0 | 225 | 9           | 3.0  | -2.0      | 13   | 6.2  |
| 1997 | 0.0    | 15.30 | 234 | 15.0 | 225 | 9           | 3.0  | -2.0      | 13   | 6.2  |
| 1998 | 0.0    | 16.34 | 267 | 15.0 | 225 | 42          | 6.5  | -2.0      | 46   | 11.7 |

Using the results of the above analyses, it is now possible to make an approximation of the NDs that have occurred over the period of the IPD data. Table 7 gives the estimated yearly variance by deducting the estimated 15% lag SD (225 variance) from the GD variance. Assuming Newell and Kishore's mean ND of -2%, this is then squared and added, from which the estimated  $\pm$  equivalent is obtained. This indicates a high of over  $\pm 30\%$  (maximum of  $\pm 33.4$  in 1983) in the early 1980s to the recent low of around  $\pm 6\%$  in the late 1990s, with a general downward trend present over the period. Of course, as a first order approximation, it is possible that the estimated  $\pm 6.2\%$  in the late 1990s was

actually the expected  $\pm 5\%$ , giving a mean square of  $m=8.3$  and lag variance of 4, leaving the maximum in the early 1980s at still over  $\pm 30\%$ , a result which clearly suggests  $\pm 5-30\%$  to be a more realistic range than  $\pm 5-10\%$ .

## 6. CONCLUSIONS

The aim of this paper was to provide a first order approximation of the accuracy of property valuations for comparison with the  $\pm 5-10\%$  threshold of tradition (Brown *et al.*, 1998). The nature of ranges was examined in relation to the uniform and normal probability density functions and the effects of bias considered. The summary statistics were examined for GDs (differences between property valuations and subsequently realised transaction prices) recorded in the Investment Property Databank (IPD) database and significant yearly changes noted in both the means and SDs. A meta analysis of previous work was conducted, which showed that all other results involving commercial GDs are consistent with statistics yielded by the IPD database. This confirms our *a priori* hypothesis that the statistics relating to the IPD database are generally consistent for valuations occurring outside the database. Of course, the application of this result cannot be expected to pertain beyond markets of similar levels of maturity. The results for the two main studies of NDs (GD adjusted for the lag in time between the valuation and transaction), with an improvement from Blundell and Ward's 1974-90 period (13% SD) to Newell and Kishore's 1987-96 period (11% SD), also suggested yearly trends of a similar nature to the GDs.

The role of intravaluer variability was also examined. Here, the variability associated with the lag was then estimated and applied to the IPD figures to give the approximation sought, suggesting  $\pm 5-30\%$  to be a realistic range in place of widely held perception of  $\pm 5-10\%$  (Brown *et al.*, 1998). As far as industrial properties are concerned, the analysis does not reveal any marked differences with commercial. This may be because, despite industrial being often more homogeneous and single tenanted, more relevant information often tends to be available for commercial. Residential valuations, on the other hand, from the very limited number of studies to date, are likely to be more accurate, probably due to being both homogeneous and information rich.

These results support the Carsberg report in emphasising the central role of the IPD and its development, the need to identify further information for publication concerning the composition and performance of valuers contributing to its indices, and the need to ensure that the knowledge gained from its use in research into the valuation process is fully integrated into the educational system. For future research on this topic, it would be beneficial to collect further data on NDs, especially in a non-hypothetical situation such as Newell and Kishore's study. This would enable this type of variability to be measured more precisely as a check on the work described here. Further useful work could also be undertaken through a closer inspection of the temporal (year on year) and sectoral differences, found here in Diaz and Wolverton and Daniel's studies of residential property and previously noted in the Drivers Jonas' comparison of retail and office/industrial properties<sup>4</sup>. As a result, it may then be possible to consider questions concerning the extent to which valuers' methods, research and abilities are improving.

## 7. REFERENCES

- Adair, A., Hutchison, N., MacGregor, B., McGreal, S. and Nanthakumaran, N. (1996) An analysis of valuation variation in the UK commercial property market: Hager and Lord revisited, **Journal of Property Valuation and Investment**, 14(5), 34-47.
- Blundell, G.F. and Ward, C.W.R. (1997) **The accuracy of valuations – expectation and reality**. Unpublished paper, The University of Reading.
- Brown, G. (1985) Property investment and performance measurement: a reply, **Journal of Valuation**, 4, 33-44.

---

<sup>4</sup> Also touching on this issue are the differences between Crosby et al's (1985) data and Adair et al's (1995) data and their relationship with temporal changes in the IPD accuracy data.

- Brown, G.R., Matysiak, G.A. and Shepherd, M. (1998) Valuation uncertainty and the Mallinson Report, **Journal of Property Research**, 15(1), 1-13.
- Carsberg, B. (2002) **Property valuation: the Carsberg report**. Royal Institution of Chartered Surveyors, London.
- Crosby, N., Lavers, A. and Murdoch, J. (1998) Property valuation variation and the 'margin of error' in the UK, **Journal of Property Research**, 15(4), 305-30.
- Cullen, I. (1994) The accuracy of valuations revisited, **Proceedings**, The Cutting Edge Conference, The Royal Institution of Chartered Surveyors, London, pp. 91-101.
- Daniels, P. (1983) Business serviced offices in British provincial cities: location and control. **Environmental and Planning A**, 5 1101-20.
- Diaz, J. and Wolverton, M. (1998) A longitudinal examination of the appraisal smoothing hypothesis, **Real Estate Economics**, 26(2), 349-359.
- Drivers Jonas and Investment Property Databank (2000) **Variance in valuations**, Drivers Jonas and Investment Property Databank, London.
- Fenton Nominees Pty Ltd v Valuer-General* (1981) 47 LGRA 71.
- Hager, D. and Lord, D. (1985) **The property market, property valuations and property performance measurement**, Institute of Actuaries, London.
- Mallinson, M. (1994) **Commercial property valuations: report of the president's working party**, The Royal Institution of Chartered Surveyors, London.
- Matysiak, G.A. and Wang, P. (1995) Commercial property market prices and valuations: analysing the correspondence, **Journal of Property Research**, 12, 181-202.
- Miles, M., Guilkey, D., Webb, B. and Hunter, K. (1991) **An empirical evaluation of the reliability of commercial appraisals**, National Council of Real Estate Investment Fiduciaries, Chicago.
- Morgan, E. (1993) Are valuations a good proxy for prices? Unpublished dissertation, Oxford Brookes University, UK.
- Newell, G. and Kishore, R. (1998) The accuracy of commercial property valuations, **Proceedings of the 4th Pacific Rim Real Estate Society Conference**, Perth.
- Parker, D. (1998) Valuation accuracy – an Australian perspective, **Proceedings of the 4th Pacific Rim Real Estate Society Conference**, Perth.
- Parker, D. (1999). A note on valuation accuracy: an Australian case study. **Journal of Property Investment and Finance**, 17(4), 401-411.
- Popper, K. (1959) **The logic of scientific discovery**. London: Hutchinson.
- Reid, I. (1985) Response to Hager/Lord. **Estates Gazette**. 274 19-21.
- Rossini, P. (1999) Accuracy issues for automated and artificial intelligent residential valuation systems, **Proceedings of the International Real Estate Society Conference**, Kuala Lumpur.
- Webb, B. (1994) On the reliability of commercial appraisals: An analysis of properties sold from the Russell-NCREIF Index, **Real Estate Finance**, 11(1), 62-6.