

FEATURE

Joe Robjohns

HM Treasury (formerly of Office for National Statistics)

Damian Whittard

Office for National Statistics

How similar are ONS's annual and monthly business inquiries?

SUMMARY

This article supports the Office for National Statistics' monthly and annual surveys reconciliation programme. It presents an analysis of the coherence of business survey returns between the Monthly Inquiry into the Distribution and Service Sector (MIDSS) and Monthly Production Inquiry (MPI) compared with the Annual Business Inquiry. A two-stage methodology is employed: firstly, analysing the aggregate data and secondly, using microdata sets matching individual business responses from the monthly and annual surveys. This analysis is reported in terms of levels and growth rates for both an unadjusted and an adjusted MPI/MIDSS series.

The Office for National Statistics (ONS) is responsible for producing different vintage estimates of gross domestic product (GDP).¹ ONS publishes four revised estimates of GDP; the first estimate of GDP output is published around 25 days after the end of the quarter (Robinson 2005) while the second estimate is published four and a half weeks later. The Monthly Production Inquiry (MPI) and Monthly Inquiry into the Distribution and Service Sector (MIDSS) are used in the production of the preliminary estimates of GDP, before the Annual Business Inquiry (ABI) is incorporated in the third estimate. The fourth estimate and first benchmark of GDP is published in the National Accounts *Blue Book* after a lag of approximately 18 months, although this estimate can be revised through stalled revisions or methodological changes (Mahajan 1997).

In order to ensure the quality of the GDP estimates and the data sources that go into producing them, ONS has implemented a programme of analysis to explore whether the data are subject to any bias. This article adds to the growing literature by comparing results from the MPI and MIDSS monthly surveys with the ABI annual survey. The article supports ONS's monthly and annual surveys' reconciliation programme started in 2000 and which was last completed for the release of the *Blue Book 2005*.

Analysis methodology

In order to assess the coherence between the monthly surveys and ABI turnover, a two-stage methodology was employed.

Firstly, an analysis of ONS's final aggregated estimate of monthly data was compared with the ABI annual series. This analysis is reported in levels and growth rates for both an unadjusted (raw data) series and an adjusted MPI/MIDSS series. Adjustments are in the main applied to manage survey design issues such as sample rotation and large reclassifications; this is done to help preserve a continuous time series of growth (Duff and Morgan 2007). The second stage uses ONS's Virtual Microdata Laboratory² to analyse a business unit's monthly response with the same unit's annual response. The microdata sets were produced using a matching technique such that businesses are only maintained in the overlap if they represent a full year of monthly data that correspond to the same period recorded in the ABI. For example, a business reporting from April to March in the ABI but only from April to December in the monthly surveys would be dropped from the overlap. The results are presented using identical methodologies, but are reported for the MPI and MIDSS separately – a full description of MPI and MIDSS sectors with their corresponding Standard Industrial Classification (SIC) codes is listed in Appendix **Table A1** and **Table A2**.

MPI results

Aggregate analysis

In this analysis, the published aggregate ABI data are used as the benchmark. The analysis is restricted to the years 1999 to 2005, as there were major differences in the sampling regime for the ABI before

then (Jones 2000). The aggregate analysis explores the issue of bias in ABI and MPI turnover in terms of nominal levels and growth rates.

Nominal values

Table 1 records aggregate annual turnover for both the ABI and MPI and includes a ratio of MPI to ABI turnover. A ratio of less than one indicates that the ABI turnover is higher than that for the MPI.

A key finding here is that there is a consistent underestimation of annual turnover in the MPI series compared with the ABI. This under-reporting is estimated to be more than 10 per cent of turnover in all of the seven years – 73 per cent of this difference is attributed to two industries, ‘food products, beverages’ (SIC 15) and ‘coke, refined petroleum products and nuclear fuel’ (23). Of the two MPI series, adjusted MPI turnover is closer in terms of level of expenditure to the ABI series.

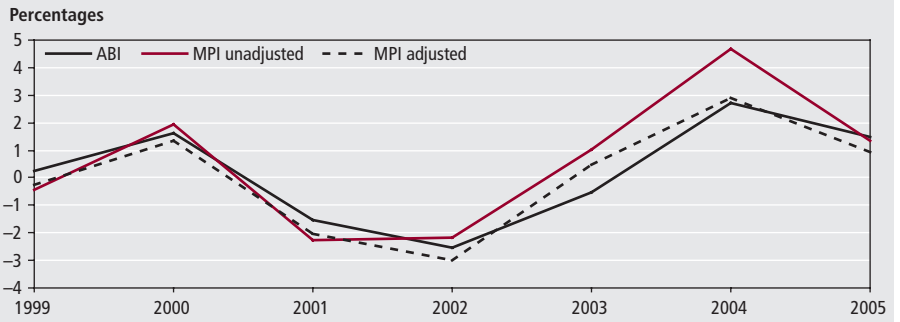
Growth rates

The level of turnover is important inasmuch as it provides a view of the economy at a particular point of time. However, of more importance is the growth rate, as it determines the long-term trend and indicates the strength within the sector as a whole. **Figure 1** shows the growth rates of ABI and MPI turnover.

The MPI growth rates are closely aligned to those of the ABI for all data points. This suggests that the MPI shares the same statistical signal as the ABI and is unbiased. In five of the seven data points, the MPI adjusted growth rate is more closely aligned to the ABI series than the adjusted growth rate, although in three of these data points the difference is only marginal.

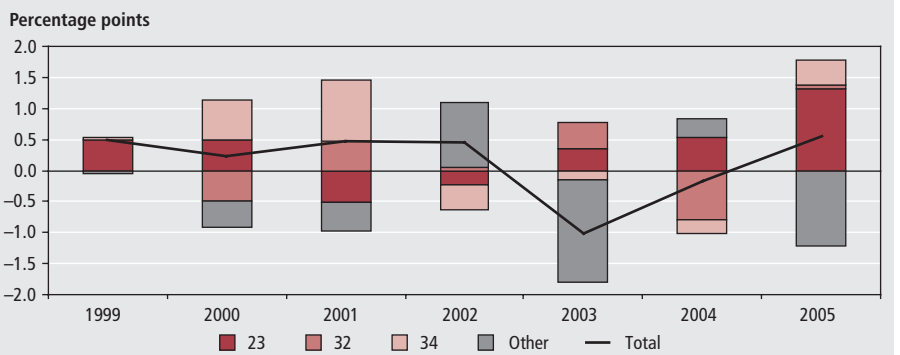
To assess the impact of individual industries on the aggregate position, **Figure 2** decomposes the differences in the ABI and adjusted MPI turnover growth rate at the industrial level.

Figure 1
MPI and ABI annual turnover growth rates



Source: Office for National Statistics

Figure 2
Contributions to differences in ABI and adjusted MPI turnover growth rates



Source: Office for National Statistics

There is no evidence of a systematic bias in any single industry. Figure 2 reports a counterbalancing effect: an inflated growth rate in one industry is offset by a deflated growth rate in another. ‘Coke, refined petroleum products and nuclear fuel’ (SIC 23) and ‘motor vehicles, trailers and semi-trailers’ (34) have the largest single differences of all the industries – they are the fifth and third largest industries in terms of output.

Noise-to-signal ratio

Table 2 expands the analysis of the congruence by comparing the unadjusted

and adjusted noise-to-signal ratio. This is calculated by comparing the difference in the variation of MPI and ABI turnover growth rates with the variation of ABI turnover growth rates (see below) – the closer the ratio to zero, the greater the statistical signal.

MPI noise-to-signal ratio

$$= \frac{\text{var}\left(\frac{MPI^t - MPI^{t-1}}{MPI^{t-1}}\right) - \text{var}\left(\frac{ABI^t - ABI^{t-1}}{ABI^{t-1}}\right)}{\text{var}\left(\frac{ABI^t - ABI^{t-1}}{ABI^{t-1}}\right)}$$

- var variance
- MPI MPI turnover
- ABI ABI turnover
- t time (year)

The ratios report a stronger signal in the adjusted series in terms of the manufacturing sector as a whole and also in the majority of two-digit SIC industries. The overall difference in variation is 0.54 points; however, this is considerably affected by the relatively poor performing industries of ‘pulp, paper and paper products’ (two-digit SIC20) and ‘recycling’ (37). At an industrial level, the adjusted series outperforms the unadjusted series in 12 of the 23 two-digit

Table 1
ABI and MPI aggregate turnover

	Turnover (£ million)			Ratio (MPI/ABI)	
	ABI	MPI unadjusted	MPI adjusted	MPI unadjusted	MPI adjusted
1999	461,771	395,350	412,066	0.86	0.89
2000	469,146	403,011	417,700	0.86	0.89
2001	461,898	393,911	409,270	0.85	0.89
2002	450,090	385,362	396,924	0.86	0.88
2003	447,637	389,232	398,825	0.87	0.89
2004	459,880	407,495	410,437	0.89	0.89
2005	466,731	413,013	414,269	0.88	0.89

Source: Author's calculation on ONS sources

Table 2

Noise-to-signal ratio between the adjusted and unadjusted MPI time series and ABI growth rates, 1999 to 2005

SIC2	Unadjusted	Adjusted
15	4.25	1.19
16	0.94	1.04
17	4.04	1.59
18	1.83	1.78
19	-0.12	-0.65
20	9.46	1.97
21	-0.67	-0.83
22	-0.66	-0.79
23	0.58	1.55
24	-0.45	0.59
25	0.79	-0.46
26	-0.88	-0.79
27	-0.14	-0.46
28	3.30	0.89
29	0.68	0.28
30	-0.20	-0.48
31	0.81	0.03
32	0.11	0.42
33	0.49	0.07
34	2.54	3.13
35	0.61	1.10
36	3.08	-0.57
37	23.36	1.39
All industries	0.70	0.16

Source: Office for National Statistics

SIC industries; in six of the industries (15, 17, 20, 28, 36 and 37), the absolute differences are over two points. In 11 of the 23 two-digit SIC industrial categories, the unadjusted series outperforms the adjusted series. These 11 industries account for over 50 per cent of manufacturing turnover. In four of the industries (19, 23, 34 and 35), the unadjusted data outperforms the adjusted series by 0.5 or more points.

Overall, the results indicate that the adjusted series are more closely aligned to the ABI series in nominal levels, growth rates and the noise-to-signal ratio. However, the unadjusted series performs better in terms of turnover growth rates in two of the seven years and contains a greater statistical signal for 11 two-digit SIC industries which, combined, account for over 50 per cent of manufacturing output. This is an important finding of the research – if the ABI is taken as the benchmark, then the noise-to-signal ratio results suggest that adjusting the figures in themselves may not add value for certain industries.

As a result of ONS's research programme, a decision was recently made to revamp the Index of Production – ONS has now moved considerably away from adjustments, with

most of the historical adjustments actually being removed. The findings from this article provide support for this change of approach, particularly at the industrial level.

The following analysis uses microdata to further explore the relationship between ABI and MPI returns, by analysing individual business responses.

Microdata analysis

This section considers the degree of fit between the reported turnover of the MPI and ABI surveys from individual businesses. The analysis uses formal measures of correlation and regression analysis. All analysis was completed using natural logarithms to 'normalise' the data, and hence improve its statistical properties.

Table 3 details the number of businesses that completed 12 months' worth of MPI returns and a matched annual ABI return, from 1999 to 2005. Just 2 per cent (150,000) of all manufacturing businesses, but one-third (7,500) of all businesses included in the overall MPI sample, are included within the ABI/MPI matched sample. The 'moving target' nature of the MPI survey means that companies are only included in the matched sample if their monthly MPI returns directly correspond with the exact period of their ABI return.

As a deliberate part of the sampling strategy, the ABI and MPI samples are individually skewed towards large employers. **Table 4** shows that a relatively low proportion of businesses are included in the sample, although they contribute a high percentage of total employment and, more often than not, turnover. With a relatively high proportion of sector employment included within the sample, it is possible to make inferences for the overall trend.

The coverage varies widely across industries. The largest two industries in terms of turnover, 'food products, beverages' (SIC 15) and 'chemical and chemical products' (24), are generally well covered in terms of both business count and employment. 'Fabricated metal products, except machinery and equipment' (28), which provides nearly 6 per cent of manufacturing output, however, is only covered in 1 per cent of businesses and 15 per cent of employment.

The ratios in **Table 5** show that even when business units are matched directly, there is a consistent tendency to underestimate MPI turnover compared with ABI by over 15 per cent on average (2005 is an outlier attributable to 'coke, refined petroleum products and nuclear

Table 3

Frequencies of matched businesses in the overlap microdata set

	Frequency (number)	Percentages
1999	3,417	2.0
2000	3,290	2.0
2001	3,227	2.0
2002	3,060	1.9
2003	2,909	1.9
2004	2,938	1.9
2005	2,696	1.8
Total	18,009	

Source: Office for National Statistics

Table 4

Percentage of total business and employment contained in the overlap data set, 1999 to 2005

SIC2	Percentages	
	Firm count	Employment
15	5	38
16	23	53
17	3	29
18	*	21
19	*	29
20	*	16
21	5	38
22	1	30
23	2	42
24	8	55
25	3	32
26	3	46
27	5	25
28	1	15
29	2	33
30	*	39
31	3	35
32	4	53
33	2	35
34	5	25
35	4	56
36	1	25
37	*	23

Note: Source: Office for National Statistics

* Removed due to ONS rules on disclosure.

Table 5

Ratio of MPI to ABI turnover

	Ratio MPI/ABI
1999	0.82
2000	0.83
2001	0.84
2002	0.85
2003	0.83
2004	0.85
2005	0.76

Source: Office for National Statistics

fuel' (23)). This is consistent with the key finding at an aggregate level, and supports the view that the lower recording of turnover in the MPI is a feature of the unit-level results, rather than associated with sampling or grossing methodology.

At an industrial level, 'food products, beverages' (SIC 15), 'coke, refined petroleum products and nuclear fuel' (23) and 'motor vehicles, trailers and semi-trailers' (34) are all systematically under-reported in nominal terms. Together, these industries account for over 74 per cent of the total difference.

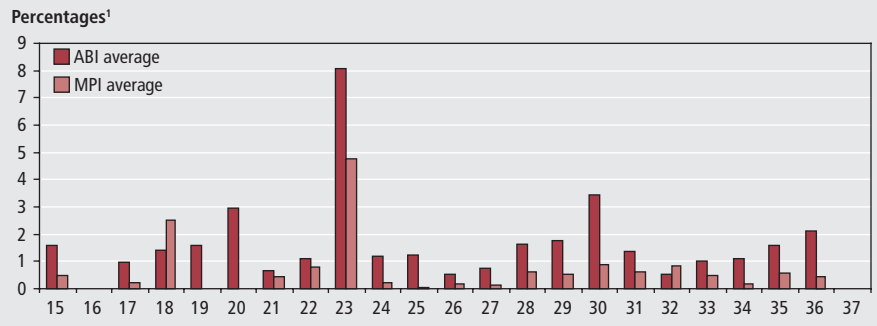
Correlation and regression analysis

It is possible to test formally the relationship between log ABI turnover and log MPI turnover by calculating a correlation coefficient. The Kendall tau rank correlation coefficient³ is used to measure the degree to which two variables are linearly related. A positive value indicates that the variables move in the same direction and a negative one in the opposite direction. A correlation coefficient of +1 or -1 means that there is a perfect positive/negative linear relationship; 0 indicates that there is no correlation.

The results of the correlations are presented in Table 6. The table shows that

Figure 3

Percentage of businesses that are ABI or MPI dominant outliers, 1999 to 2005



Note:

1 Observations with log ratio >0.5 or <-0.5.

Source: Office for National Statistics

MPI turnover is highly correlated for all manufacturing industries and is generally improving over time. As such, MPI turnover can be said to be a valid predictor of ABI turnover.

Quantile regressions were also used as a pseudo-correlation test to test the robustness of the correlations and further evaluate the extent to which two variables are similar. The full results are shown in Appendix Table A3. They confirm that the MPI is generally a good predictor of the ABI and demonstrate that the centre of the distributions of both ABI and MPI returns are unbiased.

Sensitivity analysis

In order to explore which industries were most severely affected by individual companies reporting the largest difference in ABI and MPI returns, and the extent to which the results were affected by extreme values, MPI and ABI dominant outliers were constructed. Initially, a log ratio of MPI and ABI turnover was calculated; Figure 3 plots the proportion of observations that had a log ratio of greater than 0.5 or less than -0.5 (a 50 per cent difference between the two). A ratio of greater than 0.5 indicates an ABI dominant outlier and a ratio of less than -0.5 identifies an MPI dominant outlier.

There are more ABI-dominant outliers in the majority of industries. These outliers are partially responsible for the evident discrepancy in the levels of the ABI and MPI at the aggregate level. At an industrial level, 'coke, refined petroleum products and nuclear fuel' (SIC 23) had proportionately the largest number of ABI and MPI dominant outliers.

The following analysis looks at the impact of the outliers between the two surveys on the aggregate growth series. For this analysis, a simple weighted methodology is used to produce an MPI-equivalent time series from the microdata.⁴ The same firms are then subtracted from the MPI and ABI samples and weighted up to examine the aggregate effect of the subtraction of outliers at the 1, 5 and 10 per cent tails of the distribution. These aggregates are then compared to observe what difference eliminating outliers has on the aggregate growth rates.

Figure 4 and Figure 5, respectively, chart the difference in the growth rates for the MPI and ABI microdata excluding outliers. The graphs plot the difference by taking the overall growth rate including all observations and subtracting the growth

Table 6

Kendall tau rank correlation coefficient for MPI and ABI turnover returns

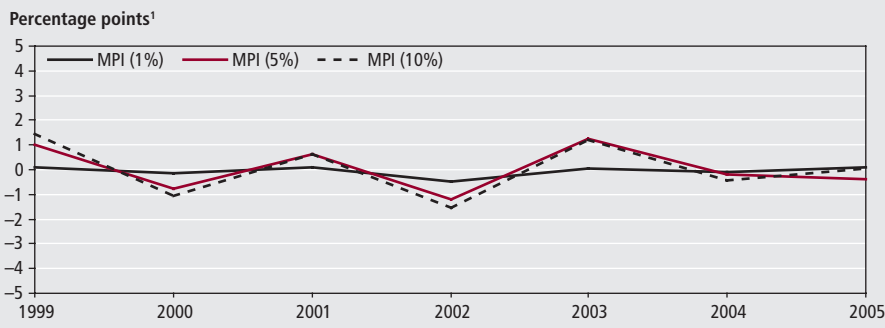
SIC2	1999	2000	2001	2002	2003	2004	2005
15	0.87	0.87	0.88	0.90	0.89	0.88	0.89
16	*	*	*	*	*	*	*
17	0.81	0.81	0.79	0.82	0.83	0.86	0.87
18	0.83	0.79	0.81	0.83	0.79	0.84	0.85
19	0.72	0.76	0.75	0.83	0.80	0.81	0.84
20	0.86	0.80	0.78	0.82	0.82	0.88	0.87
21	0.89	0.88	0.88	0.89	0.88	0.87	0.87
22	0.82	0.85	0.85	0.85	0.87	0.84	0.87
23	*	*	*	*	*	*	*
24	0.88	0.87	0.88	0.88	0.89	0.90	0.87
25	0.87	0.86	0.86	0.87	0.87	0.86	0.88
26	0.88	0.90	0.90	0.90	0.91	0.91	0.91
27	0.87	0.88	0.88	0.90	0.88	0.89	0.90
28	0.85	0.86	0.84	0.85	0.86	0.85	0.86
29	0.81	0.82	0.81	0.82	0.83	0.85	0.84
30	0.85	0.89	0.85	0.83	0.83	0.82	0.80
31	0.80	0.80	0.81	0.85	0.84	0.84	0.85
32	0.88	0.88	0.85	0.84	0.85	0.83	0.87
33	0.78	0.81	0.81	0.84	0.83	0.83	0.84
34	0.83	0.82	0.83	0.84	0.84	0.86	0.84
35	0.82	0.83	0.80	0.84	0.84	0.87	0.88
36	0.78	0.81	0.81	0.80	0.80	0.84	0.84
37	0.80	0.76	0.82	0.77	0.83	0.78	0.79

Note:

* Removed due to ONS rules on disclosure.

Source: Office for National Statistics

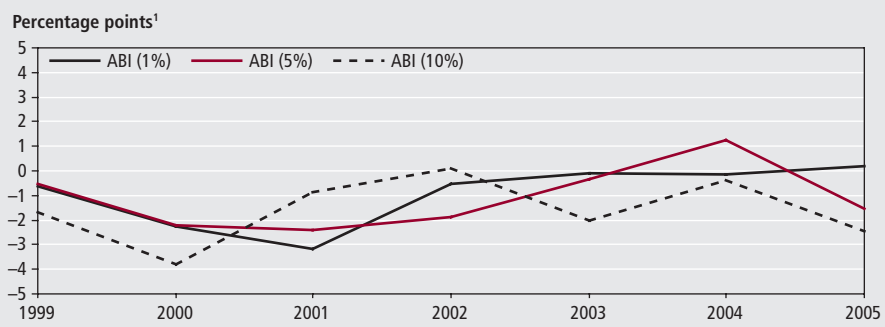
Figure 4
Percentage difference in MPI microdata turnover annual growth rates after excluding outliers

**Note:**

Source: Office for National Statistics

1 Difference from overall growth rate.

Figure 5
Percentage difference in ABI microdata turnover annual growth rates after excluding outliers

**Note:**

Source: Office for National Statistics

1 Difference from overall growth rate.

rate generated after excluding outliers from the 1, 5 and 10 per cent tails. For example, if the difference is zero, this means that omitting outliers has no effect on the overall growth rate.

By comparing Figure 4 and Figure 5, it is clear that the ABI data set was more affected by omitting outliers. Overall, however, omitting outliers from both data sets had relatively little impact on estimating growth rates. In both the MPI and ABI series, there is no evidence of any bias. This suggests that, although some companies may report considerably different turnover values in their MPI and ABI returns, they have relatively little impact on growth rates.

MIDSS results

Aggregate analysis

The first strand of the aggregate analysis is to examine whether any bias exists in the levels and growth rates of the total series in both the MIDSS and ABI turnover.

Nominal values

Table 7 compares the grossed total turnover implied by the MIDSS and ABI aggregate series. It is evident that the ABI produces

slightly higher levels than both the adjusted and the unadjusted MIDSS series but, unlike the MPI, it is consistently closer to the unadjusted figures. While the levels are of interest, it is the growth rates of these series which determine long-term trends.

Growth rates

The growth rates of total turnover are presented in Figure 6. An important point to note from this chart is that there is no evidence of bias in these growth rates between either of the MIDSS series and the ABI. It is also evident from these averages

that the unadjusted MIDSS series is closer to the ABI growth rate than the adjusted one.

Figure 7 graphs the difference decomposed into two-digit SICs. The MIDSS industries not explicitly referred to are captured as 'Other'.

The contributions depicted in Figure 7 demonstrate that there is no consistent direction of bias in any industry. SIC 51 (wholesale trade) contributes the largest difference to the ABI and MIDSS growth rate. The wholesale trade data can be considerably influenced by single large companies, particularly those working in commodities such as oil. These companies are generally multinational in nature and as such are exposed to changes in exchange rates. These changes can considerably affect the valuation of output over time.

Noise-to-signal ratio

As with the MPI analysis, an additional way of quantifying the congruence of two series is the noise-to-signal ratio. Table 8 presents the noise-to-signal ratios for both unadjusted and adjusted growth rates compared with those from the ABI. As described previously, the closer the ratio is to zero, the stronger the statistical signal.

Table 8 shows that the unadjusted MIDSS growth rates contain stronger statistical signals than the adjusted – this is true for the service sector as a whole and the majority of two-digit SIC industries. It is consistent with the aggregate and growth rate analysis reported in Table 7 and Figure 7, respectively, which showed that the unadjusted MIDSS data were better at tracking the ABI series. This is also supported from the finding in the MPI section, which reported an improved signal for certain two-digit SIC industries: together, these industries accounted for more than 50 per cent of turnover.

This is an important finding of the article and adds support to ONS's recent

Table 7
ABI and MIDSS aggregate turnover

	Turnover (£ million)			Ratio (MIDSS/ABI)	
	ABI	MIDSS unadjusted	MIDSS adjusted	MIDSS unadjusted	MIDSS adjusted
1999	862,281	842,494	782,246	0.98	0.91
2000	925,482	911,030	848,135	0.98	0.92
2001	970,158	948,385	894,417	0.98	0.92
2002	979,605	973,123	938,329	0.99	0.96
2003	1,051,876	1,013,367	987,251	0.96	0.94
2004	1,109,666	1,064,147	1,056,823	0.96	0.95
2005	1,194,590	1,157,803	1,149,570	0.97	0.96

Source: Office for National Statistics

Figure 6
MIDSS and ABI annual turnover growth rates

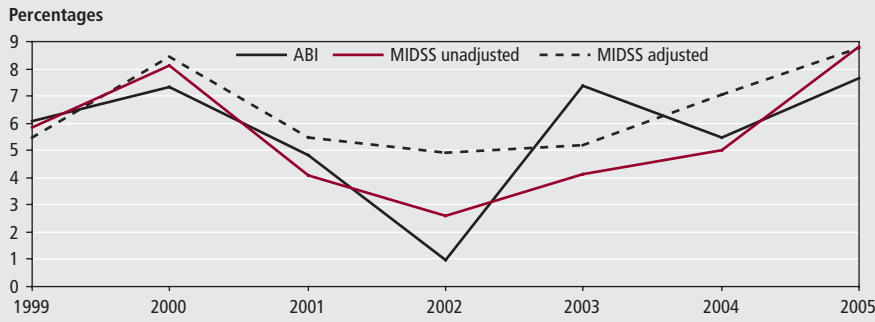


Figure 7
Contributions to differences in ABI and unadjusted MIDSS turnover growth rates

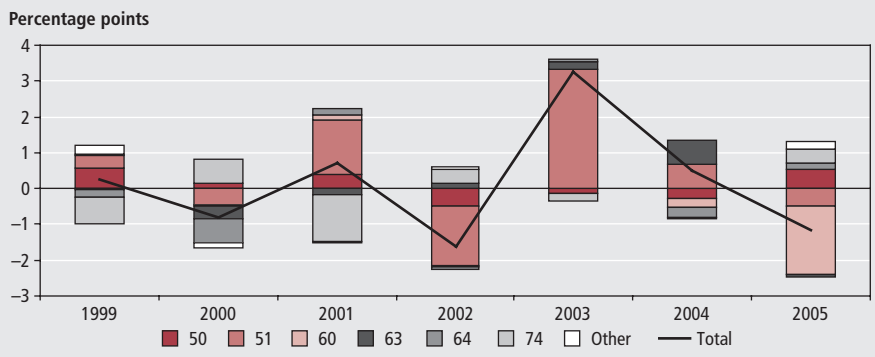


Table 8
Noise-to-signal ratio between the adjusted and unadjusted MIDSS time series and ABI growth rates, 1999 to 2005

SIC2	Unadjusted	Adjusted
50	1.60	2.23
51	0.20	0.41
55	-0.64	-0.49
60	-0.98	-0.99
63	-0.87	-0.83
64	0.10	0.86
71	-0.60	-0.58
72	-0.14	-0.58
73	-0.40	-0.29
74	0.03	-0.61
90	0.19	-0.35
92	-0.20	0.16
93	0.96	-0.53
Total	0.37	0.50

Source: Office for National Statistics

decision to substantially remove the MIDSS adjustments for the Index of Services data used in the compilation of the *Blue Book 2008*.

The next section uses business microdata to further investigate the relationship between ABI and MIDSS returns by analysing individual business responses.

Microdata analysis

This section examines individual businesses reporting turnover in the MIDSS and ABI surveys, using both correlation and regression analysis. As with the MPI micro results, all analysis was undertaken using natural logarithms.

For the MIDSS survey, there were fewer companies matched in the microdata set in the years before 2001. This is because new industries were added to the MIDSS survey in recent years. For example, in 2001, there was an improvement in the match between the ABI and MIDSS data sets when SIC 50, 55, 92 and 93 were introduced. As such, it is only possible to make robust comparisons for the microdata relating to the period 2001 to 2005.

Table 9 details the frequencies of businesses in the overlap data set in each year. The matched sample accounts for less than 1 per cent of all MIDSS businesses but approximately 15 per cent of all businesses included in the overall matched sample.

Perhaps more important than the business counts, in terms of the sufficiency of the overlap data set for supporting overall conclusions, is the proportion of total employment covered by the matched businesses. **Table 10** gives the percentage of employment: this is presented by year, there being some change on an annual basis as the survey changed structure.

Table 9
Frequencies of matched businesses in the overlap microdata set

	Frequency (number)	Percentages
2001	4,453	0.5
2002	4,541	0.5
2003	4,538	0.5
2004	4,439	0.5
2005	4,414	0.4
Total	28,473	0.4

Source: Office for National Statistics

Although the matched sample has low business coverage, the total employment coverage ranges between 29 and 37 per cent. This is indicative of the fact that it is predominantly larger businesses that are matched between the MIDSS and ABI surveys. The fact that roughly one-third of total employment is covered by the overlap means that inferences about overall trends are possible from this microdata analysis.

Correlation and regression analysis

Table 11 presents the correlation coefficients. It is evident that businesses in most industries seem to report broadly similar turnover in both surveys. The biggest industries, 51 and 74, have coefficients of around 0.95. Industries 63 and 73 are the only industries with coefficients consistently under 0.9, but these represent two of the smaller industries in terms of employment.

Quantile regressions were used to test, and subsequently confirmed, the robustness of the correlation results. The full regression results are reported in Appendix **Table A4**.

Sensitivity analysis

This section examines which industries are most affected by companies reporting the largest differences between the two surveys. **Figure 8** shows the percentage of businesses that are classed as severe outliers.

Figure 8 shows that there is a predominance of larger discrepancies for returns in industries 63, 73 and 92. SIC 73 'Research and development' and 92 'Recreational and sporting activities' also performed poorly in the regression analysis. Industry 63 'Supporting and auxiliary transport activities' performed poorly in the correlation analysis. In general, there are more ABI dominant outliers in most industries, again partially explaining the discrepancy in the levels of ABI and MIDSS at the aggregate level.

The following analysis examines the effect

Table 10
Percentage of total employment contained in the overlap data set

SIC2	Percentages				
	2001	2002	2003	2004	2005
50	23	25	28	30	31
51	30	28	31	34	34
55	16	38	32	33	36
60	49	45	47	47	48
63	45	43	43	54	58
64	59	62	63	64	74
71	38	34	32	31	35
72	23	27	25	28	33
73	40	33	30	45	42
74	30	31	34	34	35
90	50	46	61	57	49
92	27	34	38	35	25
93	17	19	18	18	17
Total	29	34	34	36	37

Source: Office for National Statistics

Table 11
Kendall tau rank correlation coefficients for MIDSS and ABI turnover returns

SIC2	2001	2002	2003	2004	2005
50	0.94	0.92	0.95	0.95	0.95
51	0.94	0.95	0.94	0.95	0.96
55	0.92	0.94	0.94	0.95	0.96
60	0.94	0.96	0.98	0.96	0.95
63	0.73	0.76	0.76	0.73	0.77
64	0.95	0.94	0.94	0.97	0.95
71	0.96	0.92	0.97	0.92	0.94
72	0.90	0.92	0.92	0.93	0.93
73	0.81	0.74	0.84	0.84	0.87
74	0.94	0.94	0.95	0.95	0.95
90	0.97	0.93	0.96	0.96	0.97
92	0.83	0.90	0.91	0.85	0.89
93	0.97	0.95	0.96	0.94	0.95

Source: Office for National Statistics

Figure 10, it is evident that the ABI data set is more affected by omitting outliers, at least in two of the four data points. However, in general, for both series, there is little change to the aggregate growth rates from taking out the businesses with large outliers. This supports the results of the MPI analysis and implies that, while there may be some businesses reporting very different turnover figures on the two surveys, this has little impact on the growth rates generated from the data.

Conclusion

For the MPI, the following conclusions can be drawn:

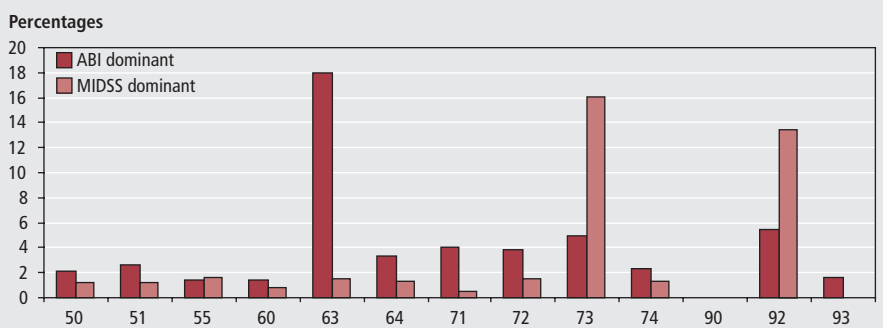
- nominal turnover is consistently under-reported in MPI monthly surveys compared with the ABI annual survey. The finding is consistent in both the aggregate and micro analysis
- MPI annual growth rates are closely aligned to ABI annual growth rates in all years – this is evident in both for the aggregate and microanalysis
- the evidence from the microdata suggests that MPI unadjusted data in a number of industries have greater statistical signal than the adjusted data – therefore it may be more appropriate to use the raw data in a number of cases. This finding will help to inform the nature of the adjustment regime, to ensure that it does not inadvertently introduce bias at the aggregate level

For the MIDSS, the following conclusions can be drawn:

- both MIDSS adjusted and unadjusted time series demonstrate a lower level of turnover than the ABI. The unadjusted series produces a slightly closer level than that including the adjustments
- the MIDSS growth rates are moderately close to the ABI growth rates for the equivalent service sector industries. There is no consistent bias in either the adjusted or unadjusted series evident in these growth rates
- the stronger statistical signal suggests that the unadjusted MIDSS series is a better indicator of ABI growth rates than the unadjusted series
- while some businesses report very different turnover figures to the MIDSS and ABI surveys, this has little impact on the growth rates generated from the data

The following joint conclusions can be drawn from the MPI and MIDSS:

Figure 8
Percentage of businesses that are ABI or MIDSS dominant outliers, 1999 to 2005



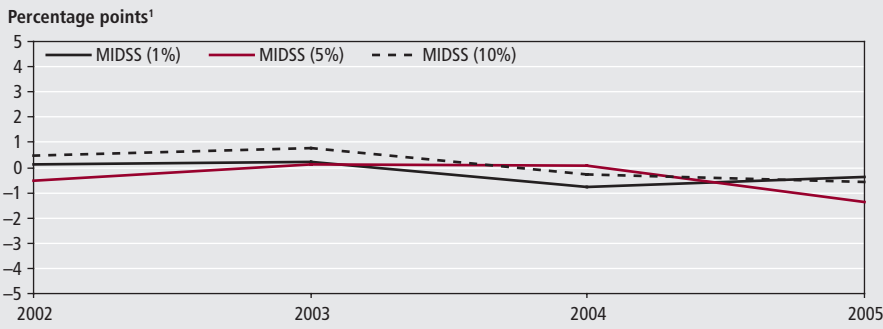
Note: Results for SIC90 have been removed due to ONS rules on disclosure. Source: Office for National Statistics

of the outliers between the two surveys on the aggregate series by constructing a time series from the microdata. **Figure 9** and **Figure 10** record the difference between the growth rate for the full matched sample and the growth rate after removing the top 1, 5

and 10 per cent of the tails. Figure 9 shows the difference in MIDSS growth rates as the businesses are subtracted; Figure 10 does the same for the ABI.

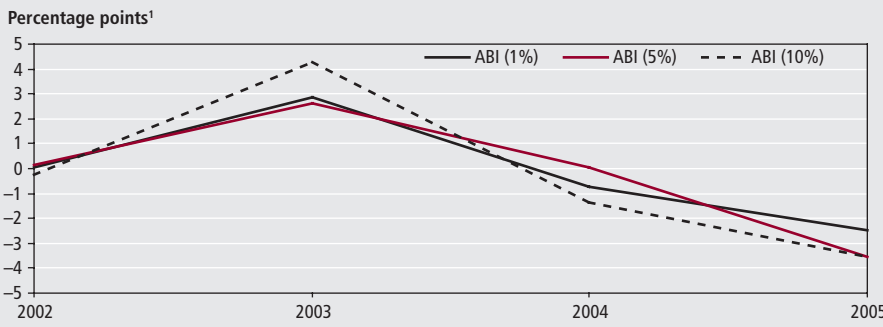
There is no evidence of any bias in either of the series. By comparing Figure 9 and

Figure 9
Percentage difference in MIDSS microdata turnover annual growth rates after excluding outliers



Note: Source: Office for National Statistics
 1 Difference from overall growth rate.

Figure 10
Percentage difference in ABI microdata turnover annual growth rates after excluding outliers



Note: Source: Office for National Statistics
 1 Difference from overall growth rate.

- matching businesses between monthly and annual surveys is feasible and produces a reasonable match, although almost entirely for larger businesses
- both MPI and MIDSS aggregate series do not produce biased growth rates when compared with ABI turnover. However, highlighted industries have been responsible for much of the difference
- businesses are generally good at reporting monthly and annual turnover consistently. There are examples of where this is not the case, but the aggregate impact would appear to be negligible

Notes

- 1 GDP can be measured using three separate approaches: production (GVA), income (GDP(I)) and expenditure (GDP(E)). GDP(O) is purely an output-based measure, which acts as a proxy for the production approach (GVA) in order to estimate quarterly GDP.
- 2 The Virtual Microdata Laboratory (VML) is a facility within the Office

for National Statistics (ONS) which enables access to restricted microdata for research purposes. Researchers from government and academia use the VML to undertake research on ONS surveys and other confidential data sets.

- 3 The Kendall tau rank correlation coefficient is a non-parametric statistic used to measure the degree of correspondence between two rankings and assess the significance of this correspondence.
- 4 A simple weighted aggregate is constructed by multiplying the turnover by the a-weight and the g-weight. The a-weight is the inverse of the sample fraction and is constructed at the level of 'year, two-digit SIC and sizeband'. The g-weight is a model weight calculated in order to account for differences in size of employment of those in and out of the sample.

ACKNOWLEDGEMENTS

The authors would like to thank Robin Youll, Felix Ritchie, Tony Clayton, High Skipper, David Hobbs, Sanjiv Mahajan and John Perry, for helpful comments. Any errors are the authors' sole responsibility.

CONTACT

✉ elmr@ons.gsi.gov.uk

REFERENCES

Drew S and Morgan D (2007) 'The launch of the Index of Services as a National Statistics', *Economic & Labour Market Review* 1(3), pp 39–46 and at www.statistics.gov.uk/ci/article.asp?id=1741

Duff S and Morgan D (2007) 'BERR Analysis: MIDSS v ABI Correlation – Early Findings', Office for National Statistics, unpublished.

Hobbs D (2008) 'What does 'turnover' mean?' Office for National Statistics, unpublished.

Jones G (2000) 'The Development of the Annual Business Inquiry', *Economic Trends* 564, pp 49–57 and at www.statistics.gov.uk/ci/article.asp?id=74

Mahajan S (1997) 'Balancing GDP: UK Annual Input-Output Balances', *Economic Trends* 519.

Mahajan S (2006) 'United Kingdom Input-Output Analyses, pp 211–59 at www.statistics.gov.uk/inputoutput

Robinson H (2005) 'Revisions to Quarterly GDP Growth and its Production (Output), Expenditure and Income components', *Economic Trends* 625, pp 34–49 and at www.statistics.gov.uk/ci/article.asp?id=1289

Smith J, Kelsoe R, Snell M, Wolley K and Drury S (2007) 'ABI2/MPI 2005 Congruence', Office for National Statistics, unpublished.

Taylor C (2004) 'Monthly Production Inquiry Process Documentation', Office for National Statistics, unpublished.

APPENDIX

Table A1

MPI industries

Manufacturing industry SIC (2003)	SIC
Food products, beverages	15
Tobacco	16
Textiles	17
Wearing apparel; dressing and dyeing of fur	18
Tanning and dressing of leather	19
Wood and of products of wood and cork, except furniture	20
Pulp, paper and paper products	21
Publishing, printing and reproduction of recorded media	22
Coke, refined petroleum products and nuclear fuel	23
Chemicals and chemical products	24
Rubber and plastic products	25
Other non-metallic mineral products	26
Basic metals	27
Fabricated metal products, except machinery and equipment	28
Machinery and equipment not elsewhere classified	29
Office machinery and computers	30
Electrical machinery and apparatus not elsewhere classified	31
Radio, television and communication equipment not elsewhere classified	32
Medical, precision and optical instruments, watches and clocks	33
Motor vehicles, trailers and semi-trailers	34
Other transport equipment	35
Manufacturing not elsewhere classified	36
Recycling	37

Table A2

MIDSS industries

Distribution and service sectors	SIC
Sale, maintenance and repair of motor vehicles	50
Wholesale trade and commission trade	51
Hotels and restaurants	55
Land transport; transport via pipelines	60
Supporting and auxiliary transport activities	63
Post and telecommunications	64
Renting of machinery and equipment	71
Computer and related activities	72
Research and development	73
Other business activities	74
Sewage and refuse disposal	90
Recreational, cultural and sporting activities	92
Extra-territorial organisations and bodies	93

Regression

A regression of a dependent variable upon an identical explanatory variable will yield the result of an intercept term equal to zero and a slope coefficient of one. Consequently, regressing ABI turnover against MPI/MIDSS turnover will produce coefficients close to these values if they are congruent variables. Cells marked with * in Table A3 denote that the values were not equal at a 5 per cent significant level.

Table A3

ABI and MPI median regressions where the joint-hypothesis test can be rejected at the 5 per cent significance level

SIC2	1999	2000	2001	2002	2003	2004	2005
15							
17							
18							*
19							
20							
21							
22							
24	*	*					
25	*	*					
26			*				
27							
28							
29						*	
30							
31		*		*			
32			*				
33							
34							
35			*				
36							
37							

Source: Office for National Statistics

Table A4

ABI and MIDSS median regressions where the joint-hypothesis test can be rejected at the 5 per cent significance level

SIC2	2001	2002	2003	2004	2005
50					
51			*		
55	*	*	*	*	*
60	*	*			*
63	*				
64					*
71					
72					
73	*		*	*	*
74		*	*	*	
90		*		*	
92	*	*	*	*	*
93					

Source: Office for National Statistics