Review of early estimates of construction output for GDP in 2003

Phase 2 – proposals for change

For further details contact:

Frances Pottier at the Department of Trade and Industry, Tel: 020 7215 1953 or email: Frances.Pottier@dti.gsi.gov.uk

or

Rob Pike at the Office for National Statistics, Tel 01633 81 2624 or email rob.pike@ons.gsi.gov.uk

Review of early estimates of construction output for GDP in 2003

Summary, conclusions and actions

1 This is the final report of a joint DTI/ONS review, set up to explain the cause of the revisions to the estimates of construction sector output growth between Q1 and Q2 2003 and to make recommendations as to how such revisions might be prevented in future. The revisions to estimates of construction sector output for the second quarter of 2003 prompted significant revisions to UK GDP estimates for the same period.

2 An interim report (published on 12th February 2004 and available on the web at www.dti.gov.uk/construction/stats and www.statistics.gov.uk) of this review:

- discussed the methods used to compile early (M1) and final (M3) estimates of construction output including the use of modelling and data collected in sample surveys
- identified the likely cause of the revisions as the unforeseen fall in recorded construction output in Q1 2003.
- made the following recommendations as to what questions should be examined in the second and final phase.

1. Research should be undertaken to explain the unexpected fall in Q1 2003;
2. Work should continue on the development of the Activity Balances model and to improve the understanding of the relationship between early survey data and that available at M3;
3. Further work should be undertaken to identify other methods of producing accurate early estimates of construction output, such as a small monthly survey.

3 Most of this report is devoted to providing a response to these recommendations. The scope of inquiry was widened to review ways of modelling construction sector output and ways of producing early estimates based on a monthly survey.

Stage 2 Summary of conclusions and actions

Q1 2003 output fall (paragraphs 3.1 to 3.16)

4 The fall in construction output during in Q1 2003 compared to the previous quarter was largely due to under-recording of construction output as a result of a processing error in DTI. Although the existence of this under-

recording was picked up within a few days of selection by internal controls and remedial action taken, this action did not pick up the full set of firms affected by the error. It has now been established that this selection error reduced the estimate of quarterly growth in the construction sector in Q1 from +0.5% to –2.6% and in Q2 raised the growth estimate from 2.1% to 5.3%.

5 When the original processing error was initially identified by DTI and remedial action taken in May 2003, DTI believed that the error had been corrected in full. However, DTI only became aware in June 2004 that the original remedial action had underestimated the effect when it emerged as part of findings from this review. DTI informed ONS at this time. The full extent of the error was identified around the end of September 2004.

6 When the construction output estimate is corrected for this omission, the consequential impact on UK GDP growth will be +0.1 percentage points for Q1 2003 and –0.1 on growth into Q2 2003.

7 This revision will be included in the 'Output and Employment in the Construction Industry' Statistical Press Release to be published on 3 December 2004 by DTI and then in Quarterly National Account First Release to be published on 23 December 2004 by ONS

Action 1:

DTI publish corrected output figures for Q1 2003 as soon as practicable in conjunction with ONS so that the improved and corrected estimates are available to Construction output customers, including notably the National Accounts.

Action 2:

DTI will maintain and develop further their commitment to quality monitoring and improvement for construction output estimates.

Modelling of early estimate (paragraphs 4.1 to 4.5).

8 Thus far no model based early estimate of construction output has been identified that can predict with high levels of accuracy the likely final survey based estimate of construction output. A particular problem for any model is the large and unpredictable variation in quarterly change in construction output.

9 There is some evidence that using a combination of New Orders or Activity Balances models may yield improvements in early estimation

although no single model's performance can be considered satisfactory at present.

10 These findings are tentative because a sufficiently long series of data for there to be confidence in the predictive powers of the models is not available. Until a longer time series can be used to evaluate the relative strengths of single and combinatorial models, the Activity Balances model is the best performing of those currently used and should be given more emphasis in formulating the M1 early estimate.

Action 3:

The DTI early estimate of construction output (M1) should place more emphasis on the Activity Balances model than has been the case in the past.

Action 4

DTI should approach Experian, the company that operates the Activity Balances model to see what later data might be added or other improvements made as a result of offering a modest delay in ONS deadlines for the early estimates.

Action 5

DTI and ONS should review the Activity Balance and other models in the light of additional data considered sufficient to produce robust analysis.

Monthly survey (paragraphs 5.1 to 5.7)

11 Data collected through a monthly survey might improve the quality of the early estimate as a predictor of the quarterly survey results. What is less clear is the sample size necessary to produce a significant improvement over model-based estimates.

12 In addition it is always necessary to consider the balance between possible improvements in estimates using sample surveys on the one hand and their costs of collection and burden on businesses on the other. Recent government statements have emphasized a commitment to lowering this burden.

13 DTI expects to continue to need ONS technical assistance to investigate ways of providing monthly estimates of construction output, either model- or survey-based. This will help the DTI meet an expected Eurostat need for monthly estimates.

Action 6

ONS will offer DTI continuing technical assistance to investigate ways of providing monthly estimates of construction output, either model- or survey-based.

Review of early estimates of construction output for GDP in 2003

1 Introduction

1.1 The Department of Trade and Industry (DTI) is responsible for the main official information on the output of the construction industry in Great Britain, which is both published quarterly by DTI and is supplied to the Office for National Statistics (ONS) to inform their releases on quarterly National Accounts. DTI publish a National Statistics release on quarterly construction output in the first week of the third month following the end of the quarter, and these figures are supplied to ONS for their *Quarterly National Accounts First Release*, published towards the end of the third month (Month 3 or M3).

1.2 DTI also supply ONS with estimates of construction output for their estimates of GDP published in the first and second months following the end of the quarter, known as the Month 1 (M1) and Month 2 (M2) estimates. In the first two quarters of 2003, these estimates did not accurately predict the figures published by DTI at M3, leading in the second quarter to a revision to GDP growth from 0.3% to 0.6%. Following the publication of the *Quarterly National Accounts First Release* for Q2 2003, ONS and DTI announced that they would be carrying out a joint review in two phases; firstly to identify the causes of the revisions; and then to identify options for improving the early estimation of construction output. This paper completes the second phase of the review.

2 The Statistics Commission Review of Revisions

2.1 The Statistics Commission carried out a review of revisions to economic statistics, including (but not limited to) the revisions to estimates of construction output and its contribution to GDP in the first half of 2003. Their report, published on 22 April 2004 can be found on the Commission's website, www.statscom.org.uk. It includes the following recommendation directly related to construction estimates:

> RECOMMENDATION (Paragraphs 22, 48):
> The Office for National Statistics (ONS) should assess more systematically the performance of the forecasting models used in compilation of the first estimate of GDP, undertake further methodological development, and make other changes aimed at greater transparency and best practice.

2.2 Both DTI and ONS supplied evidence to the review. In the response to this recommendation, the National Statistician announced that it would be considered as part of this joint DTI/ONS review of early estimates of construction.

Summary of Conclusions from Phase 1

2.3 The main conclusion of Phase 1 of the review was that it was the unforeseen fall in construction output in Q1 2003 that largely contributed to the revision in the estimate of both Q1 and Q2 construction output. This was both because of the effect it had on the estimation models, and because it caused DTI to be more cautious in forecasting a return to positive growth in Q2 2003 than might otherwise have been the case. This downward revision in the estimate of Q1 construction between M2 and M3 was not apparent in the Q1 GDP figures because of an upward revision in the services sector.

2.4 The following recommendations were made:

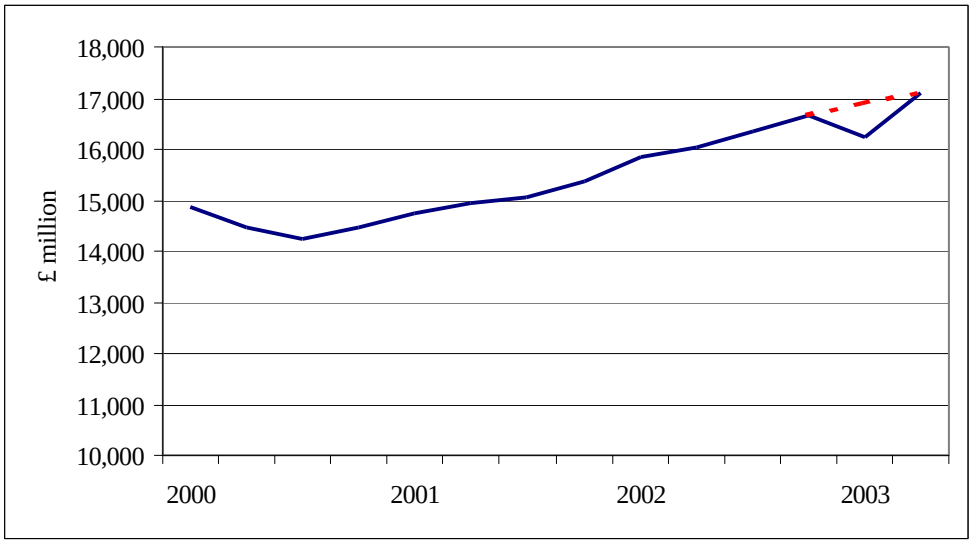
1. Work should continue on the development of the Activity Balances model and to improve the understanding of the relationship between early survey data and that available at M3;
2. Research should be undertaken to explain the unexpected fall in Q1 2003;
3. Further work should be undertaken to identify other methods of producing accurate early estimates of construction output, such as a small monthly survey.

This Phase 2 report takes forward each of these recommendations.

3 Investigation into fall in construction output in Q1 2003

3.1 Phase 1 of the review identified the unexpected fall in construction in Q1 2003 (M3 estimate) as largely contributing to the revision in the estimate of both Q1 and Q2 construction output. It recommended that Phase 2 should include an investigation into whether any problems had occurred with the surveys to cause the fall. The result of the investigation was to identify a problem at the survey of firms selection stage which caused construction output in Q1 2003 to be under recorded by £593m in current prices, or £518m in constant 1995 prices. The following graph shows the difference between published output (solid line) and revised output with the erroneous under-recording removed (dotted line) in Q1 2003.

Chart 1: Construction Output (Constant (1995) prices, seasonally adjusted): for the period 2000 –Q2 2003



What happened

3.2 It turns out that the estimated fall in construction output during Q1 2003 was due to a fault in the computer processing of construction firm data at that time.

3.3 An apparently very modest change to a computer programme produced a large, and at the time unforeseen, impact on the computer record of construction firms. The modest change was to allow, in the programme, for the theoretical possibility that a construction firm might employ 100,000 plus employees. The unforeseen consequence of this change was that the computer reset the allocated stratum for each firm. The record

it used was the return for the third quarter of 2002. Where a firm had no record of any employment, in that quarter, the computer programme treated that firm as ineligible for selection. In technical terms any companies which had not returned a form in Q3 had had their size group set to null (i.e., the field was empty).

3.4 There is a system in place, which should have stopped such a small change bringing about such a large impact. This is the issuing of a change control notice to the external IT systems managers of what is proposed so that they can assess any possible adverse consequences. In this case it was a DTI official who informed the external contractors in charge of maintaining the computer system that they wished to alter the field size in a particular firm size band so that it could accommodate an employment record which extended to six digits rather than the then five. Systems support agreed to this change.

3.5 This marked reduction in a number of firms on the register regarded as live came to light 2 days after 12,000 firms were sampled as the basis for the Q1 2003 survey. The execution of a standard register proving procedure indicated that about 66,500 firms on a register of around 170,000 had, in effect, been un-stratified. Where firms have not been allocated to a firm size band they are missed out from the sampling population. This was highly implausible. An investigation was carried out and action was taken to remedy the computer error. By selection of the Q2 2003 survey the register was accurate.

3.6 **Remedial action in the form of a systematic checking of virtually all (99.5%) of the missing firms gave rise to what was thought, then, to be the total number of firms which had been effectively removed from the register of active businesses.** At the time when the Q1 2003 final estimate of output was being compiled a correction was made for 66,500 missing firms by adjusting the grossing fractions.

3.7 The contrast in the different sizes of live firms over the three quarters Q4 2003 to Q2 2003 is shown in table 1. What is also evident from the table is the predominance, in the construction industry, of small and medium size businesses.

3.8 There the matter rested until this review, when a much more detailed and complex check of individual firm returns was carried out which identified the remaining 0.5% of firms whose returns were not checked at the time of the computer error first being noted.

Table 1: Sample and universe sizes by Stratum, Q4 2002, Q1 2003, and Q2 2003

Stratum	Q4 2002		Q1 2003		Q2 2003	
Size Group :	Sample	Universe	Sample	Universe	Sample	Universe
SG0	1,392	73,364	1,125	30,314	951	70,977
SG1	1,430	50,306	1,600	31,831	633	52,408
SG2	1,311	23,963	1,591	19,116	1,391	25,055
SG3	2,098	9,819	1,952	7,833	2,530	10,120
SG4	1,355	5,427	1,638	4,920	1,914	5,742
SG5	919	1,809	859	1,719	954	1,909
SG6	1,738	1,738	1,635	1,635	1,821	1,821
SG7	527	527	488	488	529	529
SG8	446	446	407	407	461	461
SG9	532	532	469	469	546	546
SG10	126	126	121	121	135	135
SG11	69	69	64	64	72	72
SG12 & 13	59	59	53	53	63	63
Total	12,002	168,185	12,002	98,970	12,000	169,838

3.9 On further examination it now appears that about 355 firms had been treated, in Q1 2003 as if they were ineligible for selection. All of these firms were large. All were in size bands 6-13 and because all firms in these bands are sampled there was no grossing correction. Thus although they count for a minute fraction of the total number of construction firms, their contribution to construction output is disproportionately large. Table 2 shows the amendments for each stratum once these firms are included.

Table 2: Amendments by stratum

Size group	Number of companies	Employment	Output (£ million)
6	151	6,342	105
7	52	3,588	61
8	60	5,700	117
9	67	11,926	233
10	12	5,268	93
11	7	5,873	122
12	6	14,802	238

3.10 In the course of this review the size of this underrecording of construction output has been estimated and a new, and corrected figure for Q1 2003 construction output compiled. The effects of these amendments are shown in Table 3

Table 3: Effects of the amendments to Q1 2003 construction output and GDP growth, in constant prices, seasonally adjusted

	Total Output New work	Total output R&M	Total Output All Work	M1 Change estimate	Resulting Revisions M1 – M3
Time Series of output estimates with Q103 as published (in 1995 prices)					
Q402 (£m)	9,149.1	7,522.2	16,671.3		
Q103 (£m)	8,909.8	7,326.6	16,236.4		
Q203 (£m)	9,207.9	7,890.8	17,098.7		
Resulting Quarterly Growths :					
Q402 - Q103	-2.6%	-2.6%	-2.6%	2.3%	-4.9%
Q103 - Q203	3.3%	7.7%	5.3%	-0.4%	5.7%
Time series of output estimates with Q103 corrected for selection error (in 1995 prices)					
Q402 (£m)	9,149.1	7,522.2	16,671.3		
Q103 (£m)	9,370.0	7,384.6	16,754.5		
Q203 (£m)	9,207.9	7,890.8	17,098.7		
Resulting Quarterly Growths :					
Q402 - Q103	2.4%	-1.8%	0.5%	2.3%	-1.8%
Q103 - Q203	-1.7%	6.9%	2.1%	-0.4%	2.5%
Revisions to Q103 (£m)	460.2	58.0	518.2		

3.11 The failure to include these firms in the universe at Q1 2003 meant that output in that quarter was underestimated by just over £500m. Had they been included, the contribution from construction to any revisions to GDP for Q1 and Q2 between M1 and M3 would have been much smaller.

3.12 It is possible that having correct figures for Q1 would have improved the model-based M1 estimate for output in Q2 2003. Putting the above amended figures through the Orders model produces an estimate of –0.5% change, which is not much different to the original M1 estimate of -0.4%. However, the Activity Balances model was run at the time excluding the Q1 downturn, producing an estimate of growth in Q2 of 4.4%, which suggests that had the Q1 figures been correct, this - Activity Balances - model would have performed better in Q2.

3.13 The computer systems have been made more secure. The ability to change field lengths has been prevented without first going through a thorough evaluation procedure. DTI and ONS are satisfied that the error has been rectified and suitable steps taken to prevent its recurrence.

Impact on GDP

3.14 The fall in Q1 2003 for the estimate of Gross Value Added (GVA) for the UK construction industries, as published by ONS in June 2003, was not as large as the fall in GB construction output that had been published by DTI. Consequently the impact on GDP of correcting the estimate of GB construction output in Q1 2003 is smaller than that implied by the change to the DTI's estimate of GB construction growth. The net impact on UK GDP of correcting the construction output data would be to revise the estimate of quarterly growth in real GDP up by 0.1% in Q1 2003 and down by 0.1% in Q2 2003, if there were no other revisions to components of GDP.

3.15 The reason for the difference between the DTI's GB construction output figure and the figure for GVA of UK construction industries is the way DTI construction data are used in the UK National Accounts. The DTI's GB figures are benchmarked to annual estimates of GVA and additional data is incorporated for Northern Ireland. Additionally, the estimates are taken through the National Accounts balancing process. This process is used to reconcile the production, expenditure and income measures of GDP. In the case of the estimate of real GDP growth for Q1 2003 that ONS produced in June 2003, this balancing process led ONS to apply a 'balancing adjustment' to the estimate of GVA of UK construction industries. The decision to apply this adjustment took account of related components within the expenditure measure, in particular the estimate of investment in construction, and the perceived reliability of the DTI's series at that time. This adjustment raised the level of the series in Q1 2003 and so had the effect of increasing growth

in Q1 and decreasing growth in Q2. This adjustment will be removed when the error in the survey data is corrected.

Actions

Action 1

- DTI publish corrected output figures for Q1 2003 as soon as practicable in conjunction with ONS so that the improved and corrected estimates are available to Construction output customers, including notably the National Accounts.

Action 2

- DTI will maintain and develop further their commitment to quality monitoring and improvement for construction output estimates.

4 Evaluation of forecasting models used for early estimates

4.1 DTI currently uses two forecasting models to provide early estimates of quarterly change in construction output to ONS. The more well established of these is based on New Orders data, and utilises known lags between the placement of an order and output at various stages of completion to estimate how recently placed orders should be materialising as construction output. Recent doubts about the continued accuracy of this model led DTI to develop, in conjunction with Experian Business Strategies, an alternative model, based on a survey of construction firms carried out by Experian assessing the likelihood of construction output increasing or decreasing in one quarter compared to the last.

4.2 Both of these models were assessed in some detail by NIESR as part of their investigation for the Statistics Commission, and their performance was compared with other modelling methodologies. NIESR considered that the ARIMA class of models offered a promising alternative, and in particular that a composite model combining ARIMA, the New Orders model and the Activity Balances model offered some slight improvement over using the Activity Balances model alone.

4.3 At the time of their investigation, the cause of the apparent fall in output for 2003 Q1 and its subsequent recovery had not been fully established, and only the original correction had been made for the undersampling in the data made available to NIESR. Once the full extent of the undersampling had been established and further corrections made, ONS methodologists specialising in time series analysis were asked to re-evaluate NIESR's approach using fully corrected data, and establish whether their conclusions were still valid. The report of the ONS methodologists is reproduced as Annex B.

4.4 In summary, it found that the ARIMA model proposed for inclusion in a composite model was unstable over time, and ONS proposed that a second best performing ARIMA model should be used instead. They compared a composite model using a simple ARIMA (0,1,1) with the New Orders and Activity Balances models. None of the three models was found to be an adequate predictor of construction output growth. Although the composite model appeared to perform better than the Activity Balances model alone when using Root Mean Square Error as the quality defining parameter, the composite model's predictions were further from actual output growth than the New Orders or Activity Balances models.

Conclusions.

4.5 The ONS methodological research has produced some useful results. These include:

- that no model performs clearly better than the others. Using either of RMS or prediction accuracy as assessment criteria, different models seem to perform better but performance differences are always small;
- no single model's performance can be considered satisfactory at present;
- neither has an optimal combination of models yet been clearly identified;
- there is some evidence that using a combination of New Orders or Activity Balances models may yield improvements in early estimation;
- the lack of an adequately long time series of data to use with the models may be a contributory factor in limiting their performance.

Actions

Action 3:

- The early estimate of construction output (M1) should place more emphasis on the Activity Balances model than has been the case in the past.

Action 4

- DTI should approach Experian, the company that operates the activity balances model to see what later data might be added or other improvements made as a result of offering a modest delay in ONS deadlines for the early estimates.

Action 5

- DTI and ONS should review the activity balance and other models in the light of additional data considered sufficient to produce robust analysis.

5 Monthly Survey

5.1 One of the recommendations arising from Phase 1 was that the review should consider alternative methods of deriving early estimates, including a monthly survey. Interest in a monthly survey arises not only from the desire to find a feasible alternative to a model based early estimate but also as one possible method to fulfil future obligations to Eurostat to provide monthly estimates of construction and GDP growth. ONS methodologists evaluated a number of alternative scenarios for a monthly survey.

5.2 The ONS report is reproduced at Annex C. To this, DTI has added estimates of cost (both to the taxpayer and to business). This section summarises the results.

5.3 Four alternative methodologies were considered.

1 Full monthly survey. This would leave the current survey in its present form but increase the periodicity. Currently around 12,000 firms are surveyed each quarter; the proposal would be to increase the frequency of the survey leaving the sample design the same, so that 12,000 firms would be surveyed each month.

2 Cohort sample. This option would effectively split the quarterly sample into three parts (or cohorts), each of which would be asked for monthly data. Quarterly data would be derived by the aggregation of three months' data. The partition would not be a straightforward division of the sample into three equal parts of 4,000 firms each, since those firms currently sampled on a 1-in-1 basis would continue to be sampled on the same basis each month.

3 Monthly subsample. A subsample of the quarterly survey would be selected for a monthly survey.

4 Model- based estimates. This was considered as the principal alternative to a monthly survey.

5.4 The cost implications of each of these is summarised in Table 7. The table compares each of the four options with the current quarterly survey, on both costs to business and costs to the taxpayer. Compliance costs for business are based on current assessments of the time necessary to complete survey forms and are used to complete evaluation of the cost to business for National Statistics accounting purposes. Costs to the taxpayer include staff costs and a proportion of other resources such as computing which are shared with other surveys conducted by DTI.

Table 6: Cost implications of monthly sampling

	Compliance cost	Cost to DTI
Current	2,111,000	328,000
Full monthly	6,332,000	985,000
Cohort	2,661,000	424,000
Monthly subsample - large firms	2,205,000	345,000
Monthly subsample - all firms	2,317,000	361,000

Based on 2004/05 survey control costings

5.5 No assessment has been made for the non-survey option of model-based estimates. It is assumed that this method would be relatively inexpensive, involving some initial set-up costs but negligible running costs (similar to the current model-based early estimates) and no additional compliance burden on businesses.

5.6 The evaluation of the relative strengths of the other options considered, was based on an assessment of their likely accuracy (of early quarterly estimates, final quarterly estimates and monthly estimates) and cost.

5.7 Any assessment of these options takes place against a background of a both very volatile and unpredictable path of quarter on quarter change in construction output. As a way of exploring the significance of this volatility DTI examined the returns from the larger firms in the quarterly survey, to see if they can be used with some form of modelling to provide a biased early estimator of quarterly output. Unfortunately, the results do not show a consistent bias. An estimate based on a subsample varies randomly from the full survey result. The general point and a particular illustration both point to the problem of not knowing how much confidence can be placed on a monthly survey results, no matter how large a sample, as the basis of forecasting a whole quarters outturn. The substantive question that the review was unable to answer was how much the accuracy of an early estimate of quarterly output would increase for any given increase in monthly sample size. For this and other reasons, it was not possible to say whether the extra cost of collecting monthly data could be justified.

Conclusions

5.8

- None of the monthly survey options (other than a full monthly survey) would retain the current high accuracy of the final quarterly estimate.
- A full monthly survey is the option likely to be a good predictor of the final quarterly estimate. This is the highest cost option and this cost would be difficult to justify in statistical and economic efficiency terms.
- The other monthly survey options cost less but suffer from concerns over likely usefulness for improved predictions of final quarterly estimate.
- Finally, there is no way of assessing how far the incremental benefit of a large monthly sample over a smaller one may be worth the extra cost.
- The question of how to generate monthly estimates to meet Eurostat requirements was outside the scope of this review. It is noted that DTI and ONS are expected to continue to seek ways of providing monthly estimates of construction output, either model- or survey-based.

Actions

Action 6

- ONS will offer DTI continuing technical assistance to investigate ways of providing monthly estimates of construction output, either model- or survey-based.

Annex A - Construction Output Sampling, Imputation And Grossing Methodology

A1 Production of construction output

A1.1 Construction output is a quarterly series of the output of the construction industry, in both the private and public sectors. Press Notices are published on the first Friday of March, June, September and December. The actual series is a combination of results from 2 surveys, the Quarterly Inquiry of Construction Activity and the Building and Civil Engineering Employment and Output Enquiry, plus an estimate for unrecorded output.

The Quarterly Inquiry of Construction Activity

A1.2 The Quarterly Inquiry of Construction Activity (CA) measures the output (and employment) of private contractors, covering new construction, improvements and alterations and repair and maintenance.

A1.3 Output figures include:

the value of building civil engineering and associated work, excluding VAT done by the contractor's directly employed staff which is chargeable to customers
The value of materials used, labour costs, overheads and profits
The value of work done on the contractors own initiative on buildings such as dwellings for eventual sale or lease
The value of work done on demolition and site preparation
The value of work done by the contractor on the construction or maintenance of its own premises
The value of articles made by the contractor and used in construction work
The value of any materials supplied by the contractor free of charge to subcontractors.

A1.4 The following are excluded:

The value of work done overseas
The value of work done by sub-contractors of any type
The value of any payments made to labour-only subcontractors
The value of articles made for sale or materials sold
The value of materials supplied to the contractors free of charge by firms in the industry
The value of land
Any architects or consultants fees.

Sampling and collection of Quarterly Construction Activity

A1.5 Information for Great Britain is based on quarterly returns from a sample of around 12,000 firms taken from the CISTATS universe of around 170,000 construction contractors. This universe is kept up-to-date through quarterly exchanges of information with the Inter-Departmental Business Register as well as information from data suppliers.

A1.6 Construction firms on CISTATS are assigned to size groups (SGs) based on their employment levels. Initially the employment data supplied by ONS at the company's "birth" is used, this is then updated when the company responds to a DTI survey. The sample is selected using a rotational sampling methodology with an 8 quarter rotation- rotational sampling is explained in Section 2 of this Annex. The table below shows the sampling ratios used – SGs 0 and 1 have a variable sampling ratio to ensure that the sample size is always 12,000:

Size Group	Employment	Sampling Ratio
0	1	Variable
1	2-3	Variable
2	4-7	1 in 8
3	8-13	1 in 4
4	14-24	1 in 3
5	25-34	1 in 2
6	35-59	1 in 1
7	60-79	1 in 1
8	80-144	1 in 1
9	115-299	1 in 1
10	300-599	1 in 1
11	600-1199	1 in 1
12	1200+	1 in 1

A1.7 The output information collected is broken down into New Work, Public Housing Repair and Maintenance, Private Housing Repair and Maintenance, Public Non-housing Repair and Maintenance and Private Non-housing Repair and Maintenance. The employment information is broken down into Working Proprietors, APTCs (administrative, professional, technical and clerical employees) and Operatives (manual workers).

Results Processing - Imputation and Grossing

A1.8 The following paragraphs describe the processing carried out to produce construction output from the information collected by DTI surveys. DTI uses the agreed common methodology used for all business surveys.

A1.9 Pre-imputation and Imputation – CA firms which are in the sample but have not responded have a value imputed for them. For each Size Group, the system calculates an imputation link by taking the average movement between the previous quarter and the current quarter for those firms which have responded. This is then applied to those firms which responded in the previous quarter but not in the current one. Where a firm did not respond in the last quarter an average figure is imputed, which is the same for each affected firm in the Size Group. During pre-imputation, the statistician examines the imputation link for each question in each Size Group, and adjusts it if necessary.

A1.10 Construct – a small number of the largest firms have been marked as key responders, and cannot be imputed for. For each of these firms that has not responded, the statistician must construct a response, usually based on the imputed value for the strata, unless there is better information available.

A1.11 Weights – The first stage of grossing is the calculation of the grossing weights, which will be used to multiply up the sample total to the universe total. CA uses 3 different grossing weights: model, design and outlier.

A1.12 The model weight (g) adjusts the grossing according to known independent properties of the companies sampled. The independent variable used is the latest employment value held on the IDBR and collected through an ONS survey. This information is updated on CISTATS annually, prior to selection of the Q1 sample. The formula is:

$$g = \frac{X}{\left(\frac{N}{n}\right) * x}$$

where

X = Size Group population total for ONS employment
x = Size Group sample total for ONS employment
N = population size at selection
n = sample size

A1.13 For companies in Size Groups 6 and above, where there is 1 in 1 sampling, the model weight is 1 because X = x and N = n

A1.14 The design weight (a) is the main grossing weight. The formula is:

$$\left(\frac{N}{n}\right) * \left(1 + \frac{hd}{(n-d)}\right)$$

where

h = factor for births and deaths (1 for companies in strata 0-5, 0 for companies in strata 6-13)
d = number of deaths (closedowns) recorded
N = population size at selection
n = sample size

Again for Size Groups where there is 1 in 1 sampling (i.e. SG 6-13) the design weight is equal to one.

A1.16 The outlier weight (w) ensures that responses which are deemed to be outliers (see below) are not grossed as they are not representative.

A1.17 If the output per person is greater than the outlier limit set then

$$w = \frac{n}{N}$$

where

N = population size at selection
n = sample size

otherwise w = 1

A1.18 Outlier limits, that is the cut-off point whereby everything above that level is deemed to be an outlier, are set manually for each strata every quarter, based on the output per person values for each return. For example, if an outlier limit of £60,000 is set, then all returns where the output per person is greater than this are deemed to be outliers, and will be grossed accordingly.

A1.19 So, grossed value = a*g*w*ungrossed value

where a, g and w are the grossing weights described above. Because non-responders are imputed for, returns for firms in Size Groups 6 and above, where there is 1 in 1 sampling, are not grossed.

The Building and Civil Engineering Employment and Output Enquiry (DLO)

A1.20 The Building and Civil Engineering Employment and Output Enquiry is usually known as DLO, for Direct Labour Organisation. DLOs are public sector organisations which employ their own construction workers rather than contracting out construction work to the private sector. In the past this

included the public utilities such as the water boards however now these have largely been privatised it covers local authorities.

Sampling and collection of DLO information

A1.21 As there are only about 215 DLOs, this quarterly enquiry is a census, using 1 in 1 sampling. The questions asked are the same as for the CA Inquiry.

Results Processing - Imputation and Grossing

A1.22 DLO results are processed in a similar way to CA, going through Pre-imputation and Imputation. No grossing is necessary because of the 1 in 1 sampling.

Estimate of Unrecorded Output

A1.23 The Builders Address File, which DTI uses as a universe for all its contractor based surveys is aligned quarterly with the IDBR but only includes those firms over the VAT threshold. It is clear from comparing the ONS surveys of employment with the figures for construction employment collected in the CA and DLO surveys that a significant proportion of construction employment is either self-employment or with firms not registered for VAT. The work done by these people is therefore not captured by either the CA or DLO surveys.

A1.24 In order to compensate for this missing "unrecorded" output, we make an estimate of it. This is done by taking the difference between total construction employment (including the self-employed) as measured in the Labour Force Survey, and total construction employment reported to CISTATS. This gives us a measure of unrecorded employment, which tends to be around 600,000. This is then combined with wage information and the average output per person for Size Groups 0-2 to give an estimate of the output produced by the construction workers missed.

2 Rotational Sampling

A2.1 Construction Activity sampling is carried out using a rotational sampling method. Each firm is allocated to a Size Group based on employment. The sample required from each cell is determined by the pre-set sampling fraction appropriate to the Size Group.

A2.2 Each firm is allocated a permanent random number, or PRN, when it is added to the Inter-Departmental Business Register (IDBR). This is used to produce a list of firms in each cell in random order. The first sample is the first n firms in the list, where n is the sample size.

A2.3 To work out the second sample, we need a further value, the "period to rotate" or "r". This indicates the number of periods for which we want a firm to stay within the sample, and is pre-set. For Construction Activity the period to rotate is 8, which is 8 quarters or 2 years. We then work out n/r, which gives us the number of firms which are dropped from the start of sample 1. Starting from the first firm left in the list, we take the next n firms and this makes up sample 2.

A2.4 New firms ("births") are allocated a PRN and slotted in where appropriate – the only proviso is that a firm which has started its turn in the sample cannot fall out because of a birth, and in these cases the inclusion of such a firm will be forced. Firms which have closed down ("deaths") leave the list, and the sample moves on as if they had never existed.

A2.5 This concept is best illustrated using an example. Imagine a survey with a population (N) of 100 firms, with a sampling fraction of 1 in 10 (so a sample size (n) of 10 firms) and a period to rotate (r) set to 5. We list all 100 firms in random order, and then start to sample. The first sample is the first 10 firms on the list. For the second sample, we drop the first 2 firms (sample/period to rotate = 10/5 =2), and then, starting from firm 3, we take the next 10 firms. For the third sample, we again drop 2 firms, and take 10 firms starting from firm 5, and so on. After 5 samples, all the firms in the first sample have completed their turn, and the sixth sample contains none of the firms which were included in the first sample. These firms then have a survey holiday, which lasts in this example for approximately 45 periods – the exact length of a holiday is affected by births and deaths but the formula is:

$$\text{Holiday length} = \left(\frac{(N-n)}{n}\right)r$$

Table A1: Example of rotational sampling

Sample 1	Sample 2	Sample3	Sample 4	Sample 5	Sample 6	No of periods in sample
Firm 1						1
Firm 2						1
Firm 3	Firm 3					2
Firm 4	Firm 4					2
Firm 5	Firm 5	Firm 5				3
Firm 6	Firm 6	Firm 6				3
Firm 7	Firm 7	Firm 7	Firm 7			4
Firm 8	Firm 8	Firm 8	Firm 8			4
Firm 9	Firm 9	Firm 9	Firm 9	Firm 9		5
Firm 10	Firm 10	Firm 10	Firm 10	Firm 10		5
	Firm 11	Firm 11	Firm 11	Firm 11	Firm 11	5
	Firm 12	Firm 12	Firm 12	Firm 12	Firm 12	5
		Firm 13	Firm 13	Firm 13	Firm 13	5
		Firm 14	Firm 14	Firm 14	Firm 14	5
			Firm 15	Firm 15	Firm 15	5
			Firm 16	Firm 16	Firm 16	5
				Firm 17	Firm 17	5
				Firm 18	Firm 18	5
					Firm 19	5
					Firm 20	5

Annex B –Construction Estimates

Preface:

B1 This annex to the Report is the content of an analysis into the performance of various different estimation models, commissioned by the project team and carried out by the Methodological Group at the ONS. The report is complete and has been included as drafted by the authors.

1 Introduction

B1.1 This document provides details of DTI construction growth estimate analysis. The main contacts for this analysis were James Mitchell (NIESR), Keith Folwell and Frances Pottier (DTI), Nigel Stuttard and Claudia Annoni (ONS).

B1.2 The main objectives of the analysis defined in June were categorised under two broad headings:

- Peer review the analysis conducted by NIESR;
- Provide recommendations on which model better provides early construction estimates and comment on the performance of these estimates.

B1.3 However, in August DTI identified an error in their original Q1 2003 estimate due to a system error. It was therefore not sensible to make a direct assessment of NIESR suggestions because the data that NIESR used in the analysis was erroneous.

B1.4 As a result of the error identification, it was agreed to re-run the analysis on the revised data and to focus on giving recommendations on which model provides a better performance to estimate the construction growth rate.

2 The data

B2.1 Data used in the analysis were provided by DTI and NIESR from 1999 Q1 to 2004 Q2. The dataset includes:

- M1: The preliminary estimate of construction output growth completely based on modelling new orders;
- M2: Second estimate of construction output growth based on the results of the new orders model and 50% of the survey data;
- M3: The final estimate of the construction growth data completely based on survey data;
- Activity: Estimates based on a model developed by Experian for DTI. This model is currently used by DTI only for comparison and validation purposes.

B2.2 The table below shows the variables of interest used in the analysis after the correction for the erroneous estimate in 2003 Q3.

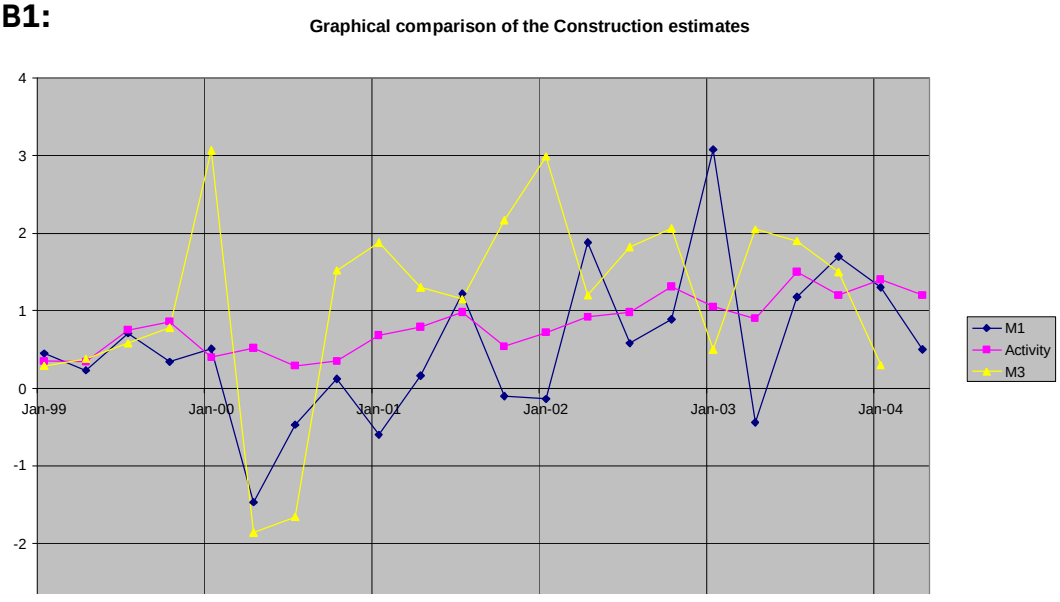
Table B1:

Date	*M3*	*M1*	*Activity (1)*	*M2*
1999Q1	0.29	0.45	0.35	.
1999Q2	0.38	0.23	0.34	0.6
1999Q3	0.58	0.71	0.75	0.5
1999Q4	0.78	0.34	0.86	0.6
2000Q1	3.07	0.51	0.40	0.5
2000Q2	-1.86	-1.47	0.52	-0.8
2000Q3	-1.66	-0.47	0.29	-0.5
2000Q4	1.52	0.12	0.35	0.3
2001Q1	1.88	-0.60	0.68	-0.5
2001Q2	1.30	0.16	0.79	0.2
2001Q3	1.15	1.22	0.98	1.2
2001Q4	2.17	-0.10	0.54	-0.8
2002Q1	2.99	-0.14	0.72	-0.1
2002Q2	1.20	1.88	0.92	1.5
2002Q3	1.82	0.58	0.98	0.6
2002Q4	2.06	0.89	1.31	0.9
2003Q1	0.50	3.08	1.05	3.1
2003Q2	2.05	-0.44	0.90	0.9
2003Q3	1.90	1.18	1.50	1.8
2003Q4	1.50	1.70	1.20	1.7
2004Q1	0.30	1.30	1.40	1.0
2004Q2	0.74	0.50	1.20	1.0

(1) The activity data is composed by NIESR estimates from 1999Q1 to 2003Q3 and DTI estimates for the remaining 3 observations.

B2.3 The graph below compares the three estimates of the construction output growth.

Chart B1:



B2.4 From the graph it is possible to see that the construction output is a very volatile series, especially in 2000 where abrupt spikes occurred. This behaviour together with the short length of the series will make it difficult, for any sort of modelling, to predict the final output growth. It is also possible to notice that the orders model is not a very good predictor of the construction output growth, whilst the activity model provides reasonable estimates of the level of the growth series, but not of the short-term movements. This is also confirmed by the low correlation of just 0.17 between M1 and M3 and by the slightly better correlation of 0.25 between Activity and M3.

B2.5 The "Review of revisions in economic statistics" looked at possible ways to improve the estimates, and in particular looked at ARIMA modelling and a weighted combination of orders, activity balances and ARIMA estimates.

B2.6 The following section shows the effect of running these two approaches on the corrected data and provides comments and recommendations on quality of these estimates.

B3 The analysis

B3.1 The main purpose of the following analysis is to run the methodology of early estimates of construction output for GDP proposed by NIESR on the revised data and to make comments and recommendations on the proposed approach.

B3.2 In the "Review of revisions in economic statistics" NIESR looked at two possibilities to improve the estimates:

- ARIMA modelling; and
- A weighted combination of the orders, activity balances and ARIMA estimates.

B3.3 We have run the analysis suggested by NIESR with data from 1999 Q1 to 2003 Q3, and used the remaining observations to evaluate the stability of the models and comment on the performance of the estimates.

B3.4 To evaluate the stability of the models consecutive runs have been undertaken considering different spans of data: first 1999 Q1-2003 Q3, then 1999 Q1 – 2003 Q4 and finally 1999 Q1 – 2004 Q1. The parameter estimates should not change dramatically if the model is stable.

B3.5 The performance of the estimates will be evaluated using the Root Mean Square Error (RMSE). This measure will not provide information on the reliability of the estimates since it will not take into account of the standard error of the estimate. However a better scoring rule cannot be used because

the new order estimates and the activity balances estimates do not have associated standard errors.

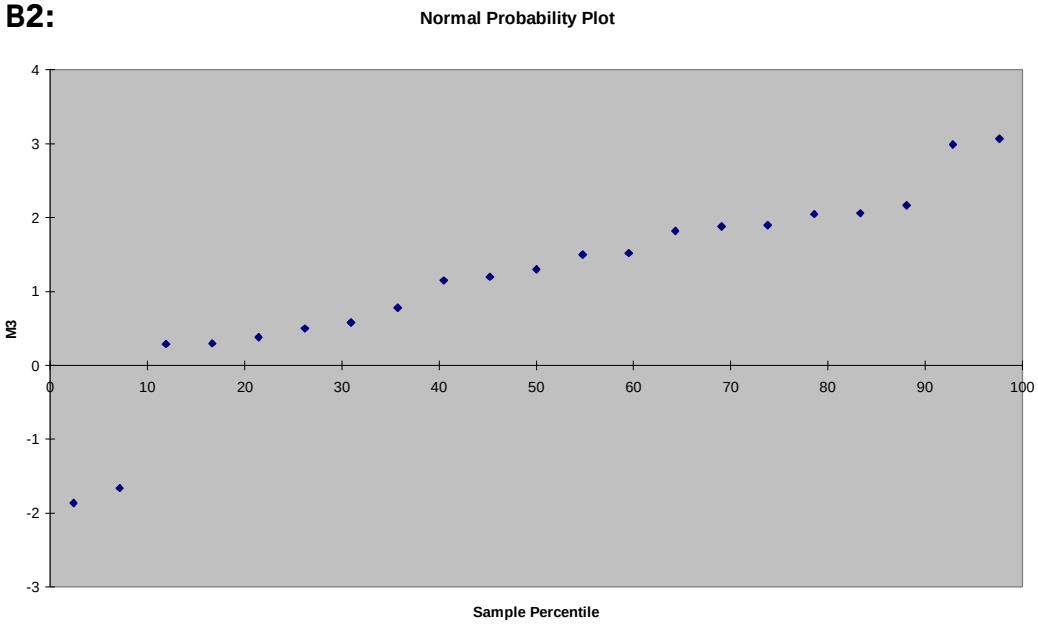
B3.6 Before undertaking any analysis, a check of the normality assumption should be made to establish the reliability of the model estimates.

Normality probability plot:

B3.7 The check to verify the normality distribution of the series has been performed with a Normal probability plot.

B3.8 From Chart B2 it is possible to see that the construction growth is fat tailed (indicating that the distribution is not normal). The effect of a fat tail distribution is usually minimised by transforming the series using logarithms, however, this is not possible with growth rates since they can take negatives values or be equal to zero. This suggests that modelling will not provide good estimates for the extremes, underestimating the two tails.

Chart B2:



B3.9 This will not strongly affect the parameter estimates of the model, and it will be possible to use the parameter estimates for hypothesis testing.

ARIMA Modelling

B3.10 NIESR compared three different ARIMA models: (0,1,1), (0,2,2) and (BIC,1,BIC). These models were applied to the following spans of data: 1999 Q1-2003 Q3, 1999 Q1 – 2003 Q4, 1999 Q1 – 2004 Q1 and 1999 Q1 – 2004 Q2, using SAS.

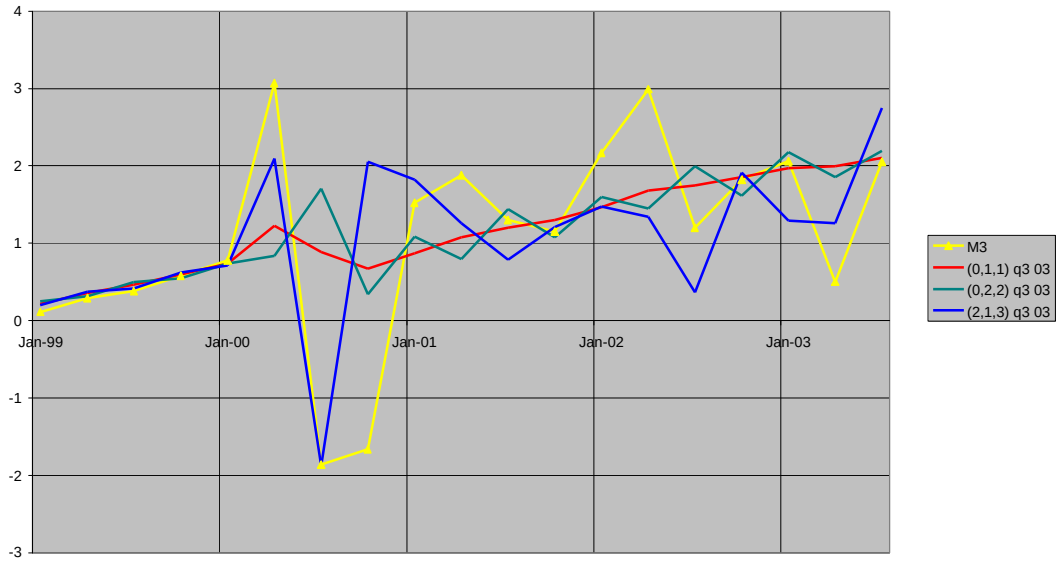
B3.11 In order to achieve best results using the ARIMA approach, two conditions need to be met. The first concerns the number of observations to be used in modelling. The generally accepted threshold of 30 data points is the minimum required. For well behaved series with a regular pattern and small volatility it is still possible to use ARIMA modelling to get a sound forecast with a shorter series. However, this is not the case for the construction series. In this case the series is not well behaved and irregular

movements dominate its short-term pattern. Working with too few data points can result in poor model identification. Since the construction series has less than the minimum number of data points generally accepted the model selection was conducted using the whole series.

B3.12 The second condition is that the data series is stationary, i.e. that the data series varies around a constant mean and variance. If the data is non-stationary, the data series needs differencing and/or the use of the natural logarithm to reduce the difference in the variability of the data. Whether the data requires a difference becomes apparent at the identification stage with software programs (i.e. SAS), whilst the use of log-transformation cannot be in this case applied because the nature of the data. Running a regression model on non-stationary variables can make results spurious.

Chart B3 shows the results of ARIMA modelling with data up to Q3 2003:

Chart B3: Graphical comparison of the Construction ARIMA estimates with data up to Q3 2003



B3.13 The table below shows the root mean square error (RMSE) calculated for each model run using three spans of data (the RMSE includes one forecast ahead of the span used):

Table B2:

	RMSE (Q4 2003)	RMSE (Q1 2004)	RMSE (Q2 2004)
M1	1.5737	1.5512	1.5164
Activity	1.2403	1.2340	1.2096
ARIMA (0,1,1)	1.2118	1.2477	1.2427
ARIMA (0,2,2)	1.2893	1.3320	1.3185
ARIMA (BIC,1,BIC)	1.2066	1.2541	1.2010

B3.14 The results of the modelling estimation on the four spans can be summarised as follows:

- (0,1,1) model had convergence problems in estimating the parameters since the moving average parameter was estimated to be very close to one (sign of non-invertibility). In addition, the t test on the parameter estimates indicated that the constant term was statistically significant whilst the moving average term did not contribute further to the explanation of the data. This suggests that a (0,1,1) model does not explain more than an estimate of the trend component of the construction growth;
- (0,2,2) model added some short-term movement in the growth estimates compared to the (0,1,1) but its performance did not improve. In addition only the constant term in the model had a significant t-value; and
- (BIC,1,BIC) model selected a high autoregressive and moving average order, which results in a non-parsimonious model difficult to relate with the underlying behaviour of the series. In addition, a different (BIC,1,BIC) model was selected when a new span of data was used. This is a sign that the Information Criteria model selection approach is not very stable with this dataset. Finally, once again, only the constant term parameter estimate was statistically significant.
- The table reporting the RMSE for the different models run on different spans shows that M1 is the worst model among the group in terms of revisions. Although the ARIMA models perform slightly better than the activity model in the first span, that does not apply to the other two spans. In fact, if data from Q1 1999 to Q2 2004 (forecasted value) are considered, the activity model performs better (for the 011 and 022 models) or as well as ARIMA.

Model identification

B3.15 Since none of the ARIMA models suggested by NIESR was satisfactory, a model research was made to identify if any other ARIMA model could be used for forecasting.

B3.16 The main tools in model identification are the autocorrelation function (ACF) together with the sample partial autocorrelation function (PACF). The two autocorrelograms were estimated for the residuals derived by removing the mean of the process from the original M3 series considering all the data available. The results are reported below:

B3.17 The two graphs show that confidence intervals are very wide due to the limited observations available, and that none of the autocorrelations are significant (outside the boundary of the two intervals). This suggests that the construction series evolves randomly around the mean of the process.

Sample Autocorrelations of the Residuals

```
      -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
        +----+----+----+----+----+----+----+----+----+----+
1                    .          |XXXX      .               0.160
2                    .     XXXXX|          .              -0.213
3                    .          |XX        .               0.098
4       - - - - - -  .          |XXXX      .  - - - - -    0.179
5                   .           |XXX        .              0.103
6                   .       XXXX|           .             -0.142
7                   .         XX|           .             -0.084
8       - - - - -   .           X|          .   - - - -   -0.024
```

Sample Partial Autocorrelations of the Residuals

```
      -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
        +----+----+----+----+----+----+----+----+----+----+
1                    .          |XXXX      .               0.160
2                    .    XXXXXX|          .              -0.245
3                    .          |XXXXX     .               0.196
4       - - - - - -  .          |XX        .  - - - - -    0.071
5                    .          |XXX       .               0.129
6                    .      XXXX|          .              -0.170
7                    .          |          .              -0.001
8       - - - - - -  .      XXXX|          .  - - - - -   -0.142
```

B3.18 All the results presented above suggest that ARIMA modelling should not be used for forecasting the construction growth rate since they do not add any extra information compared to the Activity model.

Comments and recommendations

B3.20

- There are too few observations to identify and estimate an ARIMA model. For this reason, the results of ARIMA modelling could be poor.

- (0,1,1) and (BIC,1,BIC) seems to perform better than a (0,2,2) model. However none of the models was able to track the short-term movements of the construction growth series;

- The (BIC,1,BIC) model selection method (minic option in SAS) is not a stable option to use. As a new observation becomes available a different model is fitted. In addition this method selects models with many parameters and therefore difficult to explain;

- The sample autocorrelogram and partial autocorrelogram of the distance of the construction series to the process mean do not present any significant autocorrelation. This suggests that any ARIMA models would not provide a better forecast than the mean.

- ARIMA models do not always minimise the RMSE. The activity model seems to work as well as or better than ARIMA model.

Optimal combination

B3.21 In the "Review of revisions in economic statistics" NIESR recommended to use an optimal combination of the new orders, activity balances and ARIMA predicted values to improve the quality of the early construction estimate. The following section looks at the stability of the parameter estimates of this methodology and provides comments and recommendations on its performance.

B3.22 For the purpose of this exercise an ARIMA (0,1,1) was used.

B3.23 Four regression models were run using an intercept (to correct for possible bias of the three estimates), the new orders, the activity balances and the ARIMA estimates. The first one included data from 1999 Q1 to 2003 Q3 (OEQ303), the second one included data from 1999 Q1 to 2003 Q4 (OEQ403), the third one included data from 1999 Q1 to 2004 Q1 (OEQ104) and the last one includes data from 1999 Q1 to 2004 Q2 (OEQ204). The following table reports the parameters estimated by the regression models together with their standard errors, t-values and p-values.

Table B3:

OEQ303:	***Coefficients***	***Standard Error***	***t Stat***	***P-value***
Intercept	0.2052	0.8197	0.2503	0.8057
Orders	0.0504	0.3766	0.1339	0.8953
Activity	0.9360	1.6097	0.5815	0.5696
ARIMA	0.2005	0.8318	0.2410	0.8128
OEQ403:	***Coefficients***	***Standard Error***	***t Stat***	***P-value***
Intercept	0.2375	0.7865	0.3019	0.7666
Orders	0.0346	0.3602	0.0960	0.9247
Activity	1.0062	1.5577	0.6459	0.5275
ARIMA	0.1279	0.8185	0.1562	0.8778
OEQ104:	***Coefficients***	***Standard Error***	***t Stat***	***P-value***
Intercept	0.4878	0.7639	0.6386	0.5316
Orders	0.0540	0.3633	0.1487	0.8836
Activity	0.7400	1.5441	0.4792	0.6379
ARIMA	0.0256	0.8696	0.0294	0.9769
OEQ104:	***Coefficients***	***Standard Error***	***t Stat***	***P-value***
Intercept	0.5618	0.7334	0.7660	0.4536
Orders	0.0812	0.3515	0.2310	0.8199
Activity	0.7234	1.5109	0.4788	0.6379
ARIMA	-0.0644	0.8346	-0.0771	0.9394

B3.24 From the table above it is possible to see that none of the parameter estimates has a significant t-value, a sign that none of the models explains the construction growth rate satisfactorily. However, if we consider the coefficient as the weights of the optimal combination, the Activity model is the variable that contributes the most to the optimal estimate. Nevertheless, the very poor R-square (which is less or equal to 0.11 in all the four regression models) indicates a poor goodness of fit (understandable due to the high volatility of the construction growth rate).

B3.25 The table below reports the predicted values calculated using the variable coefficients reported above together with the forecast for the remaining time periods (highlighted in yellow). At the end of the table the RMSE are also reported. They are calculated on the span of data used in the regression estimation: from 1999 Q1 to 2003 Q3 for the first model (OEQ303), from 1999 Q1 to 2003 Q4 for the second model (OEQ403), from 1999 Q1 to 2004 Q1 for the third model (OEQ104) and from 1999 Q1 to 2004 Q2 for the last model (OEQ204).

Table B4:

	Predicted OEQ303	**Predicted OEQ403**	**Predicted OEQ104**	**Predicted OEQ204**	**M1**	**Activity**	**M3**
1999Q1	0.5982	0.6312	0.7760	0.8392	0.45	0.35	0.29
1999Q2	0.6060	0.6310	0.7601	0.8057	0.23	0.34	0.38
1999Q3	1.0363	1.0738	1.0920	1.1348	0.71	0.75	0.58
1999Q4	1.1470	1.1878	1.1565	1.1766	0.34	0.86	0.78
2000Q1	0.7529	0.7481	0.8286	0.8493	0.51	0.4	3.07
2000Q2	0.8640	0.8626	0.8228	0.7443	-1.47	0.52	-1.86
2000Q3	0.6313	0.6219	0.6977	0.6816	-0.47	0.29	-1.66
2000Q4	0.6739	0.6745	0.7682	0.7875	0.12	0.35	1.52
2001Q1	0.9860	1.0061	0.9782	0.9558	-0.6	0.68	1.88
2001Q2	1.1681	1.1686	1.1056	1.0847	0.16	0.79	1.30
2001Q3	1.4241	1.4116	1.3064	1.3009	1.22	0.98	1.15
2001Q4	0.9655	0.9352	0.9117	0.8698	-0.1	0.54	2.17
2002Q1	1.1660	1.1361	1.0469	0.9865	-0.14	0.72	2.99
2002Q2	1.4975	1.4336	1.3091	1.2823	1.88	0.92	1.20
2002Q3	1.5023	1.4575	1.2849	1.2162	0.58	0.98	1.82
2002Q4	1.8484	1.8134	1.5483	1.4739	0.89	1.31	2.06
2003Q1	1.7384	1.6416	1.4768	1.4569	3.08	1.05	0.50
2003Q2	1.4248	1.3710	1.1760	1.0619	-0.44	0.9	2.05
2003Q3	2.0893	2.0436	1.7099	1.6214	1.18	1.5	1.90
2003Q4	1.8088	1.7710	1.5181	1.4413	1.7	1.2	1.50
2004Q1	1.9877	1.9504	1.6460	1.5497	1.3	1.4	0.30
2004Q2	1.7608	1.7219	1.4548	1.3398	0.5	1.2	0.74
RMSE	**1.1945**	**1.1669**	**1.1850**	**1.1662**			**.**

B3.26 If we look at the RMSE presented in the table above, it is possible to see that adding observations improves the performance of the model marginally. In addition, the RMSE calculated on the span of data between 1999 Q1 and 2003 Q3 for the optimal combination is just lower than the one calculated for the activity model, see table below.

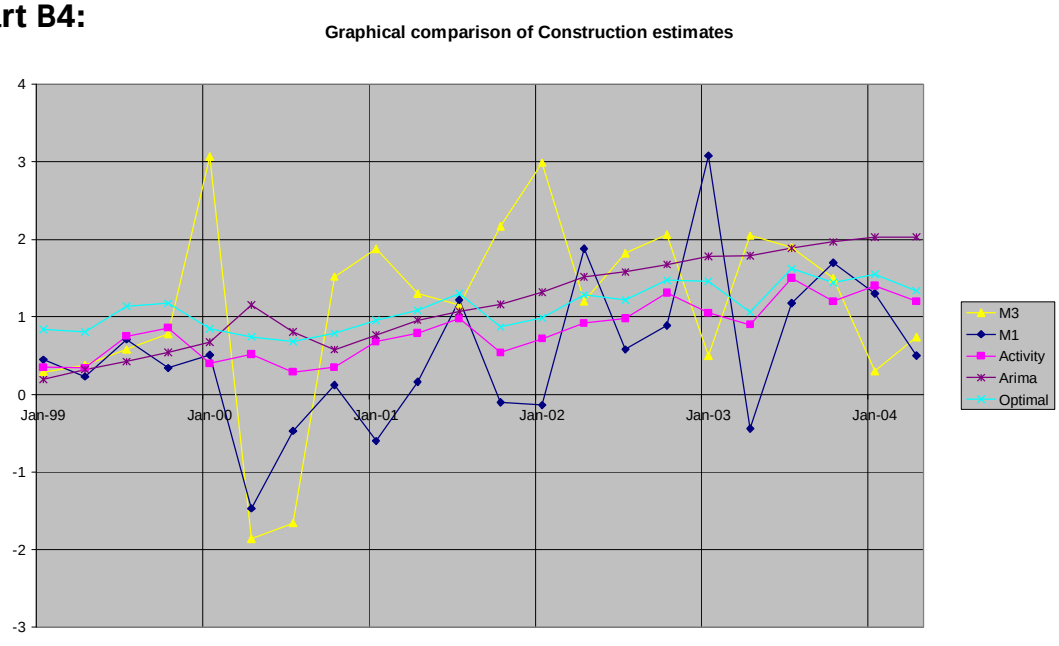
Table B5:

	RMSE
Optimal Combination	1.1945
ARIMA	***1.2403***
Orders	***1.5737***
Activity	***1.2118***

B3.27 Table B4 shows that the RMSE values are always lower than the one reported in Table B2 for any other model, but the predicted values are always further than the one provided by the order model or by the activity model.

B3.28 This is also clear in Chart B4, where the four procedures are compared with the final construction growth. In the graph it is possible to see that the optimal combination model (the light blue line) sometimes performs as poorly as all the other models. This is due to the high volatility of the construction series that makes it difficult for any sort of model to predict its short-term movement.

Chart B4:



Comments and recommendations

B3.29

- The optimal combination seems to perform better than the new orders, the activity balances and the ARIMA models if root mean square errors are used to measure the performance of the different estimation methods;

- The optimal combination model sometimes performs poorly as all the other models due to the high volatility of the construction series;

- Other possible solutions (e.g. monthly surveys) should be used to confirm or modify the quarterly estimates;

- Additional research should be carried out to improve the quality of the activity model.

Conclusions

B3.30 ARIMA models or optimal combination of different estimates do not improve the quality of DTI preliminary estimate of the construction growth rate. Other possible solutions, such as monthly surveys or further studies to improve the activity model, should be used to confirm or modify the quarterly estimates. Although the activity model performs better in terms of RMSE, it does not follow the short-term movement of the M3 series. Until more is known about the model and DTI are confident that they have the optimum combination of variables, there is no evidence that it would perform better than the present method in providing early estimates.

Annex C – Construction statistics for use in GDP estimates

Preface:

C1 This annex to the Report is the content of an analysis of different survey options, commissioned by the project team and carried out by the Methodological Group at the ONS. The report is complete and has been included as drafted by the author.

1 Introduction

C1.1 Statistics on the output of the construction industry are produced by the DTI as a quarterly, survey-based series, and also annually based on a larger survey. There are several versions of the quarterly estimates, which are produced to meet timetables for construction of versions of quarterly GDP (preliminary estimates, revised estimates and then final estimates). At the preliminary stage, there has been almost no response to the quarterly survey, and so a modelling procedure is used. This proved ineffective during one period in 2003 when an unusually low figure was forecast and incorporated into GDP, only to be subsequently revised.

C1.2 As a consequence of Stage 1 of this review, this Annex considers the suggestion that the quarterly survey is supplemented by a monthly survey. This would provide data for part of the quarter in question, and should improve the estimate of construction output used in the first estimate of GDP. This Annex outlines the options for and likely benefits and drawbacks of a monthly survey, on the basis of what might be expected to happen according to theoretical principles, tempered with some practical experience. It is not, however, based on the analysis of any data, and it would be wise to undertake such an analysis before committing to a course of action.

C2 Options

C2.1 The principle options are:

1. Full monthly survey (effectively changing the periodicity but retaining the sample size);
2. Cohort sampling, so that the current quarterly sample is retained, but one third of sample units (a cohort) are asked for a monthly response in each month of the quarter;
3. Monthly subsample – asking a small group of (probably the largest) businesses for a monthly response each month, while the remainder of the sample continues to respond quarterly;
4. Produce modelled estimates.

We now proceed to consider these in more detail.

Option 1: Full monthly survey

C2.2 This is the highest-quality solution for providing monthly estimates of construction output. It involves essentially replicating the current quarterly design but sending questionnaires to construction businesses seeking monthly information. It would be necessary to undertake a feasibility study or limited consultation exercise to determine whether businesses are able to provide output information monthly (larger businesses might be expected to have this, but smaller businesses, of which there is a preponderance in construction, may have much more difficulty). This approach would triple the number of questionnaires sent out in a quarter, and hence approximately triple the costs, both to DTI and to businesses.

C2.3 Using this approach we might expect to achieve a standard error on the monthly change which is similar to the standard error of the existing quarterly change. Because the number of businesses sampled quarterly is not affected (only the frequency of collection), the standard error of the quarterly change would also be unaffected. However, it is really the usefulness of the first two months in the quarter as a basis for predicting the quarterly figure which is at issue, and that cannot be assessed without a pilot study to collect some information monthly.

C2.4 There are some possibilities for reducing the cost from strictly three times as much, but they are either marginal reductions or have unpalatable side effects. One possibility is reduce the number of questions asked monthly, perhaps to a single question on output, with the full range of detail asked only quarterly. This would marginally reduce compliance, but still require substantial additional resources within DTI because the number of questionnaires and periodicity of processing would not be reduced. A second possibility is to reduce the sample size of the monthly survey. However, this would have an adverse effect on the accuracy of the measurement of both monthly *and* quarterly change. Other variations approach options 2 and 3 below, where they are discussed in more detail.

C2.5 Advantages

- high quality data for the first two months of a quarter from which to forecast the quarterly activity
- smallest variance of monthly change of the options considered here

C2.6 Disadvantages

- approximately triples the current survey costs (both to DTI and compliance cost for businesses)

Option 2: Cohort sampling

C2.7 In this option, the current quarterly sample is retained, but the way in which information is collected is changed, so that one-third of the sample (a cohort) is asked for *monthly* information in each of the three months of the quarter. This has a reduction in the coverage of the activity compared

with the quarterly response, in that only a third as much activity information is collected. This means that there will be a reduction in accuracy of the quarterly estimate (as a greater proportion of the quarterly activity is unknown and must be estimated). Again, the feasibility of collecting information monthly from construction businesses would need to be tested.

C2.8 There are consequences of this approach for the accuracy of the monthly estimates produced too, since the completely enumerated (one-in-one) strata would need to be split between the three months. Therefore the accuracy of the individual monthly estimates would be rather low, and the accuracy of the change between months would be very poorly estimated as we would not take advantage of overlap between units in consecutive months; this may not matter if the monthly estimates were considered solely as auxiliary (additional) information for forecasting the quarterly change, but I suspect that once a monthly series, however crude, becomes available, there will be pressure to have it available for analysis and policy formulation/monitoring, even if it is unsuitable for this purpose.

C2.9 There are some possibilities for reducing the disadvantages of this scheme, mainly by forming a hybrid with option 1 or 3. Specifically, all the largest businesses could be included *every* month, increasing both costs (compliance and processing costs) and the accuracy of the monthly estimates of level and change, while adopting the approach above of asking only a third of the smaller businesses in each month. There would still be a reduction in accuracy over the existing quarterly estimates in this case.

C2.10 Advantages

- some information available for first two months of quarter from which to forecast quarterly activity

C2.11 Disadvantages

- reduction in accuracy of quarterly estimates
- monthly estimates relatively inaccurate, since completely enumerated businesses divided between three months

Option 3: Monthly subsample

C2.12 In this option the principle is to retain the existing quarterly sample for most businesses, but to ask a subsample of businesses for monthly responses (which could then be aggregated to make a quarterly response). The obvious approach is to have the monthly sample as a panel formed from the largest businesses, which would ensure maximum coverage of activity for a given (presumably small) number of questionnaires. These are also the businesses most likely to have information available to be able to respond monthly. This approach would give an indicator (which would quite likely be a biased indicator) of the level of construction output in a particular month, but which, assuming that the bias was consistent from month to month,

would quite likely form a good basis on which to make a forecast of quarterly activity. This option could be relatively inexpensive, although does require both additional questionnaires and monthly processing. A feasibility study would be needed to check that businesses were able to provide the information monthly.

C2.13 The extent to which this approach would work in practice would depend on how consistently the construction sector reacts to economic conditions. That is, do the larger and smaller businesses have different responses at different stages of the economic cycle? If so then using the large businesses to represent what happens to the small ones could be misleading, and this could be a particular problem in construction where there are very many small businesses. Evidence for whether this happens should already be available by analysing the pattern of growth and decline in output in the quarterly survey. If this is a problem, the "biggest businesses" subsample could be replaced with a representative subsample, but with additional costs from the need for a larger number of questionnaires (as more smaller businesses would be needed for similar coverage), and more likelihood that smaller businesses will not have monthly information available.

C2.14 The approach of using a representative subsample has some advantages over the panel of the largest businesses. Firstly, it provides an opportunity to detect when smaller businesses are no longer trading, particularly useful if downturns in construction are more likely to reduce the number of businesses than to cause a reduction in output across all the existing businesses; and secondly, it allows a variance of the monthly estimate (and its change) to be calculated and therefore for some measure of the quality of the monthly series to be provided.

C2.15 Note that there is no reduction in the accuracy of the quarterly estimate because the quarterly sample is unchanged.

C2.16 Advantages

- Relatively inexpensive
- Good basis for forecasting quarterly movement

C2.17 Disadvantages

- Biased estimates of construction output monthly (so monthly series should not be analysed in isolation)

Option 4: Model construction output

C2.18 This is essentially the same as the current approach, where an estimate of quarterly construction output is made based on a model. Some work on better forecasting models has already been undertaken as part of the Construction Statistics review. The other approaches that could be taken involve use of additional information from other sources – some form

of econometric modelling. The drawback of this approach is that it is not based on current data, and as such is prone to making poor estimates at the time when hood estimation is most crucial – when there is a turning point or when patterns of behaviour in the construction industry are changing. Other available sources are correlated economic variables or administrative sources, the first being available relatively quickly but providing only some information on construction, the second typically being lagged to an extent which means they are less useful for making short period (monthly or quarterly) estimates.

C2.19 Without a clear knowledge of the potential alternative sources of indicators and correlated measures, it is not possible for me to make a sensible comment on what sorts of models, if any, might work.

C2.20 Advantages

- Cheap (no additional data collection)

C2.21 Disadvantages

- Potentially produces poor quality results when decision-making requires them to be most accurate.

C3 Additional issues

C3.1 DTI say that it is the largest businesses which typically take the longest to respond. In any monthly collection , speedy response is needed. It make take time, even for compliant businesses, before their internal systems can be developed to make the required data available to an accelerated timetable. Without acceleration only one month of a monthly series would be available from which to forecast the quarterly estimate.

C3.2 The options above focus mainly on how to collect some information monthly. Having a monthly series (or a monthly indicator of the quarterly series) is not the whole story, as there is still a need for a model to forecast the quarterly value before all the months in the quarter are available. This is a time series modelling/forecasting problem, which then needs to be tackled once the monthly information is available.