

## NSMAC 14 - minutes

### Committee members present

Martin Brand	ONS	Stephen Penneck	ONS (Chair)
Jelke Bethlehem	Statistics Netherlands	Chris Skinner	Southampton University
John Copas	University of Warwick	Sandy Stewart	Scottish Government
Harvey Goldstein	University of Bristol	Kenneth Wallis	University of Warwick
Rachel Leeser	Greater London Authority	Martin Weale	NIESR
Peter Lynn	University of Essex		
Frank Nolan	ONS		

### Presenters

Rob Elder	Bank of England	James Mitchell	NIESR
Jana Eklund	Bank of England	Neil Parkin	ONS
Salah Merad	ONS		

### Others present

Tullio Buccellato	ONS	Denise Silva	ONS
Colin Mowl	ONS	Paul Smith	ONS
Julian Chow	ONS	Nigel Studdard	ONS
Philip Clarke	ONS	Denise Williams	ONS (secretary)
Louisa Nolan	ONS (secretary GSS MAC 15)	John Wood	ONS
		Robin Youll	ONS

### Apologies

David Hand	Imperial College London
------------	-------------------------

## Introduction

Stephen Penneck welcomed everybody, and introductions were made. The minutes of the 13<sup>th</sup> NSMAC meeting were accepted without comment.

He commented on the two Deputy Chairs recently appointed to the Statistics Board. Lord Rowe-Beddoe has been appointed Deputy Chair with responsibility for the governance of the Office for National Statistics (ONS), and Professor Adrian Smith has been appointed Deputy Chair with responsibility for promoting and safeguarding the production and publication of UK official statistics.

He noted that funding has been agreed for the Census fourth page. Stephen Penneck reminded the committee that the 13<sup>th</sup> GSS Methodology Conference will take place on the 23<sup>rd</sup> of June, 2008. This is primarily for GSS members, but external delegates are welcome. More details can be found on their website:

[http://intranet/news/corp/2008/05\\_may/thirteenth-gss-methodology-conference-23-june-2008-corp.asp](http://intranet/news/corp/2008/05_may/thirteenth-gss-methodology-conference-23-june-2008-corp.asp)

Comments were invited.

Rachel Leaser asked whether departments other than ONS would be treated differently in the new UK Statistics Authority, particularly with respect to the rules for releasing data. Stephen Penneck replied that the Statistics Authority is producing a new Code of Practice for consultation. It may not be very different, and the assessment process is to be the driver for consistency. He asked for anyone with any issues about this to please bring them up. Colin Mowl said that Adrian Smith is to address issues of consistency, and noted that the Statistics Authority only has executive authority over ONS, but it will be able to exert its influence over other government departments. Stephen Penneck and Martin Brand clarified that all major government statistics can now only be released through the UK Statistics Authority. It was noted that work was in progress for the Statistics Authority publishing hub, and that, as a consequence of the review following the data loss from HM Revenue and Customs, different data release practices are now found in different departments.

John Copas asked who is selling or buying the new Methodology Consultancy Service. Stephen Penneck responded that most departments have statistical groups that are too small to do their own methodology, and that the consultancy will be a sustainable career path for London statisticians.

Sandy Stewart asked about the assessment process. Stephen Penneck said that he did not yet know how this would be treated, but discussion was ongoing between the Statistics Authority and departments. Frank Nolan pointed out that the focus was on classification of outputs as National Statistics, and that most outputs would have their assessment completed in a 3 - 5 year period. It was noted that the relevant minutes and discussions would probably be published online, to satisfy transparency concerns.

Peter Lynn asked about progress from NSMAC13, in particular about the paper on migration statistics. It was agreed that a brief progress report would be circulated.

## Actions for secretary

2a	to obtain and circulate to the committee a progress report on the NSMAC 13 paper <i>Improving statistics: a review of survey data collection at UK ports</i>
----	--

## 2.1 Qualitative Business Surveys: signal or noise?

<b>Authors</b>	Silvia Lui  James Mitchell  Martin Weale	National Institute of Economic and Social Research  National Institute of Economic and Social Research  National Institute of Economic and Social Research
<b>Presenter</b>	James Mitchell	
<b>Discussant</b>	Jelke Bethlehem	Statistics Netherlands

This paper compares the Industrial Trend Survey (ITS), a firm-level qualitative business survey produced by the Confederation of British Industry (CBI), with the Office for National Statistics' (ONS) Monthly Production Inquiry (MPI), a firm-level quantitative survey. ITS is published at the end of every target month, whereas MPI is published 25 days after the end of the target month. The motivation is to investigate whether the qualitative data can be used to improve the quantitative results, particularly with respect to timeliness. Questions arising from the paper are listed below.

- What is the quality of the qualitative data?
- How do we explain the relationship between the qualitative and quantitative data?
- Is ITS a good indicator of MPI?

### Discussion

Jelke Bethlehem began by noting that this was interesting research, and mentioned that he is involved in social, rather than business, surveys. He went on to make the following points.

- Is ITS a reliable predictor of MPI? Given the differences between the surveys, Jelke was not convinced. The two surveys have different sample selections, sample sizes and questions, and the response rate and methodology for ITS is not known. MPI is a non-seasonally adjusted level series, whereas ITS is a seasonally adjusted difference series. MPI covers only the target month, whereas ITS asks questions about the last 3 – 4 months. From his experience in conducting social surveys, Jelke pointed out that it is difficult to ask what happened in the past as respondents will not interpret the questions consistently,
  - The ITS has a 73% match rate with MPI on the Interdepartmental Business Register (IDBR). Is this matched sample representative of the whole population? Are the questions and estimates similar in both surveys? More tests should be done here.
  - Are the approaches to the seasonal adjustment of the two surveys comparable? ITS is seasonally adjusted via the questions asked, and MPI is seasonally adjusted in this paper, by introducing seasonal dummies. A further concern is that ITS asks about changes, whereas MPI is focussed on levels.
    - In the paper, growth rates are corrected in order to control outliers. However, some outliers are real, so is this the best approach? What happens if outliers are left in?
    - Jelke asked to see some simple descriptive statistics that show the correlation matrix between the movement of MPI and ITS, and the distribution of the change in MPI with respect to the movements of ITS. Quoting the serial correlation between ITS output trend and MPI turnover growth by month in Figure 1 of the paper, he observed that there is no correlation in the first

month, that it looks like respondents forget what has happened more than 6 months ago, and that there is an upward movement in month 12. What are the reasons behind this behaviour?

- On the modelling issues, Jelke asked whether, given that it is difficult to interpret the results about the time period used in the ITS, a 3-4 month lag should be included when assessing the reliability of the ITS data and when predicting ITS data from MPI results. He also commented that, although there was a lot of hypothesis testing in the paper, more information is needed about the confidence levels.

In his final remarks, Jelke queried how useful the results of this project are, and whether the ITS data is sufficiently reliable. He asked whether it would be advisable to conduct more ITS-style surveys, for rapid but less statistically sound indicators, and if this would be worth the extra burden to businesses. He finished by saying that it might be more interesting to predict MPI from ITS, rather than vice versa.

Comments were then invited from other committee members.

Chris Skinner was concerned about the interpretation of the econometric model. He suggested that as well as the t-statistics already presented, the paper should present the actual coefficients, in order to understand the underlying structure. He also questioned the interpretation of the trend question, asking whether the ITS data, which lags behind the MPI data, actually picks up the current month's data. He suggested that actual turnover (rather than difference) is investigated, perhaps using a moving average filter.

Stephen Penneck questioned the quality of the ITS survey questions, asking if any cognitive work had been done. This was not known. Various members ascertained that it is also not known who fills in the survey questionnaire, and that it could be someone quite different from the MPI respondent, which could be another potential source of incompatibility.

Ken Wallis would like to see more focus on how the time lag questions are answered in ITS, and echoed Chris Skinner's concerns that monthly differences might not be the best method of comparison. He also pointed out that the striking increase seen in month 12 of the serial correlation between ITS output trend and MPI turnover growth (Figure 1 of the paper) suggests that seasonal variations have not been removed from the series.

Sandy Stewart then raised several issues. First, he asked whether the 5 year comparison period was well-behaved, and whether the smoothed series would pick up any variations if the comparison was conducted over the whole of the economic cycle. Second, he questioned whether it is appropriate to use a smoothed series to predict a smoothed series. Finally, he queried whether MPI is actually a good predictor of, for example, turnover.

James Mitchell responded.

He replied to the committee members' queries on the reliability of the ITS data by suggesting that the way to assess this is to invert the relationship, i.e. by using the MPI data to nowcast the ITS data, using a Bayesian approach and the benchmark technique.

He noted that ITS is the property of the CBI, so gathering information about it was difficult. He acknowledged that it is less statistically sound than MPI.

James went on to suggest a new exercise. He would investigate the correlation of ITS responses at firm level (instead of aggregate level as at present). A principal components analysis would then be conducted to look at the factors driving the responses. He notes, however, that data is only available as far back as 2000.

Paul Smith then suggested that the standard seasonal adjustment programme could be used on the ITS data at aggregate level to find and remove any residual seasonality. It was stated that the seasonal adjustment programme has a problem running on qualitative data.

Ken Wallis commented that the lack of reference to the sample design of ITS was an important omission, although he acknowledged the difficulty of eliciting a full answer from the CBI. He also worried that there might be bias in the CBI survey, as small, new, fast-growing firms may not be members of the CBI.

Rachel Leeser asked whether the two surveys covered industries with the same Standard Industry Codes (SIC). James Mitchell responded that, although the ITS definition may not be as precise as the ONS definition, only SICs in MPI were included in the comparison. Rachel Leeser went on to ask whether output was the same as turnover, to which Stephen Penneck responded that they were not. MPI deals with turnover, and ITS with output, and this difference is especially significant at times of change.

**Suggestions to authors:**

<b>2.1a</b>	further investigate how the matched-pair sample represents the complete ITS / MPI samples
<b>2.1b</b>	investigate different treatments of outliers
<b>2.1c</b>	further investigate seasonal adjustment in ITS and respondents' perceptions of the time-lag questions
<b>2.1d</b>	include more statistical measures, especially confidence levels
<b>2.1e</b>	predict MPI from ITS
<b>2.1f</b>	to use actual turnover values instead of differences
<b>2.1g</b>	investigate correlation in ITS responses at firm level, and use PCA to uncover the factors driving the answers

**2.2 Comparison of Average Weekly Earnings and the Average Earnings Index**

<b>Author</b>	Neil Parkin	ONS
<b>Presenter</b>	Neil Parkin	
<b>Discussant</b>	Martin Weale	National Institute of Economic and Social Research

Neil Parkin introduced this paper by expressing his thanks to Harry Duff, who computes both the Average Weekly Earnings (AWE) index and the Average Earnings Index (AEI) on a monthly basis. He introduced the two indices, and explained that the Bank of England and the Treasury were concerned about divergence between them. He illustrated this divergence with a graph of 12-month growth rates, and described the various differences between the two statistics which might give rise to it. He said that ONS had just completed a programme of work to explain these divergences, and identified seven major terms in the conversion from AWE to the AEI. These terms are then applied step by step, to produce reconciliation tables that identify the contribution to the divergence from each term. Two of the most significant steps were briefly presented, i.e. the employment composition term and the outlier term. The importance of the order in which the terms are applied was also discussed. It was stated that analysing monthly growth rates rather than 12-month growth rates made it easier to identify the

sources of observed differences, and reduced the importance of term order. Advice was sought from the committee on the following two points:

- whether the method used was sound
- what ways of presenting the results would be most suitable

## Discussion

The discussant, Martin Weale, made the following comments.

- He recommends an approach differing from Neil's with respect to the composition term. The AEI can be specified by a term relating only to earnings growth. However, AWE contains the earnings growth term but also a term relating to employment changes and a second order, residual term. This provides a sound basis for analysing the differences between AEI and AWE.

- For the earnings growth term, the AEI and AWE use different employment weights. This is designed into the statistics, partly because they are targeting different population characteristics and partly because of different statistical judgments about such matters as outlier treatment and the use of matched pairs. The earnings growth term sometimes accounts for almost all the difference between AEI and AWE, although at other times it does not. The use of matched pairs in AEI accounts for some differences, but Martin was less concerned about the effect of randomly excluding units through the process of rotation than the effect of occasional non-response, which may lead to bias. Different outlier treatments also lead to differences in observed growth rates. Martin suspects there is a downward bias in AEI growth rates because of its treatment of outliers, and that this may create a downward bias in estimates of GDP, of which AEI is a component. The automatic treatment of outliers in AWE can lead to large revisions to published estimates.

- Re-classifications on the Inter-Departmental Business Register (IDBR) can generate some large changes in the employment term. The effects of the second order, residual term are usually very small.

In summary:

- AWE is better logically;
- issues with re-classifications in the IDBR need to be better understood;
- a sound statistical basis for imputation and outlier treatments is required.

Comments were invited from other committee members.

Chris Skinner suggested that there may be a natural term ordering according to whether the terms are pure weighting terms (such as employment composition), intermediate (such as different stratum classifications) or not related to weighting (such as the pay variable used).

Martin Brand added that looking at the monthly growth rates is very valuable because it highlights specific effects. It is of concern that they often happen in January, when the classification changes. He would like to see resources put into looking more closely at the effects of changes to classifications / cell boundaries which have occurred in January.

Paul Smith was surprised by the complexity of interactions between terms, which means that the order of the terms affect the analysis dramatically.

Harvey Goldstein suggested that it may be worth looking for patterns in the different term orderings. The effect of ordering may be reduced if some terms are grouped together, perhaps along the lines of Chris Skinner's suggestion.

Ken Wallis said that using annual rates of growth is a primitive form of seasonal adjustment. The annual growth rates are also less noisy than monthly growth rates, where the noise may disguise any real differences. He asked whether analysing the seasonally adjusted series might help.

Paul Smith suggested that complicating the analysis by introducing further differences from the seasonal adjustment should be avoided.

Martin Weale said that analysing monthly growth rates removes the 12-month persistence in annual growth rates of one-off effects.

Ken Wallis pointed out that it should be the growth rates actually used by customers that should be analysed. For econometric analysis, this may be the monthly growth rate, but for policy makers, it should be the growth rate for the last three months compared with the same three months from the previous year, as used by the Monetary Policy Committee.

Martin Weale emphasised that there are two separate questions here: how similar are AWE and AEI to each other? And how similar are these two statistics to other measures of earnings growth? In fact, AWE and AEI are more similar to each other than to other earnings measures.

Neil Parkin responded.

He reiterated that his method of economic decomposition splits the employment decomposition term into two parts. He remains to be convinced of the logic of Martin Weale's suggestion that one of these terms be fixed in the first position, but is happy to provide the reconciliation in either of the two forms.

**Suggested actions:**

<b>2.2a</b>	determine if the use of matched pairs is appropriate for AEI/AWE, and, if possible, establish conditions for the appropriate use of matched pairs for other short term indicators
<b>2.2b</b>	investigate further the effect of outlier treatments
<b>2.2c</b>	further study of the effect of the order of terms in the reconciliation, including any evidence from other studies that this effect might be expected
<b>2.2d</b>	fuller investigation of changes in classifications / boundaries of sample cells which have effect in the first month of each year

## 2.3 A state space approach to extracting the signal from uncertain data

<b>Authors</b>	Alastair Cunningham	Bank of England
	Jana Eklund	Bank of England
	Chris Jeffery	Bank of England
	George Kapetanios	Bank of England & Queen Mary University of London
	Vincent Labhard	European Central Bank
<b>Presenters</b>	Rob Elder	Bank of England
	Jana Eklund	Bank of England
<b>Discussants</b>	Kenneth Wallis	University of Warwick
	Robin Youll	ONS

Early estimations of Gross Domestic Product (GDP) tend to be revised upward. The paper presented was a response to a request by the Monetary Policy Committee (MPC) for a model to provide an alternative representation of output growth, which gave a representation of the uncertainty around early GDP estimates and which takes account of available information.

### Discussion

Ken Wallis offered the following points for further discussion:

- For ONS: can all the types of revision, including, for example, changes in methodology and re-classifications, be statistically modelled?
- For the authors: what diagnostic procedures can be implemented to check the validity and stability of the model? Can the model *predict* revisions, and which components of the model are the main drivers? Ken Wallis advised that evaluation of the model is difficult in a situation where the 'true' values of the data will never be known. Rob Elder responded that they are explicitly forecasting the end result of the revisions, and that he appreciated the difficulty of evaluation.
- What should be the allocation of labour between ONS and the Bank of England? The Bank of England paper and Quarterly Bulletin article stated that it was probably better for users to make adjustments to the ONS data to take account of bias, although the presenters could see arguments on either side. But Ken Wallis believes that ONS should have a greater role in this. He stated that the judgement about how much of the past pattern of revisions is representative of current uncertainty is not only an economic, but also a statistical issue. A co-operative approach between the two institutions is the right approach.
- Revisions due to changes in methodology and annual bench-marking are expected annually. Published results should include a full explanatory statement to this effect.

The second discussant, Robin Youll, added the following.

- The persistent bias in GDP estimates is reflected throughout OECD countries (except the US).
- ONS already provides an assessment of measurement error in the metadata published with outputs.
- The third month revision (M3) tends to add noise. Is the economy really noisy?

- External surveys can often influence ONS judgements. This is not done formally, but the ONS do compare its results with that of external surveys, which often show a strong correlation with ONS surveys. However, it should be remembered that external surveys tend to be based on significantly smaller samples than ONS surveys, and do not undergo revisions.
- Robin Youll liked the real-time evaluation of data in the paper, and the historical fan diagram. It is difficult to calculate standard errors for GDP, due to the complexity of its production, but perhaps Monte Carlo simulations could be used.
  - It would be interesting to test the model on other series.
  - He offered the following as possible contributors to the upward bias:
    - the earliest estimates use only ~40% hard data
    - changes in methodology
    - growth increases when services measures are improved. Services is a dynamic sector.
    - estimates on growth may have been too conservative in the past
  - ONS is (and should be) more concerned with improving estimates rather than forecasting revisions.
  - ONS will continue to support the publication of metadata with outputs.

Comment was then invited from other committee members.

Martin Weale argued that ONS needs to be open about the magnitude of past revisions, or risk bad press. He was concerned that the current division of responsibility could make it appear that the Bank of England was criticising ONS in difficult times, and that therefore ONS should keep the revisions process in-house. He added that Kenneth Wallis' comments about the model procedure are applicable to any time series model, not just the one proposed. Finally, he said that ONS should be sympathetic to not just looking for causes of bias, but also to implementing corrections.

Sandy Stewart raised three points. Firstly, he identified that the most important factors contributing to the revisions were methodological changes. He warned that there were significant changes due on the near future (e.g. chain linking, the change to new industry classification codes). Secondly, he pointed out that there may have been different changes in the upper and lower parts of the series. Finally, he questioned whether the assumption of stable decaying variance was realistic.

Paul Smith queried the simulation exercise in the paper. It uses the model to simulate the data, and then fits the data, so it is perhaps not surprising that it does fit. He also suggested that the difference between the predicted and observed values, and not just the standard errors of the predictions, should be evaluated.

Colin Mowl noted that although each revision has a different explanation, the revisions were still all consistently in the same direction. He went on to say that there seems to be a lot of latitude and judgement in describing the economy, and while this might be suitable for the Bank of England, who need to take into account all the data sources available, ONS would probably want a more restricted approach with less scope for individual judgement

Kenneth Wallis then said it was a question of where the divide should be. Forecasting already happens at ONS, but ONS could be more informative to its users over what has been done, and what is likely to occur in the near future.

Martin Weale pointed out that the Bank of England was not the only user of ONS data.

Rachel Leeser was concerned about transparency and confidence. If ONS introduced 'judgement' in producing the GDP, what would stop it being used in other ONS outputs? She stressed that subjective judgements should be minimised.

Stephen Penneck pointed out that judgement is used around models, even if not elsewhere. ONS says quite a lot about future developments, but it is not quantified.

Colin Mowl said that ONS can give some reasonable quantitative guidance to the revision of GDP, and it is gradually responding to feedback and increasing the amount information made available to customers. We could see good arguments for minimising the role of judgement in producing GDP estimates. But the Bank was keen to see the ONS continue to improve its methods and hopefully reduce bias in future estimates.

Rob Elder concluded the session with his response to the comments. He said that he could see good arguments for minimising the role of judgement in producing GDP estimates. But the Bank was keen to see the ONS continue to improve its methods and hopefully reduce bias in future estimates.

He said that he was confident that the techniques used by the Bank were sound, but he was concerned that the sample of real time data was small, and there was always a possibility that past experience gives a poor guide to the future.

### Suggestions to authors:

<b>2.3a</b>	to test the model on series other than GDP
<b>2.3b</b>	to investigate further how well methodological revisions can be modelled statistically
<b>2.3c</b>	to evaluate the difference between predicted and observed values, in addition to the standard errors
<b>2.3d</b>	to test the model on the economic slowdown

### 2.4 Small domain estimation in business surveys

<b>Authors</b>	Salah Merad	ONS
	Pete Brodie	ONS
<b>Presenter</b>	Salah Merad	ONS
<b>Discussant</b>	Chris Skinner	Southampton University

Salah Merad introduced the paper, which represents the start of a small-scale project as part of an initiative in the Methodology Directorate in response to the Allsop Review (2004). The review stressed the importance of producing estimates at a local level. The paper was intended to be a start-of-project advice-gathering paper. In it, the issues arising when model-based estimates are used for small domains, where direct estimation lacks precision, are discussed. The following questions were put to the committee by the presenter.

1. With respect to producing small domain estimates, when is the nested error unit-level model preferable to the area-level random effects model?
2. Is it appropriate to fit models separately in different groups, for example, by 2-digit standard industry classification (SIC) or by broad geographical regions?
3. Given that the mean-squared errors (MSEs) of design-based direct estimators and small area estimators are defined in different theoretical frameworks, is it valid to compare their MSEs?

4. Would raking/calibration help in obtaining estimates that are consistent at different grouping levels?
5. Is the estimation of the MSE of the ratio of two empirical Bayes estimators using re-sampling methods appropriate?
6. Are there any models/methods that could be considered for estimating proportions in domains with no sample data, which do not involve information on every individual in the population?
7. Would using modelled data as a covariate in a small area model give rise to technical problems in the derivation of an estimator and the estimator of its MSE?
8. With respect to small areas in particular, should we consider apportioning aggregate estimates, which are based on returned data, to geographical areas in preference to producing estimates based on estimated microdata?

Chris Skinner gave the following response.

The highly skewed nature of business survey variables can cause distinctive features to arise, which depart from those seen in social surveys. Therefore, the same methods cannot necessarily be directly applied. Nevertheless, he felt that powerful fully-enumerated covariates were potentially available. These might be variables from the business register or from tax records. He agreed that the subject of small area estimation in business surveys is potentially useful, and encouraged the authors to pursue this research. He then responded to the questions in turn.

1. Nested errors would be the more general choice over area-level models. However, it was important that sample data (used as covariates) matched the definition used for the variable at area level (from other sources). Robustness would be gained by fitting a model at the lower level but calibration would be more of a problem. The calibration procedure is easier for area level model. The availability of 100% PAYE information would be very useful for the modelling. There should be a simple comparison of the two modelling methods. However, Chris Skinner pointed out that neither model would be suitable for estimation of medians or quartiles.
2. Where groups (e.g. grouping by SIC) were sufficiently large to give good survey estimates, fitting models separately by large groups could be useful. This might also make the model more stable over time.
3. Comparisons could be made between MSEs estimated from the model-based and design-based frameworks. However, the design-based MSE estimations break down under small sample sizes.
4. Using calibration weights will allow consistency to be obtained in estimates at different grouping levels.
5. Estimation of the MSE of the ratio of two empirical Bayesian estimators can be obtained using re-sampling methods. Chris Skinner suggested using the boot-strap method.
6. A domain-level model for proportions could be used to predict values in empty domains.
7. Using modelled data as a covariate in a small area model does not affect the validity of the model, but may affect the precision obtained.
8. Modelling could be used to apportion the aggregate estimates,

Overall, Chris Skinner offered positive encouragement for this work.

The discussion was then opened up to the committee.

Harvey Goldstein made the following points.

- Spatial correlation between areas can be important. If this is ignored then this could lead to biased estimates..
- If unit level covariates are available, then the nested error model must be the more efficient.
- If there are some domains without sample data or with covariates missing, (Q6) then it would be useful to try imputation models. However, this would make the work more complex.
- If the coefficients vary across the areas, then random slope models could be tried.

- When examining the ratio of two empirical Bayes estimators (Q5), it is important to obtain these from a joint model rather than from two separate models. Also, care must be taken with very skewed distributions, for example in large cities. The extremes may need to be modelled using dummy variables.

Sandy Stewart pointed out that industry classifications could be very different at a local level from those applied at the reporting level. Businesses can have very different processes at different locations.

Stephen Penneck commented that it is important to manage users' expectations, especially with respect to small area estimation.

Rachel Leeser noted that PAYE data is in fact incomplete. For example, non-UK tax-payers are not included. This could lead to bias. She also expressed her concern over how outliers were treated.

Martin Weale supported Rachel Leeser's concern about outliers, pointing out that some are real, rather than errors.

Paul Smith noted that if the estimation worked well for local areas, there could be potential disclosure issues.

Chris Skinner was of the opinion that if only the model-based estimates are used, disclosure control requirements would be satisfied.

#### **Author actions and suggestions**

<b>2.4a</b>	the nested error model is the generally preferred choice
<b>2.4b</b>	to perform a simple comparison of the design- and model-based frameworks
<b>2.4c</b>	to investigate fitting models separately by large groups
<b>2.4d</b>	to compare the MSEs estimated from the two models
<b>2.4e</b>	to investigate the treatment of outliers

## Conclusion

Stephen Penneck concluded the meeting with some remarks on the next meeting of the committee.

It was agreed that NSMAC 15 will take place on the 11<sup>th</sup> of November, 2008, and will be at Myddleton Street. Feedback on NSMAC 14 is requested in advance of NSMAC 15.

Early suggestions for the agenda for NSMAC 15 are invited from the committee. Papers are invited particularly from ONS and the Government Statistical Service (GSS). Stephen Penneck will raise the issue with the Divisional Directors.

Suggestions were made for topics for NSMAC 15.

Frank Nolan suggested:

- disclosure control;
- data visualisation.

Rachel Leeser suggested topics from the list suggested at NSMAC 13, namely:

- mortality rates / life expectancy;
- measuring government output and activity, possibly in the fields of health or education.

### Committee actions:

<b>2b</b>	Stephen Penneck	to contact the Divisional Directors for suggestions for NSMAC 15
<b>2c</b>	committee members	to communicate any feedback on NSMAC 14 and suggestions for GSS MAC 15 in good time

**Summary of actions and suggestions:**

<b>Section</b>	<b>Participant</b>	<b>Action</b>
<b>2.a</b>	GSS MAC secretary	to obtain and circulate a progress report on the NSMAC 13 paper <i>Improving migration statistics: a review of survey data collection at UK ports</i> to the committee
<b>2.1</b>	Silvia Lui James Mitchell Martin Weale	to further investigate how the matched-pair sample represents the complete ITS / MPI samples to investigate different treatments of outliers to further investigate seasonal adjustment in ITS and respondents perceptions of the time-lag questions to include more statistical measures, especially confidence levels to predict MPI from ITS to use actual turnover values instead of differences to investigate correlation in ITS responses at firm level, and use PCA to uncover the factors driving the answers
<b>2.2</b>	ONS Methodology Directorate	determine if the use of matched pairs is appropriate for AEI/AWE, and, if possible, establish conditions for the appropriate use of matched pairs for other short term indicators to investigate further the effect of outlier treatments to further study of the effect of the order of terms in the reconciliation including any evidence from other studies that this effect might be expected fuller investigation of changes in classifications / boundaries of sample cells which have effect in the first month of each year
<b>2.3</b>	Alastair Cunningham Jana Eklund Chris Jeffery George Kapetanios Vincent Labhard	to test the model on series other than GDP to investigate further how well methodological revisions can be modelled statistically to evaluate the difference between predicted and observed values, in addition to the standard errors to test the model on the economic slowdown
<b>2.4</b>	Salah Merad Pete Brodie	the nested error model is the generally preferred choice to perform comparison of design- and model-based frameworks to investigate fitting models separately by large groups to compare the MSEs estimated from the two models to investigate the treatment of outliers
<b>2b</b> <b>2c</b>	Stephen Penneck committee members	to contact the Divisional Directors for suggestions for NSMAC 15 to communicate any feedback on NSMAC 14 and suggestions for NSMAC 15 in good time

