

MORTGAGE MARKET REVIEW

WELFARE ANALYSIS PEER REVIEW FOR FINANCIAL SERVICES CONSUMER PANEL

Jon Stern
CCRP, Department of Economics
City University, London

1 Introduction and Summary Conclusions

This peer review has been commissioned by the Financial Services Consumer Panel. I have been asked to consider the wellbeing analysis contained in Annex 1, A4 Section H, Pages A 1:75 to A1:86 of the FSA's December 2011 Mortgage Market Review Consultation Document, CP11/31. The review will focus on the Annex section in question but it will, where necessary, refer to other relevant material either in Annex 1 or in the published Consultation Document.

1.1 The remit for the review

In carrying out this review, I have been asked to consider:

- (i) The robustness of the methodology used by the FSA to capture the potential welfare benefit of those affected by the proposals;
- (ii) The robustness of the FSA's quantification of welfare costs and benefits; and
- (iii) The reliability and implications of other methods, including classical welfare analysis that could be used to capture consumer preferences.

I have also been specifically asked to consider the impact on sellers of houses as well as on buyers and what impact a more explicit consideration of the impact on house sellers might have on the welfare estimates reported.

1.2 What this peer review covers

The main focus in what follows is on the methodology employed in Annex A4 Section H. It considers:

- (a) The reliability of the chosen methodology in principle;
- (b) The way that the chosen methodology has been applied;

- (c) The robustness of the reported estimates obtained in the FSA’s application of the chosen methodology, including potential biases, margins of error, uncertainties, etc; and
- (d) The potential for the application of alternative methods of appraisal and what difference they might make.

In carrying out the review, I have paid particular attention to HM Treasury Green Book on “Appraisal and Evaluation in Central Government”. This has a recent addition (July 2011) on the valuation of non-market goods which draws on a Green Book Discussion Paper by Fujiwara and Campbell (2011). This Discussion Paper, jointly issued by HM Treasury and DWP, provides a thorough survey of the relative methods that can be applied in this area with their advantages and disadvantages. The methods discussed include analyses of well-being of the type used by the FSA for Annex 1 Section H. As the Fujiwara and Campbell paper points out, the problems of how best to value non-market goods have been addressed in various contexts, including environmental, health and transport and my review will draw on some of this literature, particularly in considering alternatives to the well-being analysis in Annex 1 Section H.

Besides these government publications, I will also draw on recent academic literature on behavioural economics and well-being analysis. This will focus on perceived strengths and weaknesses, particularly with regard to the *estimation in practice* of robust well-being estimates.

The review will make a judgment on whether and how far the FSA’s well-being analysis provides robust economic support for the MMR proposals, how it might be improved and whether and, if so, what other analysis might be needed to provide a best-practice economic appraisal.

Para 1.188 in the Overview (p. 39) states that “The CBA shows that, overall, the MMR as a whole is likely to be net beneficial”. The purpose of this review is to provide the Financial Services Consumer Panel with an informed judgment as to whether and how far this statement is justifiable, at least as regards the well-being analysis in Annex 1 Section H.

1.3 What this peer review does not cover

The review will not cover the merits or demerits of the FSA decision.

The first reason for this is that I have only been asked to consider one part of the Impact Assessment – the impacts on retail consumers in the housing market and only one part of that assessment (i.e. explicitly excluding the macro-economic appraisal). The second – and more important – reason is that the role of Impact Assessment’s (IAs) and other ex ante appraisals is to *inform* decisions not to *determine* them.

It is worth saying a little more on the second reason. The role of ex ante cost-benefit appraisals is sometimes misunderstood. Its purpose is to provide decision makers with the best available information and guidance on the economic aspects of the problem which they have to address. The key point is that it should clarify the issues and provide the best estimates (quantitative and qualitative) of the alternative options (including a ‘do-nothing’ option) with, where possible, estimates of the NPV (net present value) of the options considered.

Decision makers have additional considerations that they need to take on board so that it is by no means always the case that the highest NPV option is the chosen one – nor indeed that the highest NPV option is ‘the best choice’, given wider economic, social, legal and other concerns. What is asked of decision makers is that they pay serious attention to the appraisal results and provide a good justification as to why if they choose a different option rather than the highest NPV option(s)¹.

In the context of the MMR, the FSA’s decision clearly involved judgments on financial market stability as well as wider legal and social concerns that go much wider than the impact on the retail housing market and consumers operating in it. However, Annex 1 Section H and the formal appraisal of the options only addressed issues arising in the retail housing market.

1.4 The form of the peer review

The peer review that follows is in two parts.

The first part reviews the FSA’s well-being analysis in Annex 1 Section H, its strengths and weaknesses and the robustness of its conclusions. Hence, the first part of the review covers the first three items in the list in Section 1.2 above.

The second part covers alternative methods that could be used and makes recommendations as to how the overall CBA on consumer welfare in the retail housing market might be improved.

¹ A recent example of good practice in this area is Ofwat’s recent Future Price Limits Consultation Document. The consultation, which considered the form of price controls for the regulated water companies 2015-20, decided in favour of an option which had a middling-NPV rather than the highest-NPV options. The reasons for this choice were clearly set out in the Consultation Document. The IA had, however, clearly set out the economic considerations and charted a potential way forward in the longer term as well as providing the Ofwat Board with thorough guidance on the economic implications of the various options, including risks and uncertainties.

1.5 Summary Conclusions

The answers to the main questions posed for this review are as follows:

- (i) The methodology used by the FSA to capture the potential welfare benefits of those affected by the MMR affordability proposals (well-being analysis based on survey data) is not robust. It should have been supplemented by an alternative form of analysis, as recommended in the Treasury Green Book.
- (ii) The FSA's quantification of welfare costs and benefits is far from robust. The absence of any sensitivity analysis in the well-being appraisal is a major and crucial weakness. This is probably the single most important flaw in the appraisal.
- (iii) Applying 'classical' welfare analysis to the appraisal of the MMR proposals would have been extremely difficult and unlikely to produce robust results. Stated preference modeling, which retains much of the classic utility-based approach, would probably be helpful but perhaps not as the sole method used.
- (iv) The absence of any explicit consideration of house sellers is a weakness but it seems unlikely to make any major difference to the estimates. However, it is surprising that there was no mention of people exiting the housing market and no explicit estimation of the impact of the proposals on rental prices and their consequences.
- (v) The quality of reporting of crucial parts of the appraisal was very poor. It is bad practice not to report equation estimates and diagnostics, as in the well-being equation and elsewhere. The valuations of well-being losses were also reported very inadequately with no clear statement of monetary conversion factors.
- (vi) Annex 1, A4 and Section H failed to follow some of the main and long-established requirements of sound appraisal, e.g. as set out in the Treasury Green Book and consistently advocated by the NAO and by the FSA itself in its 2006 Guidelines. There were no specific (or quantified) objectives; no options besides the recommended option were mentioned or appraised in any way; and, as stated above, no sensitivity analysis was carried out either of the Section H appraisal components or of the final results. This is poor appraisal practice.

The overall conclusion of this review is that the analysis in Section 4 of Annex 1 Section H is seriously flawed. It does not provide a robust and well-founded set of estimates of the impact of the proposed MMR responsible lending requirements on the retail housing market or on its participants.

2 The FSA's Option Appraisal and the Well-being Analysis

This section reviews the analysis of the impact on consumer well-being reported in Annex 1 Section H.

2.1 The Appraisal Methodology

The analysis in Annex 1 Section H is seriously flawed not least because of significant weaknesses in the underlying appraisal framework. These include:

- (i) the absence of any specific *objectives* of the proposed changes;
- (ii) the absence of any reference or discussion of *alternative options* i.e. alternative variants of the proposed lending requirements and/or alternatives to the proposed requirements; and
- (iii) the absence of any *sensitivity analysis*.

The absence of any sensitivity analysis is a particularly serious flaw.

Annex 1 recognises in several places that the appraisal involves major uncertainties. However, there is no analysis of potential margins of error. The Executive Summary of Annex 1 describes the results as "... central estimates that depend on the key assumptions explained in Chapter A4 and *they [the estimates] must be viewed as sitting within broad ranges*"². (My emphasis.) Hence, it is both very surprising and a major weakness that there is no sensitivity analysis of the results of the well-being analysis. This is discussed in some detail below.

The Green Book and the supplementary Impact Assessment Guidance specify the following steps for any ex ante appraisal. These are:

A *Specify a Rationale or Purpose for the proposed policy intervention.*

This is clearly done in this appraisal e.g. in Paras 1 and 4 on page A1:2 of the Annex 1 Executive Summary ("to strengthen responsible lending", "preventing consumers taking mortgages that are clearly unaffordable")

B *Setting Outcomes, Outputs and Targets.*

Given a rationale/purpose, appraisals should specify outcomes, outputs and targets (if appropriate), preferably in quantitative terms. However, neither Annex 1 nor Section H present any defined outcomes, outputs or targets for this appraisal, either in general or in terms of specific improvement objectives.

² Annex 1, Para 26, A1:6

The FSA's recommended policy option is justified in terms of increasing well-being but there is no defined objective for how much measured well-being should be increased in qualitative terms, let alone quantitative terms.

Given the use of a quantitative analysis of well-being impacts derived from an estimated well-being equation, it should readily be possible to estimate the net benefits from improving gross well-being by defined amounts - say 25/50/75%. This would be a useful sensitivity analysis and would also help in defining options as well as in defining outputs and targets.

C Listing Options

Good appraisal practice is that the chosen option should be considered along with a number of other possible options. The Annex 1 appraisal does not do this. The recommended option is only compared against a 'do-nothing' option. No other options are mentioned let alone formally appraised anywhere in the Consultation Document. This is counter not only to Green Book advice but also to the FSA's own 2006 guidance³.

It would be absurd to suggest that all options should be fully appraised. The Green Book suggests appraising a short-list of options, implying a medium-level appraisal with detailed appraisal of the dominant option(s). The FSA document says that all high level CBAs should set out:

- What broadly are the regulatory options for the FSA?
- What are the economic and other costs and benefits of the options, relative to doing nothing?
- What is the plan for further CBA work?

The FSA document is very clear about the need for appraising more than one policy option. The 2006 document appears to refer to an initial high-level CBA sent to FSA decision-makers; but, one would expect at least some discussion in any final published appraisal of alternative options that were considered and why they were rejected.

A best-practice appraisal would include some discussion of (a) other policy options to meet the purpose and objectives of the reform; and (b) various levels of stringency of the proposed new responsible lending requirements e.g. from low-key advisory through to an outright ban on apparently unaffordable lending. Annex 1 and Section H present neither.

Any formal appraisal of other options could be done in outline rather than in detail and would provide a lot more information. Testing the impact of varying the stringency of the new requirements seems particularly important. Indeed, it seems an obvious and crucial element of robustness testing.

³ "A Guide to Market Failure Analysis and High Level Cost Benefit Analysis", FSA, November 2006.

In the absence of any mention of alternative options, the suspicion must arise that the appraisal reported in Annex 1 Section H is an appraisal of a pre-chosen option, an appraisal that has been carried out to satisfy legislative requirements. Even listing other options that had been considered with simple pros and cons and a very high-level CBA as required by the FSA 2006 guidance would alleviate this concern. If, however, this is an appraisal of a pre-chosen option, then one should expect it to be biased in favour of the pre-chosen option.

D Sensitivity Analysis

The absence of any sensitivity analysis has been mentioned above as a crucial weakness on the Annex 1 Section H appraisal.

Attention to sensitivity testing in option appraisal has significantly increased over the last 10 years or so as the use of CBAs has developed. The clearest statement of this is in para 5.69 of the Treasury Green Book (p.32) which is set out below:

“Sensitivity analysis is fundamental to appraisal. It is used to test the vulnerability of options to unavoidable future uncertainties. Spurious accuracy should be avoided, and it is essential to consider how conclusions may alter given the likely range of values that key variables may take. Therefore, the need for sensitivity analysis should always be considered and, in practice, dispensed with only in exceptional cases.”

The well-being analysis in Annex 1 Section H completely fails to recognize this need. This more than anything else means that the reported results cannot be considered as robust.

The sensitivity tests should cover, firstly, the general robustness of the estimates for the well-being model and the valuation of the estimated well-being effects (neither of which are reported); and, secondly, identification of key sensitivities for the results and examining the robustness of the central estimates to variations in those specific variables

Among the sensitivity tests that should be carried out on the well-being estimates are:

- (a) Variations in the specification of the well-being equation discussed in para 185 on page A1:81 and their consequences as reported in Table A4.8;
- (b) Variations in the overall well-being results from simulated changes in the magnitude of the estimated well-being coefficients (e.g. based on variations between equations, from taking coefficient estimates at upper and lower confidence limits, etc);
- (c) Variations in the valuation of the well-being effects discussed in para 194 on page A1:86;

- (d) Identification of key sensitivities in both the well-being equation and the well-being valuation and testing the robustness of the overall CBA results to these key sensitivities;
- (e) Variations in the degree of stringency of the affordability requirements as well as the impact of alternative policy options (as discussed above);
- (f) An estimate of the impact on sellers by simple simulation modeling of the impact of alternative estimates of the impact on the duration of delays in borrowing caused by the MMR. (See below for more on this topic.)

These and other appropriate sensitivity tests should help provide some range of likely impacts as well as just a (not very well-founded) central estimate.

It is worth emphasising that the sensitivity tests can and should be done separately and not jointly. It may be advisable to have a few concluding joint tests but it is both more informative and much less work if they are carried out sequentially.

A proper and thorough sensitivity testing exercise is probably the single most important addition needed for the appraisal reported in Section H.

2.2 The Well-Being Analysis and Estimates

This section discusses various problems with the wellbeing analysis reported in Annex 1, Section H

A *The Quality of the Empirical Estimates in the Well-being Analysis*

It is impossible to judge the quality of the estimates of the well-being analysis as, very surprisingly, the statistical results of the crucial well-being equation discussed in A1:80-82 are not reported either there or anywhere else in the document. No parameter estimates are reported and there are no tests of equation or parameter significance nor any other diagnostic results. This also applies to the debt impairment equations discussed (but not reported) in Annex 1 Section A, page A1:32 and subsequent.

Not reporting the results of estimated equations is very bad practice. The results of the well-being equation and the other Annex 1 econometric equations should be reported in sufficient detail - if necessary in an Appendix, so as to make it possible to judge the reliability and robustness of the estimated equations. In the absence of this information, it is effectively impossible to judge the quality of the reported results, their robustness or the potential for biases.

Table A4.8 reports well-being weights. However, that is a very long way from reporting the results of the underlying equation – for instance, we have no indication of coefficient

estimates, quality of fit, confidence intervals, robustness etc. In addition, the footnotes to the table state that all numbers in the table are rounded.

This issue is not just of good or bad practice. It is a particular concern for judging whether or not the well-being results in Annex 1 Section H are robust and unbiased. This is because it is clear from the empirical literature on well-being analysis that the quality of the results of such equations depends crucially on their specification e.g. on the nature and form of the income variable employed.

Fujiwara and Campbell in their DWP/Treasury survey paper find that it is very difficult to obtain reliable estimates of the effect of income on subjective well-being as represented by BHPS scores⁴. OLS (ordinary least squares) estimates consistently *under-estimate* the impact of income and, hence, *over-estimate* the effect of other independent variables that are correlated with income. But, several of the housing effects listed in Para 185 of Section H are very likely to be correlated with income, possibly substantially. Hence, the estimates of detriment from payment problems, repossessions, etc are likely to be over-estimated – possibly seriously - unless IV (instrumental variable) methods have been used so as to include predicted income rather than actual income in the estimated well-being equation.

In general, as shown in Dolan, Fujiwara and Metcalfe (2011), income effects are often under-estimated in well-being equations based on survey data. Moreover, those biases can be high - Dolan, Fujiwara and Metcalfe find that the well-being effect from OLS equations overstate the IV-estimated effect by 50% or more for a number of variables.⁵

Interestingly, Dolan, Fujiwara and Metcalfe report estimates of the loss of well-being from being perceived to be heavily in debt. This they estimate to be £37,000 per month where income is not instrumented but this falls to £17,000 per month when it is instrumented and other econometric specification changes are made to handle endogeneity issues⁶. This implies a very large margin of error on well-being estimates in general – as well as raising questions about the plausibility of the reported estimates in Section H.

At bottom, understating the impact of income directly leads to over-estimates of the impact on well-being of the housing and other control variables. But, it is impossible to judge whether or not income has been instrumented. If not, from what is reported in Section H, the estimated equation is likely significantly to over-estimate the well-being loss from the housing variables.

There are other problems with the income variable used in the Section H well-being equation:

⁴ See Fujiwara and Campbell (2011), pages 29 and 32.

⁵ See Dolan, Fujiwara and Metcalfe (2011), p. 16.

⁶ See Dolan, Fujiwara and Metcalfe (2011), p.17.

- (i) There is the question of whether income is included in absolute units or as the log of income; and secondly, there is the question as to whether or not some relative measure of income has been included as an explanatory variable. The answer to the first is unknown (Dolan, Fujiwara and Metcalfe recommend using log income rather than absolute income), but from the accompanying text, some relative measure of income does seem to be included in the equation.
- (ii) Self-employed people are a particularly important group for the new requirements. However, accurate estimates of self-employment income are notoriously difficult to obtain. They are typically under-recorded and under-reported. If that were so in this sample, the estimated coefficients on income are again highly likely to be biased in the absence of IV estimates. Even with IV estimation, one should expect problems from errors-in-variables which will at least reduce the robustness of the estimated income coefficients and may introduce biases.

The problems with the reporting of self-employment income suggest estimating the main well-being equation excluding the self-employed and comparing the results with the existing model where they are included. Other possibilities include testing for the significance of self-employment dummies (intercept and slope dummies) in the well-being equation for both employed and self-employed, including self-employment/income interaction dummies.

All of these problems will affect the reliability of the estimates of the housing and other coefficients in the equation underlying Section H as well as the estimated income coefficient per se. How badly they affect the estimates is impossible to judge in the absence of the equation estimates. However, there are clearly major issues both of robustness and bias, with potentially large upward biases in the estimates of the housing distress coefficients.

2.3 The Monetary Valuation of the Estimated Well-being Impacts

The well-being estimates discussed above are converted into monetary values by an imputation process. It seems that well-being units are converted to monetary values on the basis of the monetary value of the distress of unemployment. *This is highly debatable and is an area that crucially needs some sensitivity testing.* It is again an area where an upward bias to the estimates of welfare loss could be potentially serious.

The monetary valuation of excessive mortgage indebtedness is discussed in para 194 on page A1:86. It is a crucial but highly unsatisfactory paragraph. The following points arise:

- (i) The paragraph states that the distress from mortgage arrears is “ ... broadly of similar scale to the distress of being unemployed ...”. No

justification is given for using this variable rather than some other indicator as the proxy for the distress from mortgage arrears.

It is possible - but is unclear - whether the distress estimates are derived from the unemployment coefficients estimated in the (non-reported) well-being equation. (Note that the estimated effect of unemployment from the well-being equation is likely to be overstated unless the income variable has been corrected for endogeneity biases as discussed above.)

- (ii) As far as I can tell, the ascribed monetary value of unemployment derives from different sources. Note that the precise monetary amount is unstated and there are no proper, traceable references to the Credit-Hungry and Bread-Line Experian consumer categories.
- (iii) Unemployment features heavily in well-being analyses of the type used in Section H. Many of them ascribe very high well-being costs to the experience of being unemployed (£23,000 per month for Clark and Oswald in 2002) but Dolan, Fujiwara and Metcalfe find much lower corrected estimates, if still quite high at £12,000 per month. These are (or should be) the estimated well-being costs over-and-above the loss of income from being unemployed.
- (iv) Initial reading of para 194 suggested that the monetary cost ascribed to unemployment and hence to mortgage arrears was derived from the well-being equation. It may be, but the paragraph refers to the "... income that people lose when being unemployed ..." as the "... rough indication of the 'financial equivalent' cost of the distress associated with payments difficulties, arrears and repossession." The associated footnote suggests that this financial equivalent may be an under-estimate since it omits the non-financial costs of unemployment discussed in (iii). The procedure employed is neither clear nor replicable

The main conclusion from para 194 is that readers are given no clear presentation of what the relevant monetary amount used actually is - let alone how exactly it was derived. This is perhaps the single most important estimate in the paper but it is not reported. This is highly unsatisfactory.

In addition,

- (a) It seems as though the *same* value has been given to well-being losses from payment difficulties, arrears and repossession – if so, it is a very surprising decision. However, Table A4.8 does report different well-being *weights* for short and long-term arrears, although it does not report a weight for repossession.

- (b) There is no sensitivity analysis of any kind of variations in the monetary value used and its implications for the net well-being estimation exercise, even though the measurement of this monetary value is explicitly recognized to be highly uncertain. This is extraordinarily bad practice.

Overall, the monetary evaluation of the well-being effects is extremely poor in its exposition and far from convincing. Key statistics, references and estimates are not reported; the prose explanation is ambiguous and imprecise; and the absence of any sensitivity analysis is a very serious omission.

2.4 The Robustness of Well-being Estimates Based on Survey Data

Annex A3 on the Market Failure Analysis mentions a number of “behavioural biases” as a justification for the well-being analysis used. This is (loosely) reiterated in para 158 of Section H.

There are questions about whether or not the effects labeled as “behavioural biases” are market failures but I ignore that question as outside my remit. More importantly – and within my remit – is the question of (a) whether the well-being analysis in Section H is the best or only way of carrying out this type of analysis; and (b) whether, in order to provide robust estimates of net consumer improvement/detriment, it needs to be supplemented by other methods.

My answers to these questions are respectively No and Yes and the text below provides a short justification for the first of these. Discussion of alternative methods is left until later.

The well-being analysis based on survey based subjective estimates used in an ex ante CBA is only one of the methods of analyzing well-being impacts. The approach is described more than once as “well-established” (e.g. Para 180 in Section H). In fact, it is relatively recent and has also been subject to considerable criticism. It is certainly not the only approach that can be taken to behavioural modeling of consumer behaviour, nor to the valuation of non-market goods. In addition, as shown below, the academic consensus is that it remains a relatively controversial method of establishing well-being impacts. Hence, it is a matter of concern that alternative behaviourally based approaches are neither considered nor even mentioned either in in Section H or elsewhere in Annex 1.

There are some authoritative recent surveys of the strengths and weaknesses of the survey-based approach used in Section H, including Kahneman and Krueger (2006), Kahneman (2011) and Stevenson and Wolfers (2008), to the last of which Alan Krueger provides a very helpful discussant comment.

There are various concerns that arise in the underpinnings of this approach e.g. in its use of cardinal as opposed to the standard ordinal welfare comparisons⁷. However, the key question in this context is how far the survey data used genuinely reflect (a) well-being and/or (b) utility and/or (c) “happiness”. Krueger (2008) sensibly declines to answer the “happiness” question. However, like Kahneman, he argues that survey data on perceived satisfaction provide only flawed measures of well-being - but do seem to be systematically related to economist concepts and estimates of utility⁸.

The problem is that the key life satisfaction variable used in the BHPS and elsewhere is “... a global retrospective variable, which in most cases is constructed only when asked and is determined in part by the respondent’s current mood and memory and by the immediate context”⁹. In consequence, it has been shown in empirical experiments that it is subject to various systematic biases, including:

- Mood on the day of interview;
- Framing effects (e.g. preceding questions);
- Focus illusion (e.g. implicit comparisons); and
- Habituation/adaptation effects.

Habituation/adaptation effects are particularly relevant in the context of Section H. The issue is that people who have bad shocks (e.g. paraplegics) have a strong initial reaction in terms of reported quality of mood but this wears off considerably as time passes. In consequence, paraplegics whose limb loss was 12 months or more prior to interview report life satisfaction scores much closer to the average for able-bodied people. This is in strong contrast with able-bodied people when asked how they think a long-term paraplegic person *would* score on quality of mood¹⁰.

This habituation effect could be important in the Section H context as some of the housing and/or arrears states could last some months if not years so that habituation/adaptation effects could well alleviate any initial well-being loss.

The implication of the discussion above is not to discourage the use of well-being analysis as in Section H but:

- (a) to emphasise the need to do it carefully and with full reporting of all relevant results;*
- (b) to emphasise the critical need for sensitivity testing and similar to help establish the robustness of effects – particularly sensitivity testing to address known potential biases; and*

⁷ Cardinal welfare orderings attempt to estimate *how much* individuals prefer A to B and so on, whereas ordinal comparisons require only individuals’ *rankings* of A, B and so on.

⁸ See Krueger in Stevenson and Wolfers (2008), P.100.

⁹ Kahneman and Krueger (2008), p. 6.

¹⁰ See Kahneman (2011), p.405.

- (c) *to point to the need for other, accompanying analysis to complement survey-based well-being analysis if robust answers are to be obtained regarding any measure of well-being or welfare.*

The last point echoes the conclusion in Fujiwara and Campbell (2011), which is reiterated on p.58 of the latest version of the Treasury Green Book. I reproduce the key paragraph of the latter in full below:

“At the moment, subjective well-being measurement remains an evolving methodology and existing valuations are not sufficiently accepted as robust enough for direct use in Social Cost Benefit Analysis. The technique is under development, however, and may soon be developed to the point where it can provide a reliable and accepted complement to the market approaches outlined above. In the meantime, the technique will be important in ensuring that the full range of impacts of proposed policies are considered and may provide added information about the relative value of non-market goods compared with each other, if not yet with market goods.”

This text seems also to be strongly consistent with the academic views discussed above.

2.5 Affordability Requirements and Sellers of Houses

I have specifically been asked to consider this issue, because well-being effects on house sellers are not specifically estimated or considered in Section H.

House purchase flows involve three groups of people: first time buyer entrants, repeat buyers, and exits from home ownership. The vast majority of buyers and sellers are in the second category with the vast majority (but not all) of these transactions involving a mortgage. Exits from home ownership include exits to emigration, to nursing and care home and/or other institutions, to death – and to renting. A moderate proportion of home ownership exits will not involve a mortgage. However, none of these effects are mentioned in Section H nor are any well-being or other impacts calculated for the people affected.

The main effect on sellers is likely to be on repeat buyers, as they appear to dominate the flows. The effect would be to increase delay durations via an adverse effect on the buyer-seller matching function. This is likely to have some negative effect, but probably relatively small.

Para 6 of the Annex 1 Executive Summary states that there is no data available on mortgage delay durations. However, it would certainly be possible to do some illustrative scenario testing or sensitivity analysis. For instance, one obvious test would be to estimate how big a delay would turn the net positive effects of the proposed new requirements into a negative.

One concern that is rather brushed aside is the impact of the proposals on rental prices. Since some people will exit to rented property and one can expect higher mortgage duration delays, one should expect higher demand for rented property and, hence, higher rents. This effect is mentioned in para. 21 of Annex 1, A2, A1:16, but it is suggested there that there will be an offset via more attractive buy-to-let investment and hence that the net effect on rental prices is unclear.

The judgment above seems counter-intuitive and is not supported by any data. (Note that para 22 in Annex 1 appears to claim a benefit to first-time buyers and upgraders from lower growth in house prices.)

The FSA analysis concerns people buying houses with mortgages. However, there are also people who buy (and sell) houses without a mortgage. These transactions may occur with either one or both parties not requiring a mortgage.. The groups involved are most likely to be older house buyers and sellers (e.g. people who have paid off their mortgages or are down-sizing), or foreign buyers and sellers and/or the very rich.

It may be that these groups would be affected by the FSA proposals but they are not included in the analysis. The effect of the omission is likely to be greatest where neither the buyer nor the seller require a mortgage. It would be worth looking at the number and percentage of such transactions and, for completeness, including them in the scenario testing suggested above. However, it seems to me that the numbers involved and the impact on delay durations would have to be sizeable before it is likely to make any significant difference to the reported impacts. Nevertheless, this is clearly an effect that should have been tested in the Section H appraisal.

The omission of estimated effects on sellers of housing seems unlikely to make any major difference to the reported well-being results, although the effect on rental prices could be important for the welfare of consumers in the retail housing market¹¹. Nevertheless, these effects can and should be considered and quantified e.g. by simple simulation analysis based on plausible assumptions about the impacts on the key variables.

2.6 Conclusions on Well-being Analysis in Annex 1 A4 Section H

There are clearly major problems as regards the robustness of the reported well-being analysis. Two main features stand out: firstly, the complete absence of any sensitivity analysis; and, secondly, the failure to report key results e.g. key econometric estimation results and monetary valuation conversion factors. The absence of any clearly specified, quantitative objectives also greatly weakens the analysis as does the absence of any mention or outline appraisal of any alternative options.

¹¹ In overall UK social welfare terms, higher rental prices may – at least in part – be a transfer payment rather than a resource cost, but for consumers of housing services, it is unambiguously a cost that reduces other consumption possibilities.

There are also various potential biases with the Section H well-being analysis. However, given the inadequate reporting of key equations and conversion factors in Section H, it is impossible to judge how far the potential biases are actual biases that significantly affect the reported results. Most of the biases identified are ones that would exaggerate the net benefits, but it is impossible to judge how important they are in the light of the limited econometric information presented. Footnote 49 in Section H suggests an under-estimate of net benefits but the argument used is not persuasive for reasons discussed in 3) above.

3. Alternatives to Well-Being Analysis for the Ex Ante CBA Appraisal of the FSA's Proposed New Responsible Lending Requirements

This section considers alternatives to the well-being analysis used in Annex 1 A4, Section H. I firstly consider 'classic' cost benefit demand and supply models and then I consider other alternatives.

The key conclusions that I draw are:

- (i) Classic cost benefit demand and supply models would not be helpful in this context;*
- (ii) There are other and better possibilities, particularly Stated Preference Analysis and possibly Multi-criteria Analysis;*
- (iii) Given the relative pros, cons and costs of these other possibilities. I would recommend supplementing a revised well-being analysis of the type used in Section H by a Stated Preference Analysis – possibly a conjoint analysis.*

3.1 The Problems with a 'Classic' Welfare Analysis of the MMR Affordability Proposals

Had the FSA proposed a levy on borrowers or some other group of house purchasers, it would have been (relatively) straightforward to carry out a 'classic' ex ante welfare analysis based on estimated changes in consumer and producer surplus as derived from the demand and supply curves for mortgages - plus some consideration of dynamic, distributional and other effects. If any such option has been considered at all during the MMR exercise, according to the FSA 2006 guidelines, one would have expected at least a high-level CBA of this type to have been prepared.

However, the FSA has proposed a policy that yields some (time-varying) quantity restrictions rather than a levy or other price-based mechanism. Moreover, the restrictions are complex rather than simple quantity restrictions and the proposed changes would only affect some part of the mortgage demand and supply curves. Of course, in equilibrium, there is an implicit or explicit price dual to any quantity equilibrium but specifying and estimating the appropriate model is never straightforward and, in this case, it would be extremely difficult indeed.

The relevant theory is the theory of consumer demand under rationing. Following the classic Neary and Roberts (1980) paper, there was a flurry of papers in the 1980s trying to estimate demand for specific goods with rationing. This included Ravaillon (1988) and Ray (1989), both of which considered rationing in housing markets – the Ray paper is particularly relevant as it considers the UK housing market. It is, however, very noticeable that there seem to be few such papers since the 1980s, while the Blundell

(1988) survey of consumer demand is very clear about the major problems with rationing models as well as about the difficulties in estimating them¹².

The Abstract of Ray's (1989) paper on UK housing demand, which is based on 'virtual prices', concludes that the demand parameter estimates and welfare implications are quite sensitive to the introduction of rationed demand and the way that this was incorporated into the model. In consequence, he was unable to obtain robust results.

It is very clear from the academic literature that the appropriate models are at the frontier of economic theory and econometric practice. They may well be useful as background for thinking about policies such as the affordability proposals discussed in the MMR but they are not helpful as the basis for an Impact Assessment (IA). The results obtained are essentially determined by the form and structure of a complex mathematical economic theory model and the estimation methods applied. They are also typically highly demanding on data. That is the opposite of what is required in an IA.

3.2 Alternative Approaches to Estimating the Impact on Households of the MMR Affordability Proposals

There are various other methods that have been applied to estimating welfare impacts of non-marketed goods which may be relevant in the context of valuing the costs and benefits of these proposals. The ones that I consider in this section are (a) Stated Preference models in general with conjoint analysis as a particular example; and (b) Multi-criteria Decision analysis.

A Stated Preference Modeling

Had I been asked in advance as to which method I would recommend using as the basis of the Impact assessment in Section H, I would have advised looking first at Stated Preference Modeling (also known as contingent valuation).

Stated preference (SP) valuation involves the use of surveys in which groups of people are required to make choices between variables on the basis of WTP (willingness to pay) or WTA (willingness to accept). As explained in Chapter 3 of Fujiwara and Campbell (2011), a hypothetical "market" is set out for questionnaire respondents together with methods and/or frequency of payments. The "market" presents alternative descriptions of a good with variations in the level of the good's attributes, including the status quo (or a 'do-nothing' option). The respondents then report their (ranked) choices.

¹² I have also seen references to a 1990 paper by Geoff Meen on estimating the impact of the decline of mortgage rationing in the UK in the 1980s by means of a regime switching model. I have not seen the paper but it has recently been cited in Muellbauer and Williams (2011) who applied LIVES (latent interactive variable equation system) modelling to explore non-price credit supply conditions in Australian mortgage and other credit markets.

This method comes in various guises and has been extensively used for environmental economic issues as well as in transport economics. It seems particularly well-suited to establishing trade-offs between levels of attribute on a non-market good and cost (or benefit). Hence, in the MMR context, via SP, individuals should be able to estimate the WTA of a lower risk of arrears/repossession etc from tighter affordability conditions relative to the costs to them of changes in the likelihood of being granted a mortgage (or the likelihood of a smaller mortgage). In this way, it should – in principle - be possible to provide some kind of a utility-based valuation of how individuals would value the trade-offs with non-price affordability rules. Unlike well-being analysis, its preference basis is ordinal not cardinal as it is based on sequences of choices between options.

SP has been used widely in economic issues and also a variety of commercial and marketing applications. This means that it has been used a lot and for over 30 years, so that the problems with it are relatively well-known. The main issues arise from the survey design and valuation framework used. In fact SP is a *set of techniques* some of which seem to work better than others. In the MMR context, Fujiwara and Campbell point to the difficulties that people have when trying to convert feelings or concepts of value into monetary amounts¹³. There are, though, a number of SP techniques that do not require explicit monetary valuations but rely on rankings and these have been found useful in health economics (e.g. with QALY's¹⁴).

A particular version of SP that might be useful in the MMR context is conjoint analysis. Conjoint analysis focuses directly on trade-offs between attributes of a good rather than WTP/WTA and is much used in marketing studies.¹⁵ An interesting application of conjoint analysis is in the 2011 Demos/PWC “Good Growth” report which uses the method to compare peoples’ preferences between pure economic variables (e.g. income and employment) with other variables (e.g. health, work-life balance, the environment, etc)

However, SP modeling has its own problems and, like survey-based well-being analysis, its use is controversial. For instance, SP studies frequently produce results that are inconsistent with utility-based predictions. Diamond and Hausman (1994) present a highly critical review of SP methods and their results - as have cognitive psychologists (including Kahneman). However, later surveys suggest that there are good and bad SP studies and the method is not in itself inherently flawed¹⁶. A major issue is that WTP and WTA valuations frequently diverge; conjoint analysis, however, does not require explicit monetary valuations.

Fujiwara and Campbell (2011), Chapter 4 bring out many of the problems (including significant financial costs of SP studies) as well as the potential benefits. They conclude

¹³ See Fujiwara and Campbell (2011), p.21.

¹⁴ Quality-of-adjusted-life-years.

¹⁵ For a simple description see http://en.wikipedia.org/wiki/Conjoint_analysis. For a much more formal description, see Louviere, Hensher and Swait, “Stated Choice Methods”, CUP, 2000.

¹⁶ See, for instance, Carson (2001).

that the robustness of SP methods is a problem – but currently less of a problem than with survey based well-being analysis.

SP methods are much closer to ‘classic’ welfare analysis than the well-being analysis of Section H. The main use of SP has been to put monetary valuations on non-market goods but it has also been used to help establish non-monetary valuations e.g. of health states via QALYS. Since, the pros and cons of SP are clearly different in kind from those of the well-being analysis in Section H, my recommendation would be that a small SP study - possibly a conjoint analysis or similar variant - be used to test the robustness of the Section H well-being analysis. House sellers as well as buyers could readily be included in that.

The recommendation above is very much in line with the 2011 addition to the Treasury Green Book which advocates a case-by-case approach and suggests the use of more than one method to check the consistency of results¹⁷.

B Multi-criteria Decision Analysis

The other possible option that may be worth consideration in the MMR context is Multi-criteria analysis. This is relevant where there are important benefits to be included but to which there is no obvious (or sound) way of ascribing a monetary value – for example, the Norman churches that were threatened with demolition in the 1970s Third London Airport option appraisals. It might be argued that the satisfaction/stress benefits from tightening affordability benefits came into this category¹⁸.

Again, Multi-criteria analysis is a technique with clear snags as well as benefits. It is not CBA-based and is relatively weak on the financial and other cost side. It also relies primarily on the advice of various types of expert and can be expensive to carry out. For these reasons, it is not recommended in the MMR context, the FSA having already done a well-being analysis.

3.3 Concluding Comment on Alternative Methods

Estimation of an appropriate Stated Preference model as suggested *to complement the well-being analysis* would clearly improve the robustness of the appraisal as regards the retail housing market. It might also help reveal potential biases with the well-being model. However, it is not possible to judge which of the alternative methods (and sub-methods) discussed above is, per se, more or less biased or in what direction.

March 2012.

¹⁷ The Treasury Green Book leans towards revealed preference (RP) methods which rely on data from purchases and sales in similar markets. However, RP methods do not seem applicable in the MMR case.

¹⁸ The standard guide to multi-criteria decision analysis is a 2009 Manual published by DCLG and downloadable via http://eprints.lse.ac.uk/12761/1/Multi-criteria_Analysis.pdf

References

Blundell, R. (1988), "Consumer Behaviour: Theory and Empirical Evidence – A Survey", *The Economic Journal*, Vol. 98, No 389, pp. 16-65.

Carson, R.T. (2001) "Resources and Environment: Contingent Valuation", Elsevier.

Clark, A.E. and Oswald, A.J. (2002), "A simple statistical method for measuring how life events affect happiness", *International Journal of Epidemiology*, Vol. 31, pp 1139-1144.

DCLG (2009) *Multi-criteria analysis: a manual*. http://eprints.lse.ac.uk/12761/1/Multi-criteria_Analysis.pdf

Demos (2011), *Good Growth – a Demos and PWC report on economic well-being*".

Diamond, P.A. and Hausman, J.A. (1994), "Contingent Valuation: Is Some Number better than No Number", *Journal of Economic Perspectives*, Vol. 8 Issue 4 pp. 45-64.

Dolan, P., Fujiwara, D. and Metcalfe, R. (2011), "A Step towards Valuing Utility the Marginal and Cardinal Way", *CEP Discussion Paper No. 1062*.

FSA (2006), "A Guide to Market Failure Analysis and High Level Cost Benefit Analysis".

Fujiwara, D. and Campbell, R. (2011), "Valuation Techniques for Social Cost-Benefit Analysis: A Discussion of the Current Issues", HM Treasury/DWP.

HM Treasury (2003), "The Green Book: Appraisal and Evaluation in Central Government".

Kahneman, D. and Krueger, A.B. (2006), "Developments in the Measurement of Subjective Well-Being", *Journal of Economic Literature*, Vol 20, No 1, pp 3-24.

Kahneman, D. (2011), *Thinking Fast and Slow*, Allen Lane.

Krueger, A.B. (2008) "Discussant Comment to Stevenson and Wolfers", *Brookings Papers on Economic Activity*, Spring 2008.

Louvain, J.J., Hensher, D.A. and Swait, J.D. (2000), *Stated Choice Methods: Analysis and Application*, Cambridge University Press.

Muellbauer, J. and Williams, D.M. (2011), "Credit Conditions and the Real Economy: The Elephant in the Room", *CEPR Discussion Paper DP8386*.

Neary, J.P. and Roberts, K.W.S. (1980), "The Theory of Household Behaviour under Rationing", *European Economic Review*, Vol.13, pp 25-42.

Ofwat (2011), "Future price limits - a consultation on the framework"

Ravallion, M. (1989), "The Welfare cost of Housing Standards", *Journal of Urban Economics*, Vol 26, No 2, pp. 197-211.

Ray, R. (1989), "The behavioural and welfare implications of housing demand under rationing: The UK Experience", *Journal of Population Economics*, Vol 2, No 3, pp 211-24.

Stevenson, B. and Wolfers, J. (2008), "Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox", *Brookings Papers on Economic Activity*, Spring 2008.