# PILOT STUDY OF PHYSICAL HOUSE CONDITION AND REHABILITATION NEED

## IMPLICATIONS OF THE UNITED STATES/CANADA PILOT STUDIES AND THE WASHINGTON WORKSHOP FOR FURTHER RESEARCH IN CANADA

9 June 1981

A study conducted for the
Canada Mortgage and Housing Corporation
by
Ekos Research Associates Inc.
Suite 304, 171 Nepean Street
Ottawa, Ontario
K2P 0B4
(613) 235-7215

Ekos Research Associates Inc.

171 Nepean Street
Suite 304
Ottawa, Ontario
K2P 0B4
(613) 235-7215

July 3, 1981

Mr. Philip Brown
Manager
Housing Requirements and Conditions
Market Forecast and Analysis Division
Canada Mortgage and Housing Corporation
National Office
Montreal Road
Ottawa, Ontario
K1A 0P7

Dear Mr. Brown:

Ekos Research Associates Inc. is pleased to submit three copies of the final report entitled: "Measuring Physical House Condition and Rehabilitation Need: Implications of the U.S. - Canada Pilot Studies and the Washington Workshop for Further Research in Canada".

Sincerely,

EKOS RESEARCH ASSOCIATES INC.

Frank L. Graves

Frank L. Graves
Principal Investigator

Enclosures (3)

# TABLE OF CONTENTS

Page

## 1.0 INTRODUCTION: AIMS AND STRATEGIES

The research conducted to date does not provide final answers to all of the research questions underlying the project. This is due to two reasons: (1) the analyses of the existing data are by no means exhaustive; (2) the data are derived from a specific population, therefore, the findings are not generalisable to other populations. The purpose of this report is to provide a set of practical suggestions for future research on rehabilitation in the Canadian context. These recommendations are based upon the experience of the Canadian and U.S. pilot studies. We also incorporate some of the suggestions and criticisms of these two studies which were generated by a conference on these studies held in Washington, D.C. in May of 1981. The list of participants to this conference and the agenda are included as Appendix II to this report.

The existing survey data base generated from the Ottawa pilot study is a rich source of housing information containing many variables which have yet to be analysed in a detailed fashion. In this discussion we will sketch some of the more obvious and pressing analyses which should be undertaken using this data base. We also suggest additional strategies which require new data collection or an integration of other existing data banks. In formulating these strategies, we

have considered the results of our initial analyses, the findings of the U.S. pilot study and the proceedings of the Washington Workshop.

The balance of this paper is devoted to eleven interrelated research topics and issues. These topics are loosely ordered in terms of practical priorities. The topics include: (i) Major Findings, (ii) Further Analyses of the Census Repair Need Question, (iii) Further Work with the Seven-Point Scale and Measuring Sub-standardness, (iv) Aggregate Level Modelling, (v) Reliability Issues, (vi) Towards Models of Rehabilitation Need and Potential, (vii) Sampling, (viii) Population Forecasts, (ix) Additional Thematic Indices and Predictive Models, (x) Delivering a House Condition Assessment Methodology to Local Municipalities, (xi) Considerations for Measuring High Rise Housing Stock. Before proceeding to a discussion of these topics, we will briefly review the major findings of the Ottawa Pilot Study and present a cursory comparative discussion of the American Pilot Study.

## 2.0 MAJOR FINDINGS

It is possible to encapsulate the major findings of the study. Following is a brief summary:

### 2.1 Measuring and Defining Physical House Condition and Rehabilitation Need

Recognizing that there is no single satisfactory measure of physical house condition, the study developed a multidimensional approach. This approach was based on selecting condition indicators which were considered to be germane to the concepts of rehabilitation need and physical house condition. This approach is based on two assumptions. First, that physical house condition is not a unidimensional concept but rather involves several independent dimensions. Secondly, rehabilitation need is an even more complex concept which requires information concerning the household as well as physical dwelling. For purposes of this study, physical house condition was divided (a priori) into the exterior, interior and mechanical systems. The study also collected information on the socio economic characteristics of the household and background dwelling characteristics. A more detailed description of this conceptual organisation of the data base, and its relationship to the survey instruments, is provided in Appendix I of the Major Report.

Our analyses of the data suggest that our a priori assumption regarding the multidimensionality of the data was correct. Although there were some variations in the form and content of the dimensions and factors generated inductively by

factor analysis, there were also non-ambiguous similarities. In a sense, we can define rehabilitation need as the inverse of physical house condition. That is, the better a dwelling's physical condition, the lower its rehabilitation need. When we introduce the notions of adequacy and rehabilitation need, we realise that we must expand the purely material definition of rehabilitation need to include the social, economic and psychological characteristics of the household as well. The discussion of future research will outline an agenda for further study of this problem.

## 2.2 Comparative Analysis of Data Sources and the Value of Non-Expert Data

House condition data were provided by three distinct data sources (i.e. the person who provided the information). These were: (i) building experts; (ii) trained interviewers; and (iii) dwelling residents. Since the ratings and observations all apply to the same objects (i.e. matched observations), we were able to assess the relative quality of the different data sources in terms of reliability, validity and economy. Validity of interviewer and occupant data were estimated in the following fashion. The professional building expert's assessment of a property's physical condition was taken as the "criterion" or "truth". The validity of occupant and interviewer ratings was assessed by calculating the degree of concordance or agreement with the criterion expert rating. As the analysis shows, we found that non-expert interviewers and (to a lesser degree) occupants can provide reliable

and valid physical condition data with the proper training and proper survey instruments. Using both occupant and interviewer data, we achieve better predictions of expert ratings than we do using either source exclusively. However, the trained interviewer data is superior to the occupant data by a significant margin.

## 2.3 Production of Continuous Scale Measures

The study employed a seven point rating scale, which approximated continuous level measurement. The quasi-continuous seven point scale was superior to the traditional categorical rating of "pass - fail" on the grounds that the continuous rating scale enhanced the discriminatory powers of the rater, and hence improved the sensitivity and validity of the data. The results of the analysis - and in particular the distributional characteristics of the scale variables - indicate that the assumption of continuous level of measurement is justified. This superior level of measurement yields more information than discrete data and permits more powerful predictions. Subsequent analyses and discussions suggest that the seven point scale be modified to accomodate a pass/fail threshold level. However, this cannot be rigidly tied to program standards in a manner similar to the American approaches until standards for programs such as RRAP are more precisely defined.

## 2.4 Development of Indices of House Condition and Rehabilitation Need

The use of the interval rating scale permitted the development of indices of house condition that were continuous. These indices were constructed to range from 0 - 100 with 0

meaning perfect condition and 100 meaning the worst possible condition. This allows flexibility in the selection of different thresholds, or "cutting points", for multiple definitions of substandardness. For example, the same house condition index could be used with different threshold levels to define substandardness, in terms of: (a) minimum property by-laws of a particular municipality; or (b) a program-based operationalisation of substandardness (e.g. RRAP, CHIP, etc.). We will show how scale values can be "grounded" in terms of average repair costs or relative position (i.e. ranking) in a designated population. This facilitates the more precise classification of the housing stock into categories of rehabilitation need.

## 2.5 Evaluating the 1981 Census Question of "State of Repair"

Given the absence of Canada-wide data on rehabilitation need, the study evaluated the extent to which the 1981 Census question on "state of repair" would provide reliable and valid information on repair needs and, relatedly, on repair costs by the three question response categories: (a) dwelling requires maintenance only; (b) requires minor repairs; and (c) requires major repairs. The analysis demonstrates that occupants can provide very good answers in terms of whether or not a dwelling needed repairs. However, the distinction between minor and major repairs is much less useful although there is evidence that if certain systematic sources of bias are controlled for, certain types of occupants can make these distinctions. Undoubtedly, the phrasing of the Census Repair Need Question contributed to some of this confusion and we suggest that the distinction between major and

minor repairs could be clarified with more appropriate and precise examples.

## 2.6 Summary

The study's Pretest results can be taken as evidence that a practical methodology for collecting high quality, yet cost effective information on physical house condition is not only possible but available from the present research. However, given that this study is only based on a single sample, it must be emphasized that the Ottawa Pretest is only an important step toward the implementation of this methodology in other settings. It is also important to note the present analyses are only a basic and preliminary attempt to answer the study's major question. More detailed and sophisticated analyses must be conducted. Additional research is needed to resolve five or six issues regarding the general approach and survey instruments. For example, it is still necessary to test a French language version, to test the applicability to rural areas and cities of different size groups, and, to extend the methodology to cover additional dwelling types such as high rises.

These concerns are echoed by Statistics Canada:

> "(Even with) . . . a positive result . . . one would be rather ill-advised to proceed <u>without further research</u> to employ the method developed on the basis of this experiment in a large-scale survey intended to assess the full spectrum of housing stock."

Before considering a set of practical research problems which should improve the utility of the study, we will briefly review (in a comparative light) the approach and findings of the U.S. Pilot Study.

## 2.7 Comparative Discussion of the Approach and Findings of the U.S. Pilot Study

The American pilot study was conducted concurrently with the Canadian pilot by Abt Associates in Boston, Massachusetts. Although the studies shared essentially identical research problems, differences in the populations studied, the methodologies employed and the national policy contexts for the two studies produced significant differences between the Canadian and U.S. pilots in terms of approach and findings. The following comparative discussion details the major differences in approach and findings between the two studies.

Both studies were designed to improve and refine data bases on housing conditions at the national and regional levels for the purposes of policy and program development, program evaluation and planning at the federal level. Both studies shared the more immediate objectives of producing physical inspection instruments and testing methods for estimating rehabilitation costs. The study samples differed in the following ways. The Canadian study's site was the inner City of Ottawa whereas the American study used census tracts within the City of Boston. The Canadian sample size was about 500 properties versus 300 dwelling units for the American sample. The type of sample used by the two studies also differed. The Canadian study employed a stratified random sample of low-rise properties which was stratified by house type, tenure and dwelling condition. The American pilot also sampled low rise dwelling units. However, it used a two stage cluster sample which was stratified by census tract. Neither study will

support detailed inferences to the municipal level although the Ottawa pilot is better equipped to offer population inferences.

In terms of survey instruments, both studies used household questionnaires and technical inspection instruments. The Canadian pilot also tested a "skim" technical instrument which was administered by trained interviewers. The American instrument was organised on a detailed room by room basis whereas the Canadian instrument relied more heavily on general ratings (e.g. of interior walls and ceiling surfaces). The American instruments consequently produced significantly more variables than the 500 variables derived from the Canadian instruments.

The American instruments were organised around Section 8 standards which are very rigidly and precisely defined. This precise and rigid codification of the programs was reflected in the U.S. survey instruments and the summary measures of rehabilitation need derived from their data base. The Canadian instruments were structured to capture the content of RRAP. RRAP defines eligible areas of assistance clearly but standards are left relatively vague (vis-à-vis Section 8). Consequently, the Canadian instruments and summary measures were much more general and flexible than the American instruments. Related to this point, the Canadian study attempted to create general continuous measures of physical condition and rehabilitation need based upon individual, quasi-continuous 7-point rating scales. The

American study utilised nominal and ordinal individual ratings which produced categorical summary measures of rehabilitation need.

Both studies estimated repair costs using a cost matrix. The Ottawa study had the CMHC inspectors calculate repair costs in the field, on the inspection form itself, using an established cost matrix provided by the City of Ottawa. The American study derived repair cost estimates from ratings and estimates of material requirements. These derivations were not done in the field but derived from a computer algorithm based on a commercially produced cost estimating guide.

Just as the methodologies, policy contexts and study populations differed, so did the major findings. The repair costs estimates in the American study were quite low compared to the Canadian study. The average repair costs <u>for those units which failed</u> Section 8 were approximately $1000. whereas the overall average for the entire Ottawa sample was greater than $2000. This is despite the fact that it was agreed that the Ottawa stock was probably in much better condition than the Boston stock. The American approach to defining rehabilitation need was based on Section 8 standards and this approach identified 62% to 83% of the population as requiring rehabilitation. The Canadian study's estimates were based on relative condition, repair costs (as they compared to actual average RRAP jobs) and building experts' opinions. These methods estimated a much smaller proportion of properties

in need of rehabilitation (i.e. about 10% to 25%). The relationship between repair costs and the American summary measures of rehabilitation need was less linear and less precise than the relationship in the Canadian study. This is understandable when we consider that approximately 1/3 of the units failing Section 8 (low) standards in the U.S. study, needed less than $200. worth of repairs.

The American study found "little correlation among various classes of deficiencies" and high classification errors associated with relying on an exterior only inspection. In contrast, the Canadian study found strong intercorrelations amongst individual types of deficiencies and a significant and a usable relationship between interior and exterior. The fact that intercorrelations were not discovered by the American study is probably largely due to the lower level of measurement. In fact, the statistical basis for this conclusion was an improper Pearson correlation matrix for binary data. The lack of interior/exterior congruence is also due in part to the weaker level of measurement, although important differences in the Section 8 standards also contribute to this finding.

One of the most important findings of the American study, which was not tested in the Canadian context, was that aggregate level predictive models demonstrated much higher levels of explanation (i.e. $R^2$) than the corresponding unit level models. Models which made predictions to the level of Census tracts produced better fit and lower errors than similar models which

predicted to the level of individual dwelling units. This result is to be expected since there is far less variance (error) in areal level models. There are some significant methodological problems which accompany these areal level models which we will discuss later.

## 3.0 FURTHER ANALYSES OF THE 1981 CENSUS REPAIR NEED QUESTION

The Census Repair Need Question is destined to become an important tool for policy formation, program design and program administration. Due to the public nature of the data, it also will likely become a politically potent tool for municipalities and provinces. For these reasons, it is critical that we have a judicious understanding of the responses to this question.

Further analysis of the 1981 Census Repair Need Question is necessary. Although the crosstabular analysis reported in the Major Report provides good initial estimates of the utility of this question, further practical analyses will enhance the future value of this important data source. Our discussion is divided into three sections: (i) log-linear and logit modelling, (ii) descriptive analysis of deviant cases and marginal distributions, and (iii) multi-site testing.

### 3.1 Log-linear Models

If we consider the experts' responses to the three categories of this question as a criterion, then it will be useful to see how accurately we can predict these categories on the basis of occupant responses (and to a lesser degree interviewer responses). On the basis of the initial research findings, it is clear that a fairly accurate predictive model is possible. Since the response categories to this question are three ordered (ordinal) categories, parametric methods such as multiple regression analysis are inappropriate. However, cruder non-parametric techniques do not provide sufficiently precise answers.

There is an analytic technique available which is ideally suited for assessing how well we can predict the experts' categorical ratings from our knowledge of other categorical information, such as the occupants' response to the Census Repair Need Question, the tenure status of the occupant, the age of the building, etc. This technique has been developed recently by Leo Goodman and others and is known as log-linear modelling. The idea of using log-linear models received unanimous support from those participants at the Washington Conference who were familiar with the technique.

With log-linear modelling, we predict the likelihood of falling into a given category of a contingency table as a function of membership in another cell (or combination of cells) of the table. The technique is similar to regression although instead of predicting continous scores on the basis of our knowledge of other scores, we predict cell frequencies from other cell frequencies. Although the mathematics underlying the technique are formidable, the results can be expressed in clear non-technical language which is meaningful to a policy maker. We include a brief annotated bibliography of log-linear methods for the interested reader as an appendix.

Having conducted a log-linear analysis, we can make statements like, "the odds that a dwelling truly needs 'major repairs' (i.e. the expert would rate it as such) are 10 to 1

if we know that the occupant was a lower income renter who rates his dwelling as needing minor repairs".* This approach provides a sophisticated method for identifying systematic sources of bias and consequently a method for correcting bias in these respondents. For instance, if we know that lower income owner occupants living in single detached dwellings are three times as likely to rate properties as needing no repairs as a building expert, then we can correct aggregate municipal estimates of the proportion of low-income owner occupied single detached properties falling in the no repairs category by using a weight of 1/3. In the case where the dependent variable is dichotomous (or dichotomised) then we can use a logit model.

If we examine occupant data only, we can develop predictive log-linear models which estimate the odds that respondents with certain types of characteristics (e.g. renter, older, francophone) will rate their dwellings as needing no, minor or major repairs. Since we have the Census Repair Need data for other sites (e.g. the Atlantic Region), we can even begin to assess the impact of place and settlement type on responses to the Census question. Although this will <u>not</u> determine the relative validity of the responses given in, for example, Saint John versus Ottawa, it will tell us if the same type of dwellings typically provide the same types of responses. Further validity testing requires linked expert and occupant responses to this question.

---

* Hypothetical example.

A more immediate analysis which may yield useful results would be to conduct a descriptive analysis of extreme deviant cases. This analysis would consist of extracting those cases from the data file where the occupant made a "serious error" in responding to the repair need question. A serious error refers to the situation where an occupant and expert disagreed by two categories (e.g. the expert said major repairs were needed and the occupant said no repairs were needed). By profiling other characteristics of those occupants who make serious errors in answering the Census Repair Need Question, we may be able to make further corrections and more precisely diagnose problems of interpreting results. For example, we might decide to eliminate those responses from our estimates where we know that there is a prohibitively high probabiliy of encountering random errors. This approach was suggested at the Washington meetings.

## 3.2 Further Descriptive Analyses of Marginals and Deviant Cases

Another suggestion would be to further examine the relationship between the first order marginals of the cross-tabulation of expert and occupant responses to the Census Repair Need Question. Others have noted the relative percentages of respondents rating their dwellings as needing no, minor or major repairs are quite close to the corresponding percentages from the experts. In other words, the simple frequency distributions of experts and occupants show roughly

the same percentage of responses in the three categories. This approach may be useful only if we are not making inferences at an individual level. The crosstabular analyses showed that there were problems with this approach. Despite the fact that 21% of occupants and 23% of experts rate the Ottawa sample properties as needing major repairs, these are not the same dwellings. At an aggregate level, this may not be important. Nonetheless, we view this approach with some suspicion and feel that further testing is necessary (in particular multi-market testing and tests of statistical significance would be in order).

Related to this approach is the issue of level of aggregation for estimates generated from the 1981 Census question. We will deal with the issue of level of aggregation separately later, but we do note here that, in general, the more highly aggregated our inferences and data are, the fewer errors we will encounter. Hence, if we wish to infer to a city, regional or national level, we may be able to relax certain requirements. Further testing of aggregate models is needed.

### 3.3 Multi-Site Testing

A complete understanding of the meaning of the 1981 Census Repair Need Question awaits multi-site testing. This point was made by several participants at the Washington meeting. Unless we have linked expert and occupant responses in a variety of settings, we will not know what occupant

responses truly mean. If policy-making and program delivery are to rely heavily upon the Census Repair Need Question, then it is imperative that we know what each of the three categories mean for various combinations of respondent characteristics. For example, how does a reply of "major repairs needed" in Moncton compare with the same response in Vancouver? In other words, does the environment or settlement type alter the meaning. Similarly, are there systematic differences in the way respondents answer the question as a function of the social, economic and ethno-linguistic characteristics of the respondent and his or her household? Do francophones and anglophones answer the question the same way? We also will need to know how the environmental and background characteristics of the respondent interact. For example, are response patterns to the Census question typically different for large households in rural areas than large households in urban areas? Only after this type of analysis will we be able to usefully interpret this question.

## 4.0 FURTHER WORK WITH THE SEVEN-POINT SCALE AND MEASURING SUBSTANDARDNESS

There are three topics to be discussed in this section: (i) methods for incorporating the advantages of the categorical or pass/fail approach while retaining the quasi-continuous seven-point scale; (ii) using this approach to identify the RRAP eligible housing stock; (iii) equidistance between points on the scale and inter-market variability.

### 4.1 Methods for Incorporating Pass/Fail into the Seven-Point Scale

Much of this work can be started without any additional data collection. Most programs are conceptualised in terms of *a priori* pass/fail standards. Although a program such as RRAP does not have the same rigidly defined standards as the American Section 8 Program, many officers involved with the program tend to think in terms of pass/fail. Despite this, we would continue to argue that the continuous measures developed in the Ottawa pilot are superior to a pass/fail approach. We believe it is important to maintain the continuous approach since this provides more precise and powerful measurement. The fact that the Canadian pilot study discovered strong and significant interrelationships amongst the individual condition ratings whereas the U.S. study found none, was largely due to the continuous level of measurement. This also provided significant advantages in terms of inductive

data reduction strategies (i.e. factor analysis) and more powerful (individual level) regression models (e.g. repair costs). By modifying the quasi-continuous seven point scale, we are in position to retain the advantages of both approaches.

In order to reconcile the continuous and categorical approaches, we would suggest that in addition to defining the polar positions of the scale, a pass/fail threshold should be designated on the scale. This point would be established after considering the individual purposes of a given survey. This would retain the capability of designating alternate definitions and standards.

## 4.2 Measuring RRAP

It is possible now to go back to the Ottawa pilot study data base and define a score of 5 or more as a failing score. Given the manner in which the scale was constructed and explained (with 4 being an average state of disrepair), 5 or more probably makes sense as a failing region. We would now be able to routinely implement an *a priori* designation approach such as the Rostum approach*. The RRAP eligible population would then be estimated using this approach. Future surveys might consider building a passing and failing region directly into the 7-point scales. We suggest that even the Abt instrument could be routinely restructured to retain its ability to measure pass/fail program standards, yet provide continuous measures of physical house condition.

---

* cf. Section 3 of the Major Report.

4.3 The Issues of Scale Equidistance and Inter Market Variations

Several other points related to the question of the 7-point scale in particular, and level of measurement in general require clarification. First, the question was raised in Washington as to whether or not the difference between 1 and 4 was the same as the difference between 4 and 7. The answer is yes and this was stressed both in the training session and the nature of the 7-point scale itself (e.g. the visual prompt emphasised the equidistance of the intervals).

Another important question raised was the issue of whether the same scale value is translatable from one market to another. In other words, does a "4" mean the same in Vancouver as it does in Hull? This is really an empirical problem which awaits multi-site testing. A potential problem is the interior regions of the scale. Although the meaning of a 1 or 7 are clear and absolute, the meaning of 2 or 6 will be relative to a particular market since the "average" state of repair will vary from one site to another. One possible strategy for alleviating this potential problem would be to define some middle point on the scale in absolute terms. For example, if the scale is measuring the surface condition of interior walls, we could define walls with missing material or holes larger than a sheet of paper (8½ X 11) as a "5" or worse. Another suggestion would be to include a RRAP failing point on the scale. Grounding the two poles and some midpoint of the scale would facilitate the process of interpolating an undefined score. This approach also articulates well with the view that we should incorporate a pass/fail threshold within a 7-point scale.

## 5.0 AGGREGATE LEVEL MODELLING

### 5.1 Implications of U.S. Neighborhood Level Modelling

An important direction for future research in the Canadian context would be to consider neighborhood or other aggregate level modelling. The broad implication of the American results is that aggregate level modelling conducted at the census tract level provides much higher levels of explanation* than comparable unit level modelling. There is no reason to presume that the same types of improvements in the fit of predictive models would not occur in the Canadian context. This view was supported by the American participants at the Washington meetings.

Neighborhood level variables contributed significantly to reduced prediction errors in the American pilot and we suggest that the same result would occur in the Canadian context. Furthermore, the neighborhood is a natural unit for delivering certain types of rehabilitation programs and must be considered in any complete model of rehabilitation potential. Unit level models are the most conservative and hence the most secure initial platform for basing future models. Having demonstrated that the models used in the Ottawa and Boston pilot studies work at a unit level, we should now proceed with areal level models.

---

* The American study found that $R^2$ typically more than doubled at the neighborhood level.

Using areal level models, such as census tract models, we will be in a position to incorporate ecological census tract data as contextual variables. This should further enhance the predictive power of models estimating: the percentage of units in need of rehabilitation, the percentage of units eligible for RRAP, the absolute number of units in need of rehabilitation and eligible for RRAP, the costs of rehabilitating a given neighborhood or introducing a certain program such as RRAP, etc.

Aggregate models at higher levels of inference such as the city, province, region or nation should involve even less errors. We may find after completing the research agenda suggested for the Census Repair Need Question that (at these levels of aggregation) sufficiently precise estimates can be derived from census data or other census-like data derived from intercensal surveys, such as the Survey of Household Facilities and Equipment (HFE). In other words, the proportionate reduction in error achieved using more detailed house condition survey data is not justified given the additional expense of acquiring such data.

## 5.2 Caveats

The problems associated with aggregate level models include factors such as aggregation (ecological) bias, restricted degrees of freedom, potential misinterpretation of results and greater problems in the areas of multicollinearity and heteroskedasticity. What does it mean if we find that "average

attitudes" are a good predictor of rehabilitation needs. Similarly, there is always the danger of the consumer of these statistics shifting levels. The classic example of ecological fallacy from Robinson (1950) is a good illustration. At an aggregated level, conditions between illiteracy and race were extremely strong at a group level but weak at an individual level. The point is that group correlations cannot be used as substitutes for individual correlations, since they are rarely the same. Nonetheless, non-technical users will often leap levels in using such results. One reason that we can reduce the errors in prediction so substantially is that the absolute amount of error (variance) has been substantially reduced. Most of these problems can be handled in a well designed research plan.

## 5.3 Next Steps

Since aggregate models possess higher explanatory power, can be linked to census tract data, and are better suited for macro-level provincial, regional and national forecasting, it would seem sensible to proceed immediately with aggregate level analysis. Despite the fact that the Ottawa pilot data base does have census tract as a variable, it is not possible to aggregate by census tract to create a new neighborhood level data base. The problem is that there are tremendous differences in the number of original cases contained in each census tract and many of the survey census tracts had only a few properties sampled from them.

Furthermore, only about 20 cases (census tracts) would be available for analysis and this would make inferences questionable. Since regression analysis assumes an unrestricted range for the dependent variable and that the sample means are more or less distributed as a normal, a sample of over a hundred cases would be preferred. As a rough rule, there should be at least 10 cases per parameter contained in the model. Since we have no more than 7 or 8 terms in most of our models, we may be able to get away with 70 or 80 cases. Perhaps by moving down a level and using enumeration areas as cases, we could achieve sufficient sample size.

## 6.0 RELIABILITY ISSUES

In this section, we will discuss: (i) methods for enhancing reliability of summary measures, (ii) clarifying the distinction between reliability of a measuring instrument and the reliability of a rater. We also suggest research strategies for dealing with these issues and the outstanding issue of reliability of individual questions.

### 6.1 Increasing Reliability of Summary Measures

Although the reliability of the linear composite indices has been tested using Cronbach's alpha, further reliability analyses are advised. The reliability of the summary indices could be increased to even higher levels if we used factor scales instead of the present summary indices. This simply entails adding together the highest loading items (only) on each dimension of a factor analysis to create summary indices. Hence, for each dimension, we would have a summary factor scale. The resulting scales could then be weighted by their respective eigen values or roots. Two points are important here. First, the resulting scales would be very similar to the present summary indices. Secondly, (as David Armor has argued), the factor scales will be guaranteed to be at least as reliable as the corresponding summary indices. Cronbach's alpha is replaced with Armor's theta.* Since the factor analyses have already been conducted, the production of factor scales would be simple to achieve.

---

* cf. Theta Reliability and Factor Scaling, David Armor, Sociological Methodology, 1974-1974.

It is also quite likely that the validity of the resulting scales would be increased. This is because both validity and reliability increase as the average inter-item correlation of the constituents of a scale increase. Factor analysis and factor scaling provide a method for grouping the most highly collinear items.

## 6.2 Distinguishing Instrument Reliability and Rater Reliability

A separate reliability issue should also be dealt with. At the Washington meetings, it was clear that there was some confusion regarding the distinction between the reliability of an instrument and the person administering the instrument. Cronbach's alpha (and theta) measure the reliability of the measuring instrument, not the individual administering the instrument. A ruler may be a perfectly reliable measuring instrument, however, results may be unreliable (i.e. vary from rater to rater) due to error or bias in the raters themselves. This type of reliability problem can be dealt with most rigorously through a test-retest method. This is an expensive process whereby individual ratings of the same object are repeated by different raters to see if the same result occurs. It is impractical to go back and conduct this type of analysis on the Ottawa pilot data base. This type of exercise was conducted, at an informal level, during the training sessions. In order to approximate this approach, and to establish the reliability of individual questions (in contradistinction to summary indices) alternate approaches such as split-halves may

be conducted on the existing data base. To deal with inter-rater variance, we would suggest an analysis of variance with groups being established in terms of different interviewers or experts.

## 7.0 TOWARD MODELS OF REHABILITATION NEED AND POTENTIAL

### Introduction

Having reflected upon the results of the Ottawa and Boston pilot studies, it is our belief that the problem of measuring physical house condition is largely solved. The problems of measuring and conceptually clarifying the companion concepts of rehabilitation need and rehabilitation potential are less definitively answered by these studies. Although some progress has been made towards the issue of rehabilitation need, rehabilitation potential remains *terra incognita*. In this section, we will discuss a research strategy and conceptual model for dealing with these issues.

### 7.1 Conceptual and Policy Issues

Physical house condition is a relatively narrow concept which can be defined in terms of the structure and function of a few major subsystems of the house. It is basically a physical property of the house. The concepts of rehabilitation need and potential necessarily entail non-material as well as material components. For instance, rehabilitation need is not merely a physical property of a building, but the product of the complex interaction of the physical entity (the house) and the household (the occupants). The rehabilitation needs of a particular building will vary considerably depending upon the age, size and composition of the household. For instance, two identical physical structures, one with one occupant, the other with 10 occupants will have different rehabilitation needs. Presumably,

the crowded house will typically have higher rehabilitation needs (all other things being equal). Rehabilitation "wants" refers to the rehabilitation desires or wishes of a household. For example, a new toilet may be a rehabilitation need if there is no toilet or the old one doesn't work. A new coloured toilet is a rehabilitation want since this is a cosmetic, not a functional request. Rehabilitation "wants" will vary according to other factors such as socioeconomic status and ethnicity. For example, E.T. Hall (1956) has shown that the perception and use of building space varies greatly across different cultures. Rehabilitation potential refers to the likelihood that successful rehabilitation activity will (or can) occur in a given area. This concept introduces a further layer of complexity to the rehabilitation problem since the financial capabilities of the household, market structure, residential attitudes, and neighborhood quality as well as the broader political and economic context must all be taken into account. Furthermore, rehabilitation potential is probably best conceived as neighborhood level property, whereas physical condition and rehabilitation need operate at the level of individual houses and households.

The Ottawa pilot study data base contains a range of data which will allow us to proceed much closer to defining and measuring these important concepts. Physical house condition may be the base concept, but rehabilitation policy and program design must be informed by knowledge of needs

and potential as well. No matter the technical elegance of a system for defining, measuring and improving physical house condition, it will not succeed in the real world if it is not responsive to the social and economic components of the problem. Program take-up rates are an excellent illustration of this fact. We know that the list of technically eligible properties for programs such as RRAP and OHRP does not equal the actual list of properties which apply for and receive program assistance.

## 7.2 Research Strategy

The most basic linkage for future study would be the relationship between physical house condition and the objective and perceptual characteristics of the household. Some key questions here would be: (1) does perceived residential satisfaction increase with physical condition and what is the form of this relationship (e.g. are there diminishing returns?); (ii) what effect does physical condition of the dwelling have upon the hierarchy of housing concerns an individual or household works with; (iii) what dimensions of physical house condition are perceptually salient (i.e. have the greatest subjective impact). For instance, does the heating usually take precedence over the plumbing? (iv) How are these relationships altered by the social, demographic and ethnolinguistic characteristics of the household? (v) How does perceived neighborhood quality affect these relationships?

Having analysed the linkage between physical condition and household composition and perceptions, and the way this relationship is modified by intervening variables such as background dwelling characteristics, it is important to see how the linkages work in an extended model which includes plans. For instance, what factors most affect whether or not an individual or household is likely to plan to undertake major rehabilitation (as opposed to moving or spending money on other things). At this point, we would want to include an additional set of contextual variables which might limit these previously discussed relationships. Household finances, neighborhood quality, zoning regulation, etc. would be examples. The last link in the model would be the impact of these earlier components upon actual behavior or conduct. This would include predictive models for activities such as program take-up, good and poor house maintenance routines, moving, etc. The resulting model would constitute a good approximation of rehabilitation need and potential. It could be elaborated and refined with multi-site data, but a good beginning could be conducted with the existing data.

Rehabilitation need is largely covered by the first component of the model - the subsystem linking physical condition to household characteristics and perceptions through a series of intervening or contextual variables. The latter parts of the model including limiting factors, plans and activities

covers rehabilitation potential. With nationally representative data and further refinement, the entire model would provide a good simulation of the Canadian rehabilitation market.

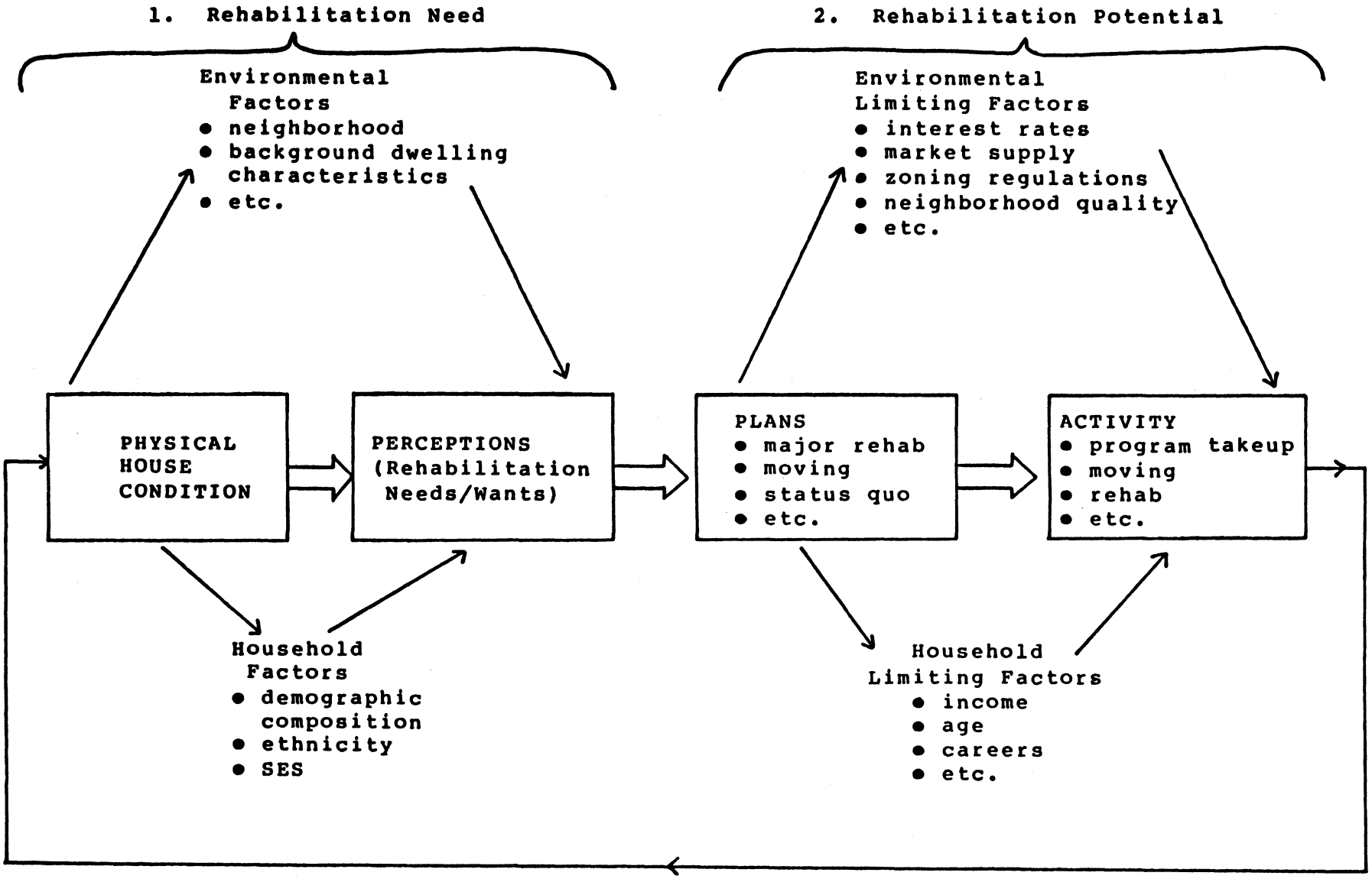
## 7.3 A Conceptual Model for Research

The basic model is displayed graphically in Figure 1.

Data resources currently exist to test the first half of the model (viz. the portion measuring rehabilitation need) in one population. One point brought up by several experts at the Washington Conference was that it would be possible to use an approach similar to the one sketched in Figure 1 as a method for identifying thresholds of rehabilitation need. It is possible to identify the threshold level where objective physical conditions become noticeable and problematic to the household. This threshold point or zone will probably vary with certain socio-economic characteristics of the household. This could be one additional input into the problem of defining a pass/fail point on a summary index of house condition (e.g. the point where conditions impinge upon the conscious perception of residential quality or the point where the resident will take action such as rehabilitation, moving, etc.). The view that operational measures (and definitions) of rehabilitation need could include both objective and perceptual indicators received support from several participants. Assuming a constant amount of government support (input) to a householder, it may be useful to give considerable

FIGURE 1

Towards Models of Rehabilitation Need and Rehabilitation Potential



weight to the characteristics and perceptions of the household as well as the material exigencies of the structure. Assuming all other things are equal, policy-makers should consider the preferences of the occupants in providing rehabilitation services and programs. Assume that two houses are in identical physical condition and also assume a fixed level of government support is available (e.g. $2,000). If one household would prefer to undertake energy related rehab whereas the other prefers a new roof, and if both are necessary and useful activities, we may want to support those activities which are most important and motivating to the occupants. The approach suggested here would give policy-makers and program designers a better idea of the needs and preferences of different combinations of households, houses and environments.

Expanding the model to include rehabilitation potential necessitates further data collection although good data on housing plans (<u>one</u> of the important variables for predicting rehabilitation potential) are available from the current survey. Rehabilitation potential should be studied at a neighborhood level. In order to establish the utility of this approach, a behavioral validation of the relationships between physical condition, household characteristics and perceptions, plans and rehabilitation activities is required. By having respondents from the planned RRAP Quality Evaluation survey complete the social survey portion of the Ottawa pilot study, we would have a sufficient sample of respondents who

did engage in program take-up. We would then be able to see the relationship between perceptions and activities.

## 7.4 Proposed Analyses

The questionnaires were designed to handle these problems with certain analytical techniques in mind. The relationship between physical condition and perception can be tested using correlation analysis, partial correlations and significance tests. The rank-order shelter scenario question asked the respondent to rank order various shelter options such as rehabilitation, moving, doing nothing, etc. Another similar question asked respondents to rank order the subjective importance of various factors deemed germane to residential satisfaction (e.g. location, physical house condition, neighborhood, etc.). Respondents also rank ordered the priority of various types of rehabilitation (e.g. cosmetic, energy conservation, etc.). Using these data and scaling technique, we could see how different types of respondents mentally organise their perceptions and images of different types of rehabilitation. There is a useful literature which shows how pertinent techniques such as non-metric multidimensional scaling can be related to market forecasting (cf. Green and Carmone, 1972, "Marketing Research Applications of Non-Metric Scaling Methods"). This analysis can also produce typologies and measures which will be valuable in predicting plans and activities. Multiple regression and log-linear modelling would be the appropriate techniques for these predictive models. With an expanded data base including neighborhood data, these models could be

elaborated to predict program take-up rates. Combined with knowledge of economic resources in a neighborhood, we would then have a good idea of what the economic (and social) costs and benefits of bringing certain types of rehabilitation programs to certain areas would be. More importantly, we would be able to predict the estimated costs of introducing a program in a given area and the likelihood of its success. It would also be possible to use the results of such an analysis to develop strategies for improving awareness and take-up of existing or potential programs. For example, information or educational programs could be more directly targeted at those potential program clients who are not aware of existing programs.

## 8.0 SAMPLING

Although many of the sampling issues are dealt with in the original research proposal and planning report, it is worth noting several points which follow from the experience of the Canadian and U.S. pilot studies. In this section, we will discuss the relative advantages of cluster and element sampling and other considerations which must be borne in mind in selecting a sampling strategy for further survey work. We will also consider the issue of what constitutes an appropriate primary sampling unit.

### 8.1 Comparative Discussion of Element and Cluster Sampling in Light of U.S. - Canada Pilots

The two stage tract sampling used by Abt was well suited for neighborhood modelling. However, the poor quality of the sampling frame they received posed problems in terms of response rates (e.g. abandoned or non-existent units). It may be problematic to equate census tract with neighborhood. This is because census tracts are not necessarily "natural" neighborhoods and if we are treating neighborhood as the unit of analysis, then we want that unit to have a certain corporate identity and internal homogeneity.* Otherwise, measuring and predicting certain properties and characteristics of that unit become meaningless. On the other hand, census tracts have the advantage of being pre-defined in non-ambiguous terms and of course can be linked to census ecological data.

---

* Census tracts are defined in terms of certain types of homogeneity but these may change radically through time and also the variables for grouping census tracts may not be those most germane to rehabilitation potential.

If a good sampling frame, such as a municipal tax roll, is available then an element sample is desirable. A probability element sample usually equates the unit of inference with the primary sampling unit. This is not to say that an element sample cannot support a neighborhood level of analysis. If we had stratified by census tract, as well as tenure, dwelling type and condition, in the Ottawa pilot sample, we would have easily been able to undertake aggregate census tract modelling. In a small sample, certain tradeoffs in the selection of stratifiers are inevitable. A cluster sample (e.g. where census tracts of city block are the primary sample units) has the advantage of economy, and in the absence of a known listing of the entire population, practicality. It is analytically more complex since the sampling units are of unequal size. It is also a less efficient sampling strategy.* Maximum precision and efficiencies will be achieved through a series of mini stratified random samples. It may also be desirable to use a series of subsamples. For example, a 1 in 4 subsample of basements linked to a full exterior inspection plus household interview could be beneficial. A small subsample of technical inspections should be conducted to calibrate the survey results achieved be trained interviewers.

---

* Efficiency here is used in the technical sense to refer to high precision (low variance) per element. Although this is the agreed upon meaning in the general statistical literature some sampling books use efficient to mean economic (i.e. low unit cost per fixed unit precision). Cluster samples are less efficient (in the statistical sense) because they have higher element variance (cf. L. Kish, 1965:150, also pp.161-164).

## 8.2 Selecting a Primary Sampling Unit

One outstanding issue is the question of an appropriate primary sampling unit. At the element level, the choices are - dwelling units, properties or buildings. Properties and dwelling units are problematic since many house condition problems - particularly serious structural problems are at a building level. Furthermore, the task of estimating costs is simplified if we use the building as the sample unit and unit of analysis. However, programs are delivered at a property level and Census data are collected at the household level. Selection of an appropriate primary sampling unit must consider these issues and decide on an individual basis. The requirement of an individual study must be considered in terms of factors such as desired precision, level of inference, available sampling frames, intended future uses, budget, etc. before this decision can be made.

## 9.0 POPULATION FORECASTS

Using the weight adjusted Ottawa pilot study data, it will be possible to produce population estimates concerning the levels of physical house condition, rehabilitation need and rehabilitation costs in the Ottawa inner city. Using the variances of the statistics employed, we will be able to specify confidence intervals around these predicted values.

By extending this approach we might be able to make more general forecasts concerning the long term demand for rehabilitation. In order to infer to other areas from the Ottawa data it will be necessary to employ a series of reasoned assumptions and to utilise available data from other sites. Scanada has developed forecast estimates for the demand for renovation in the Halifax market and extended the results to the nation. We would propose alternative statistical and theoretical approaches to the problem.

Ideally, we would prefer time series data to establish assumptions about the longevity of materials in certain environment, house maintenance practices and economic capabilities. Broader social, economic and demographic changes must be considered in any forecasting exercise (e.g. what role will increasingly limited energy supplies play in shaping the rehabilitation vs. new construction markets). Conceivably, a variety of existing

research, data and literature could be organised to assist in producing a national forecast model. Obviously, new multi-site data as well as expert opinion data would improve any forecasting exercise.

## 10.0 ADDITIONAL THEMATIC INDICES AND PREDICTIVE MODELS

In addition to constructing factor scales, there are a range of other useful summary indices which should be created from the Ottawa pilot data base. In the planning report, we discussed the importance of certain "themes" germane to physical house condition. These included themes such as "control of air and moisture" and "structural integrity". The data base is designed to routinely operationalise these themes according to the method described in the planning report. The resulting measures could constructively be used as both dependent and independent variables in future predictive models. We suggest that models predicting structural condition only will be more successful than models predicting overall physical house condition.

Another set of variables which should be analysed is the set of repair history and improvement history variables. These variables can serve as important control variables in predicting repair costs and physical house condition. It may also be useful to predict repair and improvement histories as subsequent inputs to rehabilitation market forecast models. A crude preliminary analysis of some of these variables shows that they are significantly associated with physical house condition and repair costs.

It is also suggested that additional separate modelling using occupant data only and interviewer data only be conducted. A preliminary regression analysis conducted recently

indicates that over 33% of the variance in repair costs can be accounted for using only the individual interviewer condition ratings. The corresponding occupant models provide about 20% levels of explanation. It should be noted that repair costs is the most demanding dependent variable and the American unit-level models only achieved around 26% levels of explanation with neighborhood data as well. These results merit further attention.

## 11.0 DELIVERING A HOUSE CONDITION ASSESSMENT METHODOLOGY TO LOCAL MUNICIPALITIES

It was clear from the Washington Conference that future rehabilitation research efforts may well bifurcate into two separate problems - (i) the problem of producing useful regional and national estimates, (ii) the problem of delivering a packaged methodology to municipalities. Although research on these two topics may be linked to a certain extent (e.g. multi-site testing), most of the strategies described in this report deal principally with the first topic.

The inspection methodology used in the Ottawa pilot study does work and should be tested in other municipalities. In conducting further testing, it is important to bear several considerations in mind. In Ottawa, the study initially intended to use municipal property standards officers (PSOs). Several problems became apparent. One of these was that PSOs would be in a difficult position since they would feel compelled to report major infractions that they encountered in the field. Naturally, this would not be ideal for a research survey where respondent cooperation was required. A volunteered response should not be rewarded with a fine. Another serious problem was that the survey work severely interfered with the PSO's normal work load. Consequently, CMHC inspectors were used in the Ottawa pilot. We suggest that both of these problems can be effectively dealt with in future developmental work. The solution is to conjoin additional instrument testing and survey work with ongoing monitoring

of municipal standards. By developing a regular survey which randomly identifies the worst portion of the housing stock, and having the PSOs administer the inspection instrument to those properties, municipalities would be able to both monitor and enforce local standards and acquire a good survey base. We would suggest that this effort be a regular, ongoing survey conducted at regular intervals. This would eliminate seasonal variations and provide a routinised and continually updated house condition information base for the municipalities.

We would suggest that a coordinated approach linking the resources of local municipalities and CMHC is appropriate. The municipalities would provide a sampling frame (e.g. tax rolls*) and PSOs to administer the instrument. CMHC would provide the instruments, a training package, sampling plan, data base management, data analysis and population estimates. The system could be delivered on a cost recovery basis to the municipalities. Both CMHC and the municipalities would have access to the resulting data base and forecasts. With increasing use, the system could be simplified and fine-tuned. For example, the Ottawa pilot has produced a simplified inspection form consisting of some thirteen items (questions).

---

* Tax rolls are often a preferred sampling frame since the elements are the actual physical properties rather than households which are used by the Census.

It is important that CMHC coordinate the system, and in particular, the training package, if the results are to be meaningfully compared from one centre to another. Once the system is in place and functioning properly, it could produce valuable social indicators of housing quality and housing needs. This would permit comparisons of various municipalities through space to see where resources are most urgently required. These indicators would also facilitate comparisons through time to assess progress towards certain goals and to assess the effectiveness of programs such as RRAP and CHIP.

For developing a sampling strategy for this approach, it will be necessary to have some base for pre-selecting the population in need of rehabilitation. This could be done at a census tract level using the 1981 Census Repair Need Question and possibly other data sources such as a revised HFE. Some combination of exterior only plus interior or exterior plus a subsample of basements would be advisable. It would also be recommended that future efforts to deliver a municipal inspection methodology be linked to multi-site testing of the 1981 Census Repair Need Question.

## 12.0 CONSIDERATIONS FOR MEASURING HIGH RISE HOUSING STOCK

As noted in the Major Report, the methodologies tested in the Ottawa pilot study are not necessarily generalisable to other types of housing with different types of respondents in different types of environments. Since the Ottawa pilot only dealt with low-rise stock and since a large portion of the Canadian population resides in high-rise dwellings, the question of how we modify this approach to deal with high-rise is a pressing one.

We do not have empirical grounds for suggesting research strategies here, but we can itemise a list of issues and concerns which must be considered. From the perspective of physical house condition, the building, rather than the property or dwelling, is the most sensible unit of analysis. However, if we consider rehabilitation need as the interaction of the physical dwelling with the household, then the dwelling unit may be considered the most appropriate unit of analysis.

In measuring the physical condition and rehabilitation needs in high-rise buildings, there are several important factors which must be borne in mind. The findings of the Ottawa pilot regarding the relationship between interior and exterior obviously do not apply. At the level of an individual dwelling unit, the ratio of interior to exterior will be very different than for a low-rise (i.e. instead of a roof and four exterior sides, there may be only one exterior side wall).

The mechanical systems for high-rise units will typically pose very different rating problems than those in low-rise units. They will be much larger, considerably more complex and centralised. There will be mechanical systems which are typically not evident in low-rise dwellings (e.g. alarm or sprinkler systems, garbage disposal chutes, etc.). Rating these will require special technical expertise which will be beyond the capabilities of trained interviewers. It is likely that certain problems with mechanical systems can be diagnosed through occupant factual observations (e.g. how many times did the power fail...?), and evaluations (e.g. how satisfied are you with the heating...?). With a sample of occupant responses from a single high-rise of this sort, we may be able to eliminate the capricious errors possible when we rely on a single respondent, as is the case in a single detached family home. Additional problems in gaining access to areas containing the mechanical systems are likely, given that high-rises will often be owned by large corporations with no interest in participating in such a survey. The same applies to gathering financial data pertaining to the building.

Different types of themes may be applicable to high-rise apartments. For example, the issues of occupant privacy and acoustic transmission from one unit to another are crucial issues in a high-rise apartment. Similarly, as the size of the residential building increases, the common areas of the building play an increasingly important role in affecting occupant health, safety and satisfaction.

APPENDIX I

## Annotated Bibliography and List of Log-Linear and Related Techniques


Allison, Paul D. 1980 "Analyzing Collapsed Contingency Tables Without Actually Collapsing", American Sociological Review 45 (February), pp. 123-130.

The author demonstrates that similar results to log-linear models fit to collapsed contingency tables can be obtained by fitting log-linear models to the full table, if certain parameters pertaining to the collapsed table are excluded. The approach is recommended as particularly useful for the analysis of recursive systems of categorical variables.

Bishop, Y., Fienberg, S., and Holland, P. 1975 Discrete Multivariate Analysis: Theory and Practice, MIT Press, Cambridge, Mass.

A comprehensive, highly theoretical and highly technical presentation of various techniques for discrete multivariate analysis.

Davis, J. 1974 "Hierarchical models for significance tests in multivariate contingency tables: an exegesis of Goodman's recent paper", pp. 189-231 in H.L. Costner, Ed., Sociological Methodology 1973-1974, San Francisco: Jossey-Bass.

A non-technical review of the hierarchical significance testing a component of Goodman's system (the other component is log-linear modelling).

Fienberg, Stephen E. 1977 The Analysis of Cross Classified Categorical Data, MIT Press, Cambridge, Mass.

A good discussion of logits and log-linear models given a less technical presentation than Bishop, Fienberg and Holland text.

Goodman, L.
1970 "The multivariate analysis of qualitative data: interactions among multiple classifications", Journal of the American Statistical Association 65: 226-256.

A seminal work by Leo Goodman.

Haberman, Shelby J.
1978 Analysis of Qualitative Data Volume I Introductory Topics, New York, Academic Press.

A recent review of the field with social scientific examples by a colleague of Leo Goodman's at the University of Chicago.

Holt, D.
1979 "Log-linear models for contingency table analysis: on the interpretation of parameters". Sociological Methods and Research 7 (February) pp. 330-336.

A good cautionary discussion concerning the interpretation of log-linear parameters.

Knoke, D.
1975 "A comparison of log-linear and regression models for systems of dichotomous variables", Sociological Methods and Research 3 (May): 416-434.

This article shows that the estimated parameters of a log-linear and regression model will be similar if the split on the dependent variable is not extreme.

Knoke, D. and Burke, Peter J.
1980 Log-linear Models, Sage University Papers Series: Quantitative Applications in the Social Sciences, John L. Sullivan (ed.), Beverly Hills, Sage Publications.

A non-technical introductory publication for the uninitiated. This text defines and discusses the terms of log-linear analysis and gives numerous examples.

Swafford, Michael
1980

"The parametric techniques for contingency table analysis: a non-technical summary". American Sociological Review 45 (August) 664-690.

A non-technical presentation of a direct comparison of three techniques for contingency table analysis; linear probability, logit, and log-linear modelling. The paper highlights the availability of linear probability modelling and describes errors common to published applications of log-linear modelling.

# Workshop on Measuring House Condition and Rehabilitation Need

May 4-5, 1981

## Participant List
(Revised)

David Carr
Decatur Housing Authority
1808 E. Locust Street
Box A1409
Decatur, Illinois 62525

Steve Casey
Department of Statistics
Rutgers University
New Brunswick, New Jersey 08903

Philip Morrison
Department of Regional Science
University of Pennsylvania
3718 Locust Walk CR
Philadelphia, Pennsylvania 19104

Judith Feins
John Kirlin
Antony Phipps
Abt Associates
55 Wheeler Street
Cambridge, Massachusetts 02138

David Koons
Bureau of the Census
U.S. Department of Commerce
Washington, D.C. 20233

Mary Nenno
National Association of Housing
and Redevelopment Officials
Division of Policy Development
2600 Virginia Avenue, N.W.
Suite 404
Washington, D.C. 20037

Jonathan Sunshine
VA Scholars Program
107 Veterans Administration
Central Office
810 Vermont Avenue, N.W.
Washington, D.C. 20420

Anthony Yezer
Department of Economics
George Washington University
Washington, D.C. 20052

Arthur Young
Housing Division
Bureau of the Census
U.S. Department of Commerce
Washington, D.C. 20233

Laurent Hodes
George Mansfield
Howard Sumka
David Varady
Division of Community Conservation
Research
Office of the Assistant Secretary
for Policy Development & Research
Department of Housing and Urban
Development
451 Seventh Street, S.W.
Washington, D.C. 20410

John Gardner
Sue Neal
Evaluation Division
Office of the Assistant Secretary
for Policy Development & Research
Department of Housing and Urban
Development
451 Seventh Street, S.W.
Washington, D.C. 20410

Howard Hammerman
Jill Khadduri
Jennifer Stucker
Housing Assistance Research Division
Office of the Assistant Secretary
for Policy Development & Research
Department of Housing and Urban
Development
451 Seventh Street, S.W.
Washington, D.C. 20410

Robert Gray
Policy Development Division
Office of the Assistant Secretary
for Policy Development & Research
Department of Housing and Urban
Development
451 Seventh Street, S.W.
Washington, D.C. 20410

Duane McGough
Kathy Nelson
Housing and Demographic Analysis
Division
Office of the Assistant Secretary
for Policy Development & Research
Department of Housing and Urban
Development
451 Seventh Street, S.W.
Washington, D.C. 20410

John Simonson
Office of Economic Affairs
Office of the Assistant Secretary
for Policy Development & Research
Department of Housing and Urban
Development
451 Seventh Street, S.W.
Washington, D.C. 20410

Philip W. Brown
Katherine Kemp
Market Forecast and Analysis Division
Canada Mortgage and Housing Corporation
Montreal Road
Ottawa, Ontario K1A OP7

Paddy Fuller
Statistical Services Division
Canada Mortgage and Housing Corporation
Montreal Road
Ottawa, Ontario K1A OP7

Robert Hale, Jr.
Residential Improvement Division
Canada Mortgage and Housing Corporation
Montreal Road
Ottawa, Ontario K1A OP7

Naomi Lee
Federal Statistical Activities
Secretariat
Station 3A, Third Floor
Statistics, Canada
Jean Talon Building
Tunney's Pasture
Ottawa, Ontario K1A OT6

Boris Mazikins
Housing Survey & Dissemination Section
Consumer Income and Expenditure
Division
Statistics, Canada
Jean Talon Building
Tunney's Pasture
Ottawa, Ontario K1A OT6

Gilles Boivin
Canadian Association of Housing and
Redevelopment Officials
Service de la Restauration de la
Conservation du Patrimoine Residentiel
330 Rue St. Paul Est
Montreal, Quebec H2Y 1C6

Dale Bishop
MLC Associates
309 Cooper Street
Suite 210
Ottawa, Ontario K2P OG5

Eric Moore
Department of Geography
Queen's University
Kingston, Ontario

Frederick Ermuth
Francis Graves
Ekos Research Associates, Inc.
171 Nepean Street, Suite 304
Ottawa, Ontario
K2 POB4 Canada

The American University
College of Public and International Affairs
Center for Urban Public Policy Analysis



# AGENDA

## Workshop on Measuring House Condition and Rehabilitation Need

May 4-5, 1981

Sponsored by U.S. Department of Housing and Urban Development and Canada Mortgage and Housing Corporation.

### Monday, May 4

| | |
|---|---|
| 9:00 - 9:30 a.m. | Introduction<br>Overview of the research and its design; history and objectives; the Cambridge workshop; cooperation and comparisons, U.S. and Canada; goals of this workshop.<br>- CMHC and HUD |
| 9:30 - 10:45 a.m. | Overview of the study approach and its major findings<br>- Abt Associates, Inc.<br>- Ekos Research Associates, Inc. |
| 10:45 - 11:00 a.m. | Coffee |
| 11:00 - 12:00 noon | Review of the housing measures, survey instruments and survey procedures used. Discussion.<br>- Abt Associates, Inc. |
| 12:00 - 1:00 p.m. | Lunch<br>Blue Room - Mary Graydon Center |
| 1:00 - 2:00 p.m. | Review of the housing measures, survey instruments and survey procedures used (continued). Discussion.<br>- Ekos Research Associates, Inc. |
| 2:00 - 2:15 p.m. | Coffee |
| 2:15 - 4:30 p.m. | Analysis and results of the measurements.<br>Derived measures of need and cost; comparison of inspector and occupant evaluations; comparison of interview and inspector evaluations (Canada); comparison of interior and exterior conditions. Discussion.<br>- Ekos Research Associates, Inc.<br>- Abt Associates, Inc. |

AGENDA
(Continued)

Workshop on Measuring House Condition and Rehabilitation Need

May 4-5, 1981

Tuesday, May 5

| | |
|---|---|
| 8:45 - 9:45 a.m. | Analysis of the 1981 Canadian Census question. Discussion.<br>- Ekos Research Associates, Inc. |
| 9:45 - 10:00 a.m. | Coffee |
| 10:00 - 11:30 a.m. | Comparison of U.S. results to existing measures. AHS composite measures; predictive models. Discussion.<br>- Abt Associates, Inc.<br>- Ekos comments |
| 11:30 - 12:30 p.m. | Lunch |
| 12:30 - 2:00 p.m. | Investigation of alternative strategies for collecting data and estimating physical house condition and rehabilitation need (e.g.: full-scale technical inspection; skim technical inspection; household interviews).<br>- Discussion |
| 2:00 - 4:00 p.m. | Directions for further research and analysis. Outline of a research design for the best alternative(s) given policy constraints.<br>2:00 - 3:00 - Canada - CMHC<br>3:00 - 4:00 - U.S. - HUD |
| 4:00 - 4:30 p.m. | Conclusions and wrap-up. |

## Summary of the American (Boston) Pilot Study

The Boston pilot study was designed to test new methods for assessing housing condition and the need for rehabilitation in the nation's housing stock. The study derived measures of housing condition, rehabilitation need and rehabilitation cost and used inspection data from a sample of 290 housing units in 30 lower-income census tracts in Boston, Massachusetts. The feasibility of using interview data similar to those collected by the Annual Housing Survey to predict the derived measures for individual dwelling units and for neighborhoods was explored. In general, the study was guided by five research objectives:

1) Develop a survey instrument designed to measure physical house condition and rehabilitation need;

2) Develop a method of estimating rehabilitation costs from the data;

3) Assess the instrument design and its use in a one-city pilot test of 300 housing units;

4) Assess current housing condition measures and identify the problems;

5) Assess the feasibility of estimating physical housing conditions, rehabilitation need and rehabilitation cost using data from the Annual Housing Survey.

The data were collected during a ten-week period in the fall of 1980. Two-person survey teams were sent to each dwelling. One person administered an AHS-type interview to an adult member of each dwelling while a physical inspection of the dwelling was completed by a trained evaluator.

The findings with respect to housing condition showed that sixty-two percent and eighty-three percent of the houses were inadequate according to narrow (easy) and broad (strict) interpretations of Housing Quality Standards and Acceptability Criteria. The most frequently encountered deficiencies were electrical hazards, dangerous porches, stairs, railings and ceiling surfaces, and inadequately heated bedrooms. There was little correlation among the classes of deficiencies and exterior inspections alone would have missed four out of five units failing the narrow interpretation.

Based on a measure of rehabilitation need which emphasised health and safety factors, forty-one percent of dwelling units were found to be in need of major repair; thirty-six percent needed moderate repair; fifteen percent needed minor repair; and eight percent needed no repairs. Among dwellings in need of rehabilitation, fifty-six percent required exterior repairs (only ten percent required structural repairs), while interior repair needs were concentrated in the categories of room surface and structure (fifty-six percent) and electrical systems (sixty percent).

Using a standard manual of rehabilitation costs for the Boston SMSA, average repair costs were $961 for the narrow interpretation and $1,073 for the broad interpretation. The median costs were $367 and $440 respectively. About one-third of the units failing the narrow measure could be repaired for less than $200 with another forty-four percent being repairable for between $200 and $1,000. The two most frequently failed

standards (illumination and electricity, and structure and materials) comprised two-thirds of all deficiencies but contributed only one-third of the total repair costs.

An analysis of the correspondence of interview and inspection results showed a relatively high degree of inconsistency, with the percentage of cases in disagreement usually falling in the twenty to forty percent range. This disagreement was highest for wall and ceiling conditions and electrical problems and lowest for plumbing facilities and egress in case of fire. Using the inspection-based (narrow) measure as a reference, four frequently used aggregate measures of condition from the AHS showed a relatively high incidence of classification error. In the aggregate, classification errors occurred for thirty-eight to forty-five percent of the sample, thus leaving substantial room for improvement in the use of AHS data to estimate housing condition and rehabilitation need.

The development of predictive models of housing condition and rehabilitation need appear to be feasible and to have useful policy implications. Predictive models, at both the unit and census tract level, were developed using inspection-based cost and condition data as dependent variables, and (1) occupant perceptions, (2) characteristics of the structure and household from the AHS and (3) census and survey data about the neighborhoods as independent variables. The unit based regression models showed modest levels of predictive power. Models for three cost measures explained from

twenty-six to twenty-eight percent of the variance in costs, while models for three condition measures accounted for thirty to thirty-two percent of the variance. Neighborhood-level models of condition and costs showed relatively high levels of predictive power, despite the small sample size of thirty census tracts. Models were developed to predict the proportion of units in a tract with rehabilitation need, the proportion failing condition measures, the average rehabilitation costs and the average number of deficiencies. These models explained between thirty-seven and seventy-seven percent of the variance with all three sets of independent variables contributing significantly.

The results of this test of modelling feasibility are encouraging. However, the pilot sample was drawn from a single site and a relatively homogenous housing stock and further research must explore the effects of introducing heterogenous elements (city and suburban locations, multi-family dwellings) more typical of the American urban housing stock.

ABT ASSOCIATES INC.
55 WHEELER STREET, CAMBRIDGE, MASSACHUSETTS 02138
TELEPHONE • AREA 617-492-7100

AAI #81-20

# FINAL REPORT ON THE MEASUREMENT AND PREDICTION OF HOUSING CONDITION AND REHABILITATION NEED

Contract #H-5002 - Task Order #2

July 27, 1981

Principal Authors:

Antony A. Phipps
Judith D. Feins
John A. Kirlin

Prepared for:

U.S. Department of Housing and Urban Development
Office of Policy Development and Research
Washington, D.C.

| Contract Manager | Quality Control Reviewer | Management Reviewer |
| --- | --- | --- |
| Antony A. Phipps | James E. Wallace | Helen E. Bakeman |

APPENDIX IV

# PROJECT REPORTS

## PILOT STUDY OF PHYSICAL HOUSE CONDITION AND REHABILITATION NEED

1. Planning Report for the Study of Physical House Condition, Rehabilitation Need and Potential. Submitted to Statistics Canada by CMHC, 5 June 1980.

   - Objectives
   - Content and analysis
   - Design of the information collection project.

2. Report on the Development of the Survey Instruments, 31 July 1980.

   - Benchmark in development of 4 major instruments
   - Inventory of research concepts, elements and items
   - Matrix relating instrument items to major areas of substantive concern
   - Survey instruments
   - Proposed analysis.

3. Report on Survey Procedures, 15 October 1980.

   - Documents methods of data collection, survey logistics, procedures and quality controls
   - Presents objectives, approach and schedule of training sessions
   - Includes interviewer manuals.

4. Report on the Quality of the Data Collected, 15 September 1980.

   - Assessment of data characteristics and quality in terms of response rates by questions, by properties/dwellings, etc.
   - Data coding and editing instructions
   - Research design considerations to enhance data reliability and validity.

5. Photo Essay, 17 March 1981.

   - Examines the relationship between internal and external dwelling characteristics and the derived house condition indices.

6. Critical Evaluation of the Survey Instruments, 17 March 1981.

   - Presents (preliminary) simplified survey instruments to measure physical house condition and rehabilitation need.

7. Major Report of the Pilot Study of Physical House Condition and Rehabilitation Need, 22 April 1981.

- Measuring physical house condition
  - Global ratings
  - Summary scales
- Measuring rehabilitation need
  - Operationalization of RRAP
  - Analysis of repair costs
- Linear models of condition (various data sources)
  - Prediction of expert from occupant and interviewer data
  - Relation between exterior, interior and overall measures
- The 1981 Census Repair Need Question
- Answering the Core Questions Underlying the Research.