Govdocs
Can6
Man1
M510.8
T43
no.2

Mincome Manitoba

Manitoba Basic Annual Income Experiment

The Sample Design and Assignment Model

Derek P. J. Hum
Michael E. Laub
Charles E. Metcalf
Donald Sabourin

UNIVERSITY OF MANITOBA
DOCUMENTS
MAR 10 1980
LIBRARY

Technical Report No. 2

1979

THE SAMPLE DESIGN AND ASSIGNMENT MODEL

OF THE

MANITOBA BASIC ANNUAL INCOME EXPERIMENT

by

Derek P. J. Hum
Charles E. Metcalf
Michael E. Laub
Donald Sabourin

MINCOME MANITOBA

Technical Report No. 2

FOREWORD

The Manitoba Basic Annual Income Experiment is designed to evaluate the economic and social consequences of a guaranteed annual income program based on the concept of negative income tax. Of particular research interest is the labour supply response of individuals and families containing non-aged, able-bodied members. The Experiment is a jointly-funded project of the governments of Canada and Manitoba and was collectively designed by researchers and officials of Mincome Manitoba, the Department of Health and Social Development, Manitoba, and the Policy Research and Long Range Planning Branch of the Department of National Health and Welfare, Ottawa. Mincome Manitoba is the agency established to administer the project and is solely responsible for all experimental operations. Seventy-five percent of the cost of the Experiment is funded by the Government of Canada; twenty-five percent is funded by the Province of Manitoba.

ACKNOWLEDGMENTS

The work on sample design and sample assignment for the Manitoba Basic Annual Income Experiment represents the distilled efforts of many participants and advisors. First and foremost the authors wish to acknowledge an intellectual debt to Professors John Conlisk and Harold Watts, whose assignment model served as the prototype for our own work. The authors also wish to acknowledge the substantial contribution of others to this effort, especially Mr. Brian J. Powell, Director, Experimental Research Unit, Department of Health and Welfare, Ottawa, and his staff, whose advice and assistance was immeasurable. Additional assistance and support on the development of the model was provided by Mr. Marius Lewin and other staff members of Mincome Manitoba. Among those whose advice was extremely valuable are Mr. David Kershaw of Mathematica and Professor David A. Wilton, University of Guelph, Ontario. Professor Mordecai Kurz, Stanford University, California, and Dr. D. Lee Bawden, Program Director, Human Resources and Income Security, the Urban Institute, Washington, provided helpful comments and much encouragement.

The opinions expressed herein are those of the authors and should not be construed as representing the opinions or policies of the Province of Manitoba, Canada or any agency of either government. This paper was drafted by Dr. Derek Hum in collaboration with Mr. Donald Sabourin, drawing freely from additional material prepared by Dr. Michael Laub and Dr. Charles Metcalf. Responsibility for any misleading interpretations or errors lie with D. Hum and D. Sabourin.

THE SAMPLE DESIGN AND ASSIGNMENT MODEL OF THE MANITOBA BASIC ANNUAL INCOME EXPERIMENT

THE SAMPLE DESIGN AND ASSIGNMENT MODEL OF THE MANITOBA BASIC ANNUAL INCOME EXPERIMENT

I. Introduction

The Manitoba Basic Annual Income Experiment is a large scale project jointly funded by Canada and Manitoba primarily designed to study the labour supply response of households to a guaranteed annual income. The design and conduct of the experiment includes selecting participants from a number of sites and assigning them to alternative negative income tax programs for a period of three years. The participants assigned to one of the several treatment programs are given income-conditioned payments for the three-year duration. In addition a control group was selected. The control group is not entitled to a guaranteed annual income but will provide valuable information for comparison purposes.

The design of the experiment involves many issues. The scope of the present paper is limited, however, to that of the sample design and allocation model of the experiment.[1] Both treatment and control group households provide valuable information but because the control group receives no payments the cost of an observation for the control group is very low compared to an observation assigned to any one of the treatment programs. In addition the payments received by a household in any one of the treatment programs over the course of the experiment may be quite substantial. In the context of the sample design and assignment problem the issue is to simultaneously determine (i) the total sample size, (ii) the allocation between controls and treatment categories, and

---

[1] Other design issues are described separately in "The Objectives and Design of the Manitoba Basic Annual Income Experiment" (in preparation).

(iii) among the latter, the assignment among the different treatment plans -- all within a fixed budget constraint.

The detailed manner in which a sample is chosen and allocated to various treatments is a fundamental issue in experimental design. Once the number of experimental treatments to use has been decided, some analysis approach must be specified to allocate the sample. The common analysis-of-variance (ANOVA) procedure would entail assigning to each experimental treatment an equal number of observations, given equal cost for each sample point. Despite possible modifications to account for the differing costs per treatment observation, unequal expected variances, different stratum populations, and unequal importance of experimental treatments, the ANOVA approach to sample allocation may not be efficient nor desirable for a variety of reasons.

The sample selection and assignment process of the Manitoba Basic Annual Income Experiment uses a more complex approach. The model uses census data on the population as well as information obtained in pre-experimental interviews of the potentially eligible population. The assignment model employed is a formal technique for optimally allocating observations or sample points among various experimental treatments in order to maximize the value of the information generated by the experiment. The basis of the assignment model is a rigorous benefit-cost analysis of the alternative sample allocations that are feasible within a given budget constraint. In this application assumptions are made explicit and benefits are measured in terms of reductions in the variances of certain predicted values. Costs are measured in monetary terms reflecting the financial budget constraint.

A formal assignment model provides the experimenter with a structured means of making two types of complex trade-off decisions. First, the experimental treatment in question is an income-conditioned negative income tax (NIT) program. Since several different NIT treatments are involved in the experiment, the costs of assigning an observation to a treatment cell varies with both the characteristics of the family (income, size, etc.) as well as the parameters of the NIT treatment. Thus the overall sample size itself is a variable that depends upon the pattern of allocation of observations. The question is: to what extent is the experimenter justified in trading expensive observations for more inexpensive observations in order to increase overall sample size, bearing in mind the extreme cost variations per observation that typify a multi-year income maintenance experiment?[1]

Still another trade-off arises from the fact that the treatment plans vary in terms of their degree of importance, whether indicated by research or policy relevance reasons. If some treatments are more important than others, then the experimenter should allocate the sample in order to improve the ability to predict the response over the most relevant region of design points. But, to what extent is it necessary to substitute sample points with relatively low policy importance in favour of sample points with relatively high policy importance?

---

1 It must be emphasized that contrary to most experiments wherein the costs of sample points are usually roughly equal, costs of various sample points in a negative income tax experiment can differ greatly. For example, the ratio of cost between the most expensive design point and a control design point is approximately 24 to 1. Therefore, the cost trade-off for different design points is not one-for-one but may be as high as 24 to 1.

The formal assignment model used by the Manitoba Basic Annual Income Experiment is an adaptation, albeit with substantial modifications, of the Conlisk-Watts model for optimizing experimental design for estimating response surfaces. While the analytic procedures of the formal assignment model were followed as a basic guideline, in certain instances minor adjustments were made on the basis of external considerations and other information judged important.

II. A Brief Review of the Experimental Design[1]

II.1 Experimental Objectives Influencing Design

The income maintenance experiment in Manitoba was designed to accommodate certain research objectives. The primary research objective of the Manitoba experiment is to investigate the impact of various guaranteed income programs on the behaviour of recipients, most particularly the labour supply response of recipients to income-conditioned payments. Although the design of the experimental payment system adapts a classical negative income tax (NIT) concept and the major research thrust is towards labour supply response, the experiment will generate evidence regarding a wide range of responses and will be relevant for a broad spectrum of income assistance programs having basic support levels and tax rates.

II.2 The NIT Program Features

The NIT system of the experiment provides for an annual support level (G) which depends upon both family size and composition, and represents the payment (guaranteed annual income) the family would receive if it had no other source of income or wealth. As the family's income increases, the payment for which it is eligible declines at a rate determined by a constant offset tax rate (t). Under this system families will continue to receive positive payments at a declining rate

---

[1] This brief review of the experimental design is included primarily to render the document self-contained. Full details of the experimental design are available in a separate report, "The Objectives and Design of the Manitoba Basic Annual Income Experiment" (in preparation).

up to the point where their income exclusive of the NIT is equal to the breakeven level (B). The breakeven level, the support level and the offset rate are algebraically related: $B = G/t$ so that specifying any two of these parameters determines the value of the third for constant tax rates.

II.3 The Experimental Treatments

The two basic experimental parameters are the support level and the negative tax rate since the primary research objective is to predict response (especially labour supply) to a range of NIT programs whose characteristics differ in terms of support levels and negative tax rates. In addition, the total costs of any income maintenance program will depend in large part on the magnitude of the actual support level and particular negative tax rate adopted. Accordingly the program characteristics of most interest and importance both for labour supply research as well as program cost estimates are the basic guarantee and tax rate.

Three different support levels and offset rates were employed in the experiment in order to allow detection of non-linearities in response. Eliminating the combination of the highest guarantee level with the lowest tax rate yields eight possible combinations of financial treatment plans. However, the addition of a control group not eligible to receive payments results in a total of nine experimental plans to which selected families can be assigned.

II.4 Integration with Other Tax Transfer Programs

The guarantee support level and the offset tax rate are the only experimental program parameters varied directly; other basic parameters are maintained constant over all financial treatment plans. To ensure that participants in the experiment are in fact facing the support levels and tax rates designated by their financial treatment, the multitude of other support levels and tax rates must be taken into account and neutralized for those families enrolled in the experimental NIT program and also receiving other income-conditioned government tax and transfer programs. The basic experimental design issue is that of maintaining control over the marginal tax rate faced by participants and the level of other government transfers. This is necessary since the problems associated with estimating the response of individuals who are subject to more than one tax rate as a result of simultaneous participation in several income-conditioned programs with differing eligibility conditions are not insignificant. Accordingly, all other government transfers received by an experiment participant are subject to a 100% tax and all other income-conditioned taxes paid by participants are subject to rebate, thereby integrating the NIT program with existing tax transfer systems while simultaneously ensuring that experiment participants face the support level and the tax rate of their financial treatment.

II.5 Features of the Payments System[1]

The support level is the amount guaranteed to the family over twelve months if it has no income at all. Payments, however, are made

---

[1] A complete description of the payments system is set forth in separate reports under preparation.

to participants on a monthly basis. Families report total income from all sources monthly and are paid a sum each month based upon reported total income in the previous month plus an adjustment amount. The adjustment amount results because reported income above the breakeven level in any month is carried forward and counted as income in later periods, when the income falls below breakeven. However, the accounting period, that is, the time interval over which income is measured for the purposes of determining payments, is one calendar year. Consequently, actual monthly payments received by the family over the accounting period are reconciled with their entitlement based upon the total income received within the same period. Underpayments to families are corrected and overpayments are recovered.

II.6 Characteristics of the Population and Sample Structure

The population of interest for the experiment is much more narrow than the population which one might visualize eventually encompassed by a universal guaranteed annual income program. Bearing in mind the primary research objective of labour supply analysis, cost and efficiency considerations dictate that certain individuals not be considered part of the eligible population for experimental reasons. Accordingly, the aged, the institutionalized, the disabled, and others are excluded on the grounds that the experimental treatments are unlikely to produce any relevant work behaviour response of interest.

Similarly, the research objectives serve to guide the classification and stratification of the population of interest. The population of interest extends over a variety of family types and a range of normal incomes. The aspects of family structure along which the sample is stratified include: (a) double-headed family; both heads working, (b) double-headed family; only one head working, (c) single-headed family, and (d) single (unattached) individuals. In view of the requirement to define a population of interest relevant for labour supply analysis, the resulting stratification reflects the hypotheses that unattached individuals will react differently from individuals with family responsibilities, that intact families will behave differently from single-headed families, and that significantly different labour supply response will be observed between secondary workers who were and were not working prior to the experiment. In addition, because it is expected that families with different income levels will respond differently to the experimental treatments, families are also stratified in terms of a discrete definition of estimated normal income; that is, an income measure from which estimates of the transitory components have been removed.

The population of interest is also confined to residents of Manitoba, and the sample selected is drawn from (a) the City of Winnipeg, (b) several smaller rural Manitoba communities, and (c) a community in Manitoba where all families and individuals are eligible for enrollment in the program. The samples drawn from Winnipeg and the several smaller communities are referred to as the urban dispersed and rural dispersed sample, respectively, in that the sample size is small relative to the communities from which they are drawn. Sample points in the last category are referred to as observations from the saturation site.

II.7 Special Experimental Design Issues

The income maintenance program in Manitoba is limited to three years of payments to participants. This may represent a problem for the interpretation of research results and the prediction of effects of permanent programs because it implicitly assumes that the experimental families will make the same adjustment within three years that they would to an indefinite length national program. Unfortunately budget constraints did not allow the design to incorporate the experimental horizon as a direct experimental variable.[1] While this problem is not one that can be neglected, it is expected that a three year program is long enough such that most experimental families will make a long-term adjustment.

For many families the process of regularly reporting their income or responding to periodic interviews may have an impact on their behaviour in addition to the effect of receiving payments. The reporting requirement may lead families to have more accurate views of their income and financial situation than they would otherwise. For policy purposes, it is important to be able to distinguish between these two purposes, responses. Accordingly, a portion of the control group also provide monthly reports of their income, thereby making the act of reporting itself an experimental variable.

Randomly drawn dispersed samples can be used to determine individual responses which are not confounded by macro or community effects. In addition, the effect of dispersion is to isolate treatment families

---

1 The Seattle and Denver Income Maintenance Experiments in the United States have a variable-duration dimension in design. Experimental families are assigned to financial treatments for either a three, five, or twenty-year duration.

so that NIT program parameters can be experimentally varied within the dispersed site. At the same time this very isolation may place treatment families in a highly artificial environment. In recognition of this the income maintenance experiment in Manitoba instituted a saturation site wherein all members of the community are eligible for payments. This single saturation site should provide a more accurate approximation to a real-life situation, provide valuable administrative and operational experience, and possibly generate information on community effects. Among its disadvantages is the impossibility of multiple financial treatments since any horizontal inequity introduced into a single saturated community would quickly become obvious to the participants. As well, a single saturation site precludes a pure control group for comparison. Since budget and other constraints ruled out the possibility of multiple saturation sites, saturation control in the experiment is achieved by dispersing control and treatment families on the same financial program as that of the saturation site throughout a number of Manitoba communities and their surrounding rural municipalities. Notwithstanding methodological problems the saturation site provides, at a minimum, the possibility for research and administrative experience into substantive issues of behavioural response topics in connection with the NIT program exactly comparable to conventional economic analysis but in a highly realistic social laboratory setting. Beyond that minimum, there is also the possibility of comparisons with the control families in the rural dispersed sites.

III. The Formal Assignment Model[1]

In designing an experiment to measure the behavioural effects of alternative income maintenance programs, a major design problem is how to maximize precision in estimating these effects given limited resources (financial, etc.) while also observing additional constraints. The purpose of a formal assignment model is to optimize the allocation of sample points among different experimental cells given that observations placed in different cells yield different benefits with respect to some objective function and represent commitments to incur different levels of costs. The formal assignment model employed by the Manitoba Basic Annual Income Experiment is composed basically of the following components; namely (1) a basic regression model, (2) a set of admissible regressor rows, (3) an objective function, and (4) a set of constraints on the choice set.

III.1 The Response Function and Regression Model

It is assumed that the behaviour under investigation can be described by a quadratic response function and that the objective is to design an N-observation sample to optimally estimate the response function:

---

1 The model employed is an adaptation of the Watts-Conlisk model. The exposition in this section draws heavily from Conlisk, J. and H. Watts: "A Model for Optimizing Experimental Designs for Estimating Response Surfaces", American Statistical Association Proceedings, Social Statistics Section, 1969, 150-6; Conlisk, J.: "Choice of Response Functional Form in Designing Subsidy Experiments", Econometrica, Vol. 41, No. 4, 1973, 643-656; and an unpublished manuscript by J. Conlisk. The notation in this paper corresponds to that of Conlisk and Watts.

$$(3.1) \qquad y_i = f(Z_{1i}, \ldots, Z_{ki})\, \beta + e_i \qquad (i = 1, \ldots, N)$$

where y_i is the dependent or response variable, the Z_{ji} are design variables subject to experimental control, either directly or indirectly, β is a vector of coefficients, e_i is a random error term, and $f(\cdot)$ is a known function so that equation (3.1) is linear in the unknown coefficients β. More specifically (3.1) may be written, in matrix notation, as the conventional regression model:

$$
\begin{array}{lllll}
 & y = X\beta + e & & & \\
(3.2) & E(e) = 0 & & \mathrm{Var}(e) = \sigma^2 I = E(ee') & \\
 & b = (X'X)^{-1} X'y & & \mathrm{Var}(b) = \sigma^2 (X'X)^{-1} &
\end{array}
$$

where y is the dependent variable vector, X is the regressor matrix, e is the error vector, β is the coefficient vector, b is the least squares estimate of β, and $\mathrm{Var}(\cdot)$ is the variance matrix operator.

III.2 The Design Space and Admissible Regressor Rows

The k independent variables Z_{ji} subject to experimental control are the design variables, and a given choice of the matrix $Z = [Z_{ji}]$ constitutes a particular design. The problem may be described as one of choosing some design Z, and therefore X, such that together with the response y generated by the experiment, efficient inferences about β can be made.

The k design variables Z_{ji} are subject to experimental control. They are selected, as in a normal regression context, on the basis of having an hypothesized effect on the response (dependent) variables. Consequently, the design variables are combinations of:

(a) socio-economic (stratifying) variables (family type, income level, family size, number of earners, etc.)

(b) treatment variables which are varied directly (guarantee level, offset tax rate).

Specifying a value for each of the k design variables yields a design point. Each observation on the design variables (each row of Z) will be constrained to lie at one of the m pre-chosen (and distinct) design points in the design space. Let the m design points be designated by the m k-element row vectors $Z_1, \ldots, Z_m$, and n_i be the number of observations taken at the i^{th} treatment. The design matrix will then comprise n_1 rows like Z_1 , n_2 rows like Z_2 , etc.

The regressor matrix X depends, row for row, on the design matrix Z , so that corresponding to each distinct row of Z (admissible row) is an admissible regressor row for X . The mapping from a row of Z to a row of X depends on the hypothesized functional relationship of the design variables on the response. The X and Z matrices are identical (except for an intercept term) only for the case in which the relationship is hypothesized as linear with no interaction terms. Usually, however, the design variables will contain terms of degree higher than one as well as interaction terms.[1] In sum, each row of Z represents an observation at a particular design point for which there is a corresponding admissible regressor row and each of these m distinct rows of Z (and also X) may be represented a number of times. Consequently if X_i is the i^{th} admissible regressor row corresponding to the i^{th} row of the design vector Z, the regressor matrix X will be composed of n_1 rows of X_1, n_2 rows of X_2, etc.

The admissible rows of the design matrix Z represent design points for which there are corresponding admissible rows of X. The m design points may be viewed alternatively as strata; the response is expected to be heterogeneous across strata and homogeneous within strata. Generally, the experimenter must choose the location of the design points so as to give the region of interest in the design space adequate coverage. By limiting the coverage of this region to m distinct design point locations, the design choice problem of an N-observation sample reduces from choosing the Nk elements of Z to the much simpler one of choosing m non-negative numbers $n_1, \ldots, n_m$. Conveniently, note that the total sample size N and the regressor cross-product matrix are given in terms of n_i and x_i

$$(3.3) \qquad N = \sum_{i=1}^{m} n_i \qquad X'X = \sum_{i=1}^{m} n_i x_i' x_i \quad .$$

---

1 In general the number, k, of independent variables in a design row vector need not be equal to the number, k, of regression coefficients. For example, with $k = 2$ such that $f(Z_{1i}, Z_{2i}) = (1, Z_{1j}, Z_{1j}^2, \mathrm{Log}(Z_{2j}))$, this leads to the $k = 4$ regression form: $y_i = \beta_1 + \beta_2 Z_{1j} + \beta_3 Z_{1j}^2 + \beta_4(\mathrm{Log}(Z_{2j})) + e_i$.

III.3 The Objective Function

Criteria to rank alternative designs is embodied by specifying an objective function to optimize. Adopting an estimation viewpoint the goal of the experiment may be stated as accurately estimating a vector $P\beta$ of linear combinations of the elements of β where P is a known matrix, the rows of which specify the linear combinations of regressor variable values for which a prediction is desired.[1] Therefore the objective is to minimize some measure of the prediction error of linear combinations of the estimated regression equation, which can be different for each design point.

The best linear unbiased estimator of $P\beta$ is $Pb = P(X'X)^{-1}X'y$ so it is natural to specify the objective function in terms of some scalar function of the matrix estimation error $\text{Var}(Pb)$.[2] In particular, assume the experimenter wishes to minimize a weighted sum of

---

1 The matrix P is chosen to reflect the region of prediction interest. For example, choosing $P = I$, the identity matrix, implies a primary interest in estimating the elements of β; that is, the individual terms of the coefficient vector. An alternative choice might be $P = (x_1, \ldots, x_m)'$ which implies a primary interest in estimating the height of the response surface above each of the m design points.

2 This assumes equal error variances in the regression model. If error variances are assumed to differ across the design points such that the error variance corresponding to the i^{th} regressor row x_i is $\sigma^2 \cdot \nu_i$, the experimenter can specify the values for ν_i and incorporate such information into the objective function. This is equivalent to a weighting process to remove heteroscedasticity from the regression equation by increasing the sample sizes of those cells with higher variances. After some experimentation it was decided to employ the equal variance assumption. First, no good estimates for ν_i were available. Additionally, the unequal variance effect is already inherent in the cost function since lower costs (payments) are associated with higher income cells, everything else being equal. Since lower costs imply a larger allocation to that design point, consequently, the additional effect of adding variance weights would have placed too much emphasis on low or non-payment households.

the variances of the elements of Pb . The objective function may be written as $tr[W\ Var(Pb)]$ where $tr(\cdot)$ is the trace operator and W is a positive definite diagonal weight matrix whose diagonal elements measure the relative importance (policy relevance) to the experimenter of the elements of $P\beta$. Accordingly, substituting from (3.2) and (3.3) and multiplying by σ^{-2} , the objective function can be stated as

$$\begin{aligned}\phi(n_1, \ldots, n_m) &= \sigma^{-2}\ tr[W\ Var(Pb)] \\ &= \sigma^{-2}\ tr[P'WP\ Var(b)] \\ &= \sigma^{-2}\ tr[P'WP\ \sigma^2\ (X'X)^{-1}] \\ &= tr[D(\sum_{i=1}^{m} n_i x_i' x_i)^{-1}]\end{aligned}$$

where $D = P'WP$.[1]

III.4 Constraints on the Choice Set

The experimenter's choice set will generally be constrained to a proper subset of all non-negative integer m-tuples $(n_1, \ldots, n_m)$. The design point sample sizes n_i must satisfy the budget constraint $\sum_i c_i n_i \leq C$ where c_i is the cost of one observation at the i^{th}

---

1 Minimizing the trace function $tr[Var(WPb)]$ may be viewed as minimizing a weighted sum of variances of elements of b. Note, however, that $tr[Var(WPb)]$ may also be written as $E[(b - \beta)'(P'W'WP)(b - \beta)]$ so that minimizing a weighted sum of variances of the regression coefficient estimates is equivalent to minimizing a weighted expectation of a quadratic loss function. For a further discussion of this and alternative forms of the objective function see Conlisk, J. and H. Watts, "A Model for Optimizing Experimental Designs for Estimating Response Surfaces", American Statistical Association Proceedings, Social Statistics Section, 1969, 150-6.

design point, n_i is the number of observations at the i^{th} design point and C is the total available budget. In addition there may be further external or arbitrary linear constraints delimiting the choice set: written $L(n_1, \ldots, n_m) = 0$. These additional constraints may involve minimum lower bound values, absolute maximum limits, specific budget shares allocated to various design points, etc.

III.5 Formal Statement of the Design Model

The experimental design problem for regression analysis consists of four basic components: (1) a response function form and regression model, (2) a set of design points and the corresponding admissible regressor rows x_i , (3) the objective function matrices P and W , and (4) budgetary and other constraints on the choice set. Formally, the problem can be stated:

$$\text{Minimize} \quad \phi(n_1, \ldots, n_m) = \mathrm{tr}[P'WP(\sum_{i=1}^{m} n_i x_i' x_i)^{-1}]$$

subject to

$$(3.5) \qquad \sum_i c_i n_i \leq C \quad ; \quad L(n_1, \ldots, n_m) = 0$$

$$n_i \geq 0 \qquad i = 1, \ldots, m \quad .$$

This is a well-behaved programming problem involving minimization of a convex objective function over a set of linear constraints.[1] In sum, the experimental design and sample allocation model has been redefined in terms of a flexible and tractable mathematical programming problem.

---

[1] Although the problem is strictly one in integer programming, in practice the unknowns were treated as continuous.

III.6 Choice of Response Function in Design: Relative Efficiency

The true behavioural response functional form is usually not known in advance. This problem is also aggravated by the fact that the design assignment generated by the programming model is extremely sensitive to the response functional form incorporated in the objective function. In addition quantitative techniques for assessing benefits and losses associated with different forms of the response function are often lacking.[1] We discuss briefly the two issues of the choice of the response function to employ in the experimental design and the problem of comparing alternative design assignments.

The sensitivity of the optimal design to the choice of response function is principally attributed to a paramount concern with efficiency although differential costs of observations are also important. For example, if the shape of the response function is assumed to be linear, the optimal design under equal costs per observation would assign all observations at boundary design points. On the other hand if the response function is non-linear, the optimal design must also assign some observations to interior treatment points in order to estimate curvature. In general, the higher the degree of the assumed polynomial describing the response surface and the greater the assumed irregularities, the more interior design points must be covered. In the extreme case where the response function consists solely of a set of dummy variables indicating whether or not the observation lies at a particular treatment point, the model is identical to a one-way

---

1 For a discussion of three possible approaches, see Conlisk, J., "Choice of Response Functional Form in Designing Subsidy Experiments", Econometrica, Vol. 41, No. 4, 1973.

analysis of variance design and all treatments receive substantial allocations.[1]

Given the sensitivity of design to alternative specifications of the response function some measure of comparative efficiency of different designs is useful. The relative efficiency of any design allocation $(n_1', \ldots, n_m')$ to a reference design $(n_1, \ldots, n_m)$ may be measured by the ratio of objective function values: $\phi(n_1', \ldots, n_m')/\phi(n_1, \ldots, n_m)$. The objective function $\phi(n_1, \ldots, n_m)$ can be seen upon inspection to be homogeneous of degree minus one in its arguments. This will permit the derivation of certain scale properties, most notably the fact that for a given percentage allocation $(n_i/N, \ldots, n_m/N)$, any change in the total budget C will result in an equiproportional change in the value of ϕ. For example, a doubling of the budget will double the sample size of all treatments and halve the value of the objective function. Since the objective function is a measure of variance magnitude, the relative efficiency measure can be given the following interpretation. If the ratio of objective function values associated with two designs is, for example, one half so that one assignment has .5 efficiency relative to another, a doubling of the budget is required to bring the one design up to the accuracy level of the other. Accordingly the relative efficiency measure can be used to make pair-wise comparisons of alternative design allocations.

---

[1] Just as the linear form assumes no curvature is present, the analysis of variance response form can be viewed as allowing for maximum curvature since it permits the response surface to have a different height over each design point. As well, although the general design model does not usually have an explicit solution, the special case of a response function consisting exclusively of exhaustive binary variables can be shown to give the usual one-way analysis of variance allocation; that is, design point sample sizes n_i proportional to weight, w_i, and inversely proportional to the square root of the cost of the observation, $c_i^{\frac{1}{2}}$.

III.7 Non-orthogonality of the Assignment Model Allocation[1]

Orthogonality is often stressed as a desirable feature in sample allocations because it permits hypotheses concerning a single design characteristic to be tested without having to control for variations in other design characteristics. In the context of a negative tax experiment having three support levels and three tax rates, orthogonality would be obtained if the probability of assignment to a specific support level was independent of the tax rate assignment. Phrased differently, the probability of assignment to a given support level/tax rate combination would be the product of the separate probabilities associated with the support level and tax rate. Such a design would allow simple tests of the effects of independent variations in support levels and tax rates as well as tests of interaction effects.

In a negative tax experiment the costs for different sample points vary greatly and the assumed response function is complex. Under these circumstances orthogonality can be shown to result in a highly inefficient and non-optimal sample allocation.[2] Particularly, the costs of alternative design points is highly correlated with the design characteristics. Consequently an orthogonal design which ignores this correlation is highly undesirable and inefficient. First the total sample size which can be afforded with a given budget is usually drastically reduced.

---

1 For a detailed discussion of this issue, see Metcalf, Charles E., "Sample Design and the Use of Experimental Data", Chapter V, Part IV, Final Report of the New Jersey Graduated Work Incentive Experiment, from which this section draws substantially.

2 See Conlisk, J. and H. Watts, "A Model for Optimizing Experimental Designs for Estimating Response Surfaces", American Statistical Association Proceedings, Social Statistics Section, 1969, 150-6.

Additionally, estimation of the hypothesized response function is impaired if important cells required for estimation receive too small an allocation relative to an optimal design associated with the response function.

The cost considerations can be partly taken into account by assigning "lower" probabilities to plans with either higher support levels or lower tax rates. However, this still does not allow the experimenter control over the ratio of the support level to tax rate, which ratio affects the level of payments. The risk therefore is that key plans with respect to a given response function will receive a smaller allocation than required and other plans of lesser importance with respect to the response function will receive larger allocations than desired. Further, if the key plans are among the less expensive ones (lower guarantees and/or higher tax rates) the total sample size is reduced and the efficiency with which the hypothesized response function can be estimated is decreased.

The design space of the Manitoba Basic Annual Income Experiment combines a form of orthogonality with a focus on a central design point. Each of the treatment parameters (support level or tax rate) is varied given the single, central value of the other. The distribution of treatment characteristics is non-orthogonal in the sense that the probabilities of assignment to specific support levels and tax rates are not independent since certain support level/tax rate combinations appear with zero probability. However, because of the prior restriction that each treatment parameter be varied given a single central value of the other, the regressor matrix associated with the specified linear relationship is orthogonal.

The design space of the Manitoba Basic Annual Income Experiment also contains two additional design points in addition to the design described above. The inclusion of these two design points increases the range of plans over which inferences can be made and adds variability to the design space in a regression context.

III.8 Divergence Between Initial Allocation and Final Treatment Sample Size

The formal assignment model employed by the Manitoba Basic Annual Income Experiment specifies the optimal allocation of sample points on the basis of the best prior estimated specifications of the response and pattern of sample development over time. In other words, the effective treatment sample sizes available for analysis at the end of a sufficiently lengthy experiment could differ substantially from the starting sample sizes recommended by the formal assignment model. This can occur due to attrition of treatment units, splitting and recombination of units, creation of new treatment units and misclassification of units. In sum, then, the effective treatment sample sizes may differ significantly from the starting treatment sample sizes and it is only the starting treatment sample sizes over which the experimenter has control.

The formal assignment model is sufficiently flexible in adapting to the problem of divergence between initial and effective treatment sample sizes. One possibility is to specify or estimate how the pattern of attrition, splitting, recombinations, etc. will vary across treatment cells and over time and incorporate this information in the formal model. Given some specification of these retention probabilities the formal assignment model can then generate optimal starting treatment sample sizes based upon desired effective treatment sample sizes. This approach requires for its success accuracy in predicting retention rate patterns over time and incorporation of such assumptions prior to execution of the experimental treatment.[1]

Under some circumstances it may not be possible to provide acceptably accurate estimates of retention rates prior to the start of the experiment. As well it may be necessary on occasion to substantially modify prior estimates or incorporate design changes to the experiment. In such situations the formal assignment model may be applied in the context of a sequential sampling strategy.

III.9 The Assignment Model in a Sequential Sampling Context

Suppose the experimenter already has some observations for the m design points and let the number of these initially given observations

---

[1] The Seattle and Denver Income Maintenance Experiments adopt a more elaborate approach of specifying conditional probability functions for attrition and splitting than did the Manitoba Basic Annual Income Experiment. See Conlisk, J. and M. Kurz, "The Assignment Model of the Seattle and Denver Income Maintenance Experiments", Research Memorandum No. 15, Stanford Research Institute, July, 1972. Retention (non-attrition) rates for the Manitoba assignment model were assumed to vary with the payment level, increasing linearly from .88 for controls to .96 for treatment units receiving payments of $1,000 or more per year. The typical attrited unit was assumed to cost half as much as an observation remaining intact for the duration of the experiment, and to contribute one-fourth as much as an intact observation to experimental objectives. Compared to the cost of an intact observation, the assumed attrition rates reduced the cost per initially assigned observation by two to six percent, and amount of usable information by three to nine percent. For further details see Appendix A.1.

be $n_1^o, n_2^o, \ldots, n_m^o$. The initially given observations may represent, for example, the current effective treatment sample sizes. Suppose also that an additional sample must be designated to supplement the initially given observations. The additional sample may be required because of design modifications to the experiment and/or the current effective treatment sample sizes differ from the final desired effective treatment sample sizes by an amount considered unacceptable.

Given the same response functional form to be estimated and the identical basic design matrix Z, the formal assignment model (3.5) can be written:

$$\text{Minimize} \quad \phi(n_1^o + n_1, \ldots, n_m^o + n_m) = \mathrm{tr}[P'WP(\sum_{i=1}^{m} (n_i^o + n_i)x_i'x_i)^{-1}]$$

(3.6) Subject to

$$\sum_i c_i n_i \le C \quad ; \quad L(n_1, \ldots, n_m) = 0$$

$$n_i \ge 0 \qquad i = 1, \ldots, m \quad ,$$

hence the assignment model could be adapted to accommodate a sequential strategy and to design a sample allocation to supplement a set of initially given observations.

IV. Specification of the Mincome Manitoba Assignment Model: Winnipeg Site

The main objective of the Manitoba experiment was to study the behavioural response of recipients to various NIT schemes. The application of the assignment model was a formal technique for optimally allocating observations or sample points among the various experimental cells or strata in order to maximize the value of the information generated by the experiment. The details of the assignment model that require specification are as follows:

(1) a basic regression model (the hypothesized response function),

(2) a set of m design points (admissible regressor rows, x_i),

(3) a set of prediction points represented by rows of the matrix P (the linear combination(s) of regression coefficients for which prediction error is to be minimized),

(4) a set of prediction-point weights represented by diagonal elements of the matrix W (the relative importance of each of the design points),

(5) a set of design-point costs (estimated cost of an observation for each of the design points),

(6) a total budget amount, C,

(7) additional internal and external constraints.

IV.1 The Regression Model (Response Function)

The behavioural response function of the experiment was hypothesized to be of the form:

$$R = F(E^*, G, M, t, B) \tag{4.1}$$

where

R = response of interest

E^* = the "normal" level of income for the household that would apply over the pre-experimental period in the absence of any negative taxes (adjusted for family size and number of adults in the household)

G = support level or guarantee (amount of payment if all other income equals zero)

M = expected level of payments for the household calculated on the basis of pre-experimental income

t = negative tax rate (offset tax rate)

B = a dummy variable which
= 1 if normal income is below breakeven point (G/t)
= 0 otherwise.

Having specified the response function for the regression model, the objective function for the assignment model is simply the weighted sum of the variances of elements of the regression coefficient vector. More specifically, the objective function used for the Mincome Manitoba assignment model involved a weighted sum of four objective functions defined over the four (family type) population groups, each with a separately identified response function: The set of weights employed are listed in Table IV.1 below and represent the approximate distribution of family types in Canada.

TABLE IV.1

Family-Type Weights Applied to Population Group

Family Type	Weight
Double-Headed Multiple Earner	.35
Double-Headed Single Earner	.35
Single-Headed Households	.10
Unattached Individuals	.20
	1.00

Conceptually, the use of separate response functions (one for each family type) can be viewed equivalently as the inclusion of family type group interaction dummies on all other coefficients; that is, it assumes that the estimated response function across family types will yield significantly different regression coefficients for all variables. The same 17-term response model was assumed for each of the four family types and may be written:

$$R = B\cdot(b_1 + b_2 E^* + b_3 E^{*2} + b_4 G + b_5 G^2 + b_6 M^2 + b_7 M^3 + b_8 t + b_9 t^2 + b_{10} E^*\cdot G + b_{11} E^*\cdot M + b_{12} E^*\cdot t + b_{13} G\cdot M + b_{14} G\cdot t + b_{15} M\cdot t) + (1 - B)(b_{16} G + b_{17} t) \qquad (4.2)$$

The above model is linear in parameters and includes all possible interaction terms. A quadratic regression form was used since there was no evidence for assuming a linear response function. Consequently, allocating points to the interior of the design space allows the experimenter to test for the presence of curvature. The linear payments term, M, was dropped because of its redundance with E^*, t and G.[1] The above-breakeven dummy reflects the hypothesis that households receiving no payments would respond differently from those households receiving some payments. The above-breakeven dummy implies that above-breakeven households are used in estimating two of the seventeen terms in the regression model, whereas below-breakeven households are used in estimating the remaining fifteen terms in the regression model.

IV.2 The Set of Design Points

There are four design variables. They are: (1) the guarantee level, G ; (2) the offset tax rate, t ; (3) the normal income level, E^* ; and (4) the family type, f ; hence a design point is a value for the quadruplet (G, t, E^*, f) . It follows from the regression model (4.2) that the i^{th} design point (G_i, t_i, E_i^*, f_i) is of the following form:

---

[1] Aside from a component which reflects the amount of the positive tax rebates, M is totally redundant with E^*, t and G. The inclusion of M as a variable would have introduced multicollinearity in the regression equation.

$$x_i = (1,\ E_i^*,\ E_i^{*2},\ G_i,\ G_i^{\,2},\ M_i^{\,2},\ M_i^{\,3},\ t_i,\ t_i^{\,2},\ E_i^*G_i,$$

$$E_i^*M_i,\ E_i^*t_i,\ G_iM_i,\ G_it_i,\ M_it_i)$$

(4.3) if the household is below breakeven, and

$$= (G_i,\ t_i)$$

if the household is above breakeven.

A total of 162 design quadruplets (G, t, E^*, f) and corresponding x_i were specified, made up of nine (G, t) combinations (including a control group) at each of eighteen family type income level cells.

The E^* levels chosen are specified in Table IV.2, and are measured in dollars per year and adjusted (except for single individuals) to a family of size four.

TABLE IV.2

Pre-experiment Normal Income Cells

(Normalized for a Family of Four Members)[a]

Family Type	Pre-experiment Normal Income Interval	Mean Assumed for Interval When Needed (i.e., E^*)
Double-Headed	0 - 2999	1300
	3000 - 4999	4000
	5000 - 6999	6000
	7000 - 8999	8000
	9000 - 12999	11000
Single-Headed	0 - 999	400
	1000 - 2999	2000
	3000 - 5999	4500
	6000 - 9999	8000
Single (unattached)	$0 \leq E < 999$	400
	$1000 \leq E < 1999$	1500
	$2000 \leq E < 2999$	2500
	$3000 \leq E < 4999$	4000

[a] For single individuals, the normal income cells are not normalized for a family of four, but represent actual level of income.

The nine experimental plans are shown in Table IV.3. The eight treatment plans differ only with respect to the basic parameters G and t, the guarantee level and the offset tax rate respectively.

TABLE IV.3

Experimental Plans: Treatment and Control

		Tax Rate (t)		
		.35	.50	.75
Guarantee (G)	3800	Plan 1	Plan 3	Plan 6
	4600	Plan 2	Plan 4	Plan 7
	5400		Plan 5	Plan 8

Plan 9 = Controls

The offset tax rates and guarantee support levels were chosen to bracket the ranges thought to be of greatest interest for policy. Tax rates vary from a low of 35% to a high of 75%, which is similar to the ranges in existing welfare programs. Support levels also tend to bracket the range of greatest policy relevance. Other considerations used in choosing the support levels (tax rates) were: (1) that the range of support level (tax rates) be broad enough to permit separate measurement of the income effect; (2) that the lowest support level yield an acceptable degree of domination over the existing welfare programs; (3) that no tax rate be so high as to substantially remove the work incentive;[1] and (4) that the several guarantee level/tax rate combinations adopted

1 The point at which a tax rate becomes sufficiently high to substantially remove the work incentive is open to discussion. Specifically, a tax rate of 100% or higher is excluded.

in conjunction with the sample points assigned to them, satisfy a budgetary constraint.

IV.3 The Set of Prediction Points (P Matrix)

Given the set of design points and their corresponding x_i's a P matrix must be specified which will determine the linear combinations Pb of the regression coefficients b whose variance is to be minimized. The objective of the Manitoba experiment was viewed as estimating the height of the response surface over each cell (design point). The prediction matrix P becomes in this case the X matrix; that is, the regressor matrix itself. Therefore, Pb is equivalent to Xb, the expected value of each of the points in the design space.

IV.4 The Set of Prediction Point Weights (W Matrix)

All that remains to be specified to construct the objective function $\phi = tr[(P'W'WP)\ (\sum_i n_i\ x_i'\ x_i)^{-1}]$ is the diagonal matrix of policy weights W. The sizes of the diagonal elements represent the relative importances to the experimenter of the elements of Pb; that is, the relative importances of the prediction points.

The policy weights for each of the design points is the product of three components; namely,

(1) income class weights: reflecting the income distribution of the Canadian population, adjusted for family size,

(2) treatment weights: reflecting the relative policy importance of each of the (G,t) combinations, and

(3) pyramidal weights: so designed because policy interest centered more heavily on households near the poverty level (the working poor) than on those with very low or no income (and thus not likely to exhibit a substantial labour force response) or on those with much higher incomes. These weights include a piecewise linear component which rises from one at zero income to two at the lowest support level; and to zero at four times the support level.

The treatment weights are set forth in Table IV.4 below as well as the income class weights combined with the pyramidal weights in Table IV.5.

TABLE IV.4

Treatment Weights

		Offset Tax Rate		
		.35	.50	.75
Support Level	3800	.10	.30	.15
	4600	.05	.20	.10
	5400	-	.05	.05

TABLE IV.5

Income Class Weights

(Combined with Pyramidal Weights)

Family Type	Pre-experimental Interval	Weights
Double-Headed (Multiple Earner)	0 - 2999	.10
	3000 - 4999	.30
	5000 - 6999	.21
	7000 - 8999	.19
	9000 - 12999	.20
		1.00
Double-Headed (Single Earner)	0 - 2999	.16
	3000 - 4999	.25
	5000 - 6999	.25
	7000 - 8999	.21
	9000 - 12999	.13
		1.00
Single-Headed	0 - 999	.12
	1000 - 2999	.47
	3000 - 5999	.33
	6000 - 9999	.08
		1.00
Single Individual	0 - 999	.21
	1000 - 1999	.40
	2000 - 2999	.26
	3000 - 4999	.13
		1.00

Control Weights

Finally, the weight placed on control households in a given socio-economic strata $(E^*$, f space combination) is the square root of the sum of squares of the treatment weights in that socio-economic strata, i.e.,

$$\left[\sum_{i=1}^{8} (\text{treatment weight})_i^2 \right]^{\frac{1}{2}}$$

The above weighting scheme is the appropriate one under the assumption the definition of the experimental objective is the measurement of behaviour in each treatment cell as a first difference from that in the corresponding control cell.

Further, the weight on any treatment cell falling below break-even was doubled before being squared in order to place more weight on the below breakeven cells. Thus in relative terms, a control cell with many of the corresponding treatment cells above breakeven would have a smaller weight than a control cell with most of its corresponding treatments below breakeven; the relative importance of a control cell thus depends on the treatment weights in its socio-economic strata.

IV.5 Additional Constraints Imposed on the Choice Set

It has been noted that the sample assignment is sensitive to the response form specified by the experimenter. This illustrates a design dilemma as Conlisk notes:

"Prior beliefs about continuity of behavioral response...(may) suggest that (the response function) takes a simple continuous form such as the linear form. However a design based on such a simple form leaves the experimenter very ill protected if response surface irregularities require a more complicated form...At the other extreme, if the experimenter gives himself the complete protection against irregularities provided by specifying the analysis of variance form for (the response function), the resulting design is inefficient in the likely event that a simpler form is the true one..."[1]

(i) Treatment Minima Constraints

One approach to the problem that no design can be efficient with respect to all response functional forms is to impose treatment minima constraints. This approach would represent for the experimenter a balance between the attraction of a simple continuous response function and the desire for protection against irregularities of response.

The Mincome Manitoba assignment model incorporated a lower bound constraint of two observations in each treatment cell into the optimization procedure in an attempt to provide a form of "insurance" against

---

1 See Conlisk, J., "Choice of Response Functional Form in Designing Subsidy Experiments", Econometrica, Vol. 41, No. 4, 1973, for a further discussion of the relative efficiencies of alternative design functions.

incorrect specification of the response.[1] Using the unconstrained allocation as a norm, the imposition of lower bound constraints led to a 6.9% reduction in the theoretical efficiency of the sample design, relative to the specified response function.[2]

(ii) External Constraints on Sample Availability and Budget Shares

Difficulties in finding sufficient numbers of households possessing certain characteristics led to the placement of upper bound constraints on four of the socio-economic strata. The four strata include the bottom two income strata for double-headed multiple earner families and the second and third income strata for single unattached individuals. The imposition of the household availability constraints reduced the sample size by 4%. The upper bound constraints represent a further 7% reduction in theoretical efficiency of the sample design, relative to the specified response function.

---

1 An alternative approach to specifying treatment minima constraints was also tried, namely, requiring each cell to have at least 50% (25% in the case of above-breakeven and control families) as many observations as it would receive under an analysis of variance allocation. The adopted approach whereby a minimum of two observations per cell was imposed is far more restrictive. Indeed, the allocation resulting from simultaneous application of the two constraints above was very similar to that obtained from the second taken by itself.

2 The treatment minima approach in the assignment model may be viewed essentially as a two-step hedge process against mis-specification subject to two constraints, one on the acceptable probability of Type II error and one involving the minimization of Type I error. The approach adopted by the Mincome Manitoba assignment model was suggested by Perry Gluckman and is based upon (1) the principle of two samples per stratum design outlined in G. Elfving, "Optimum Allocation in Linear Regression Theory", Annals of Math. Stat., 1950, and (2) the possibility of estimating the standard error of regression coefficients in two samples per strata designs employing jackknifing procedures.

Finally, certain constraints on budget shares were specified in order not to place too much emphasis on non-payment (that is, above breakeven or control) families. Because non-payment households are relatively inexpensive a modest budget share can still accommodate a sizeable sample of controls and above breakeven families.[1] Consequently, arbitrary budget share restrictions that not more than 5% of the budget was to be absorbed by above breakeven households, and not more than 7% of the budget to be absorbed by control households were imposed. The resulting allocation substituted non-payment families in favour of below breakeven households on a five-to-one ratio and produced an incremental 3.8% reduction in theoretical efficiency of the sample design relative to the specified response function.[2]

IV.6 Design Point Costs and Budget

The optimization procedure of the formal assignment model involves minimizing an objective function subject to the following constraint:

$$(4.4) \qquad \sum_{i=1}^{162} c_i n_i \leq C \quad ,$$

---

[1] For example, prior to the imposition of the budget share constraints, the assignment model allocated 58% of the sample to non-payment groups although this portion of the sample absorbed only 16.5% of the budget. The subsequent allocation incorporating the budget share restrictions is also illustrative of the differential relative costs of observations. The restrictions led to a reduction in sample size of 174 observations in the non-payment group which permitted an increase of only 34 families below the breakeven level.

[2] Also, because of the constraint that the controls absorb not more than 7% of the budget, the weights placed on the controls outlined earlier become important in relative terms only and not in actual value.

where c_i is the cost of one observation at the i^{th} design point, n_i is the number of sample observations assigned to the i^{th} design point (the unknowns), and C is the total budget constraint.

The budget constraint established for the Winnipeg Site sample was \$4 million. The cost function, c_i, for the i^{th} design point, is described more fully in Appendix A.1. The cost function specifications include:

(1) Costs that vary with sample size, but not with attrition or response to treatment. These costs cover the variable cost components of the pre-experimental interviews (screener and baseline) and enrollment.

(2) Costs that vary with sample size and attrition, but not with response to treatment. The primary component here is the cost of periodic interviews for the duration of the experiment.

(3) Costs that vary with hypothesized labour supply response and sample size, before adjustment for attrition.

(4) Hypothesized retention rates for families in cell i for the duration of the experiment.

(5) Payments cost per year for a family in cell i <u>before</u> adjusting for attrition, but after adjustment for hypothesized labour supply response.

(6) Variable components of administrative costs made in payments system as a proportion of total payments made.

(7) Duration of experiment with regards to payments and surveys.

(8) Guarantee level and negative tax rate for i^{th} cell.

(9) Average positive income tax rate for the i^{th} cell.

(10) Proportion of experimental time period that the average attriting unit remains in the experiment.

IV.7 Final Allocation of Assignment Model

The final allocation of the assignment model for the Winnipeg Site has the following features:

(1) The total number of observations (required enrollees) is 1014.

(2) The distribution of the sample by family type is
58.5% of the households are double-headed
15.7% of the households are single-headed
25.8% of the households are single unattached households.

(3) The distribution of the sample by breakeven and control status is
50.9% of the households are allocated to below breakeven cells
9.3% of the households are allocated to above breakeven cells
39.9% of the households are allocated to controls.

Full details on the distribution of the final sample assignment are set forth in Appendix A.2. In addition, the distribution of enrollment completions for the Winnipeg site is given in Appendix A.5.

V. The Rural Sample Segment of the Experiment

V.1 The Rural Saturation Site

There are distinct advantages as well as shortcomings in randomly dispersed samples designed to measure behavioural responses in an experimental setting. On the other hand, special problems are associated with the strategy of saturation sampling wherein all households normally resident in a selected community are deemed eligible for participation in the income maintenance program for the duration of the experiment.[1] Prominent among the many advantages of a saturation sampling strategy is the fact that the experimental environment created is less artificial than that achieved by dispersed sampling methods. In recognition of the potential benefits of information obtained from saturation sampling, the Manitoba Basic Annual Income Experiment included a single saturation site (Dauphin, Manitoba) in addition to the now standard dispersed sampling approach.

The saturation site provides for a single treatment plan (G = 3800, t = .50) since any horizontal inequity introduced would quickly become communicated to all members in the community. All households who are normally resident in the saturation site are eligible for income maintenance payments at any time during the course of the experiment. However, periodic interview information is

---

1 A brief discussion of the relevance and methodological issues involved with the saturation component of the experiment is contained in Chapter II, Section 7 of this report: "Special Experimental Design Issues". Further details are available in "The Objectives and Design of the Manitoba Basic Annual Income Experiment" (in preparation).

collected only for a proper subset; namely those having estimated normal annual income less than $9000, prior to the start of the experiment, not disabled or otherwise ineligible, etc.[1]

V.2 Specifications of the Mincome Manitoba Assignment Model: Rural Dispersed Sites

In addition to the dispersed sample in Winnipeg and the saturation site of Dauphin, an additional $1.3 million was budgeted for the inclusion of a rural dispersed sample which could provide an analytical bridge between the saturation and Winnipeg samples. The rural dispersed sample would include a single treatment plan with a guarantee level of $3800 (for a family of four) and a 50% offset tax rate. This treatment is the same as that available in the saturation site and is among those available for the Winnipeg sample. The rural dispersed sample would also include a control group.

Since only one payment treatment was being considered, the identification of a response function as specified for the Winnipeg sample is not relevant. The allocation of the rural dispersed sample was therefore based upon a dummy variable approach. The rural dispersed sample employed the same eighteen socio-economic cell definitions as the Winnipeg sample but, having only one treatment and one control group, results in 36 distinct cells or design points as opposed to 162 cells

---

[1] In addition to the residency requirement, citizenship or landed immigrant status was required for eligibility for payments. Without prejudice to their eligibility status for payments, certain individuals were excused from participation in periodic interviewing because they were thought to have little work response behaviour of research interest. The totally disabled and elderly were excluded on these grounds.

for the Winnipeg sample. A cell or sample point for the rural dispersed sites consequently consists of the triplet (E^*, f, d), where d is a dummy indicating treatment or control cell, E^* is level of normal income, f is the family type. In all cases, the treatment minima constraint of two observations was enforced for each cell, except where the sample availability constraints took precedence. The family type policy weights, the income class weights and the pyramidal weights were identical to those used in the Winnipeg allocation model. The control cells as well as the above-breakeven cells received 25% of the weight assigned to the payments cells (the below-breakeven cells). The cost function for the rural dispersed sites was the same as that used for the Winnipeg site except for inclusion of an additional term reflecting the incremental cost per initially enrolled unit attributable to the selecting a dispersed observation outside of the Winnipeg site. The final theoretical allocation is included in Appendix A.3 and reflects the sample availability constraints which had to be imposed because of the difficulties in obtaining sufficient numbers in several of the strata. The budget share constraint of not more than 7% to be allocated to controls and not more than 5% to the above-breakeven cells was not imposed for the rural dispersed sites. The distribution of enrollment completions for the rural dispersed sites is given in Appendix A.6.

VI. The Supplementary Portion of the Sample: Description, Design and Allocation

VI.1 Background

Shortly after the beginning of payments to experimental participant units in the Manitoba Basic Annual Income Experiment preliminary indications revealed an unacceptably small number of observations below the breakeven level. This meant that, even if there were sufficient variability in the amounts that low-income families would receive from the experiment, there might not be sufficient numbers of these families at the end of the experiment to estimate their labour supply response efficiently and with an acceptable degree of precision. In addition, the attrition rate proved higher than anticipated and if assumed to continue would have led to smaller treatment sample sizes than desired for research purposes. Further, additional information that became available led to certain desired design modifications and sample changes. More specifically, it was discovered that certain groups of households of policy relevance had been systematically excluded from selection and a correction was recommended. As well, sample size and other considerations led to a design decision to eliminate an ineffective treatment plan.[1]

---

[1] Design modifications involved more than simply specifying financial treatment plans characterized by a (G,t) combination. However in the context of the sample assignment issue the result is that one of the previously designated eight treatment plans was eliminated. This decision was based upon two factors. First the financial benefits of the experimental plan in question were slight when compared to the alternative available to the participant of attriting and accepting existing welfare payments. Second, given the size of the treatment sample and our estimate of the potentially available population it appeared unlikely that an augmented sample could be drawn at acceptable cost to yield the desired final treatment size. (Continued)

As a result of these and other considerations, the Manitoba Basic Annual Income Experiment designed and augmented portions of the sample. The supplementary sample was restricted to the urban dispersed site and confined to those household types and income strata for which the current treatment sample sizes were thought to be inadequate. The supplementary sample received experimental payments for a duration identical to that of the initial sample but commencing one elapsed calendar year after the first payments to the initial sample.[2]

VI.2 Selection of the Supplementary Sample

The supplementary sample involved re-sampling from parts of the original Winnipeg sampling frame in an attempt to strengthen the original sample with respect to both the completeness of the frame and the sample distribution relative to the assignment model quota of the original sample. As in any sequential design, the selection and allocation of the supplementary sample depends almost solely on the distribution relative to the assignment model allocation of the original Winnipeg sample at the time of the supplementary sample assignment.

---

1 (continued)

Plan 6 (G = 3800, t = .75) was eliminated and all observations were placed in Plan 7 (G = 4600, t = .75) to minimize noise associated with individuals changing support level and tax rate. The individuals therefore retained the same tax rate but received an upward adjustment of their support level. For the purposes of the supplementary sample allocation, the assignment model requirements for Plan 6 were transferred to Plans 7 and 8 (G = 5400, t = .75) in a 2-to-1 ratio. It should be noted that the elimination of Plan 6 does not preclude the possibility of estimating non-linearities associated with tax rates and guarantee levels.

2 The progress of the original sample through time, its status at the time of selection of the supplementary sample, the method and procedures of selection of the original and supplementary portions of the sample will be discussed in separate reports.

The supplementary sample was designed in such a way as to minimize the gap between the assignment model cell requirements and the projected cell distribution of the original sample one year later. This revealed that (i) the sample distribution after one year was most deficient in the first three income levels of each family type; (ii) the fourth income level for double-headed households was nearly filled; and (iii) the highest income level cells were overfilled. Consequently households in the first three income levels were chosen; the double-headed households in the fourth income level were chosen with a lower probability and no households were selected in the highest income level. Except for the original baseline recontact portion[1] of the supplementary sample (where all family types were chosen), only double-headed households were selected for enrollment in the supplementary sample.

In general, the assignment of treatments to observations within a given socio-economic stratum was, subject to the sample availability constraint within that stratum, directly proportional to the following two quantities:

(i) the number of sample points required in that treatment based upon the original assignment model allocation. This quantity was used as a measure of the relative importance of that cell as well as a measure of the cost of that design point, and

---

[1] The baseline interview was the first extensive pre-enrollment interview administered to households in the original sample. A portion of the supplementary sample consisted of those Winnipeg households who were selected for the baseline interview as part of the original panel but did not complete it (refusals, moves, not contacted, etc.). As this was essentially a recontact effort of the original sample, it was decided to select households from all family types if they met the income requirements.

(ii) the ratio of (i) above and the current cell size after supplementary sample augmentation. This was a measure of the gap between the observed and required numbers and also incorporates a measure of the expected attrition.

VI.3 Design and Allocation of the Supplementary Sample

The allocation of treatments to selected households was done separately for each family type. For single individuals, the allocation used was the same as that of the original sample, namely the assignment model probabilities. The allocation used for double-headed and single-headed households was a function of the following variables (within each family type):

R_{ij} -- the number of sample points required in treatment i and income level j, based upon the original allocation of the assignment model.

N_{ij} -- the number of sample points from the original sample who were in income cell j, with treatment i at time of supplementary sample assignment.

T_{ikm} -- proportion of the sample points in income cell k (at time of enrollment), who were in income cell m and treatment i at time of supplementary sample assignment ($k \leq m$). If $m < k$, m is set equal to k.[1]

---

1 Setting m equal to k when m was less than k was done on operational considerations. Very little loss of generality resulted from this assumption of no misclassification to a lower income level. The proportion of observations in any normal income cell where a misclassification to a lower income cell occurred ranged from zero to a maximum of 17%, the average being 6%.

O_{ij} -- number of sample points in normal income cell j who are to be assigned to treatment i in the allocation process (i.e., the unknowns).

S_j -- total number in income cell j who are available for the supplementary sample.

Specifically, the required assignment is that allocation which, for family type m and normal income cell j, minimizes

$$\phi = \sum_{k=1}^{9} \frac{R_{kj}^2}{N_{kj} + \sum_{n=1}^{j} (T_{knj} \cdot O_{kn})} \tag{6.1}$$

subject to the sample availability constraints

$$S_j = \sum_{i=1}^{9} O_{ij} \quad . \tag{6.2}$$

Employing the calculus method of Lagrange multipliers this requires selecting O_{kn} and the multiplier λ so as to minimize

$$\begin{aligned} & \phi + \lambda(S_j - \sum_{i=1}^{9} O_{ij}) \\ & = \sum_{k=1}^{9} \frac{R_{kj}^2}{N_{kj} + \sum_{n=1}^{j} (T_{knj} \cdot O_{kn})} + \lambda(S_j - \sum_{i=1}^{9} O_{ij}) \quad . \end{aligned} \tag{6.3}$$

If we let g_i be the gap or shortfall for treatment i in income cell j as a proportion of its target after augmentation, the optimal allocation of the sample points (the O_{ij}'s) can be shown, by differentiating (6.3) with respect to O_{ij} and λ, to be the allocation satisfying:

$$(6.4) \qquad \frac{\sqrt{T_{1jj}}}{(1-g_1)} = \frac{\sqrt{T_{2jj}}}{(1-g_2)} = \cdots\cdot = \frac{\sqrt{T_{9jj}}}{(1-g_9)} \quad .$$

An additional constraint of not more than 35% of the sample points to be allocated to controls was imposed for double-headed households. Appendix A.4 gives the theoretical allocation of the supplementary sample by plan, family type and level of normal income. The distribution of enrollment completions in the supplementary sample is set forth in Appendix A.7.

VII. Conclusion

A distinguishing characteristic of the Manitoba Basic Annual Income Experiment is its interest in estimating the behavioural response of households to a range of tax and guarantee levels over the entire design space. This is in contrast to a demonstration or pilot project whose aim is usually to establish the feasibility or best administrative procedures of a given prototype program. Because of the substantial costs of observations involved in multi-year negative tax experiments, both in terms of payments as well as data collection, the issue of sample size is extremely important, given that every experiment must operate under a budget constraint. Further, negative tax experiments exhibit extreme cost variations per observation.

The assignment model employed by the Manitoba Basic Annual Income Experiment incorporated specific assumptions and constraints. The resulting optimal design is seen to depend upon the relative costs of observations at different design points, the overall budget constraint, the hypothesized response surface and the objective function specified. Associated with the gain in efficiency of any optimal design is some risk of mis-specification and the fact that a design optimal with respect to one objective may be non-optimal with respect to other considerations. The assignment model insures that such decisions are made in a structured manner on the basis of explicitly stated assumptions and carefully specified objectives.

Appendix A.1 The Cost Function for the Assignment Model of the Winnipeg and Rural Dispersed Sites

Implications of Costs for the Assignment Model

The purpose of the assignment model is to optimize the allocation of dispersed sample points among the various experimental cells given that observations placed in different cells yield different levels of benefits with respect to the objective function and represent commitments to incur different levels of costs. It is important to estimate accurately both the magnitude of these costs per cell and their relative values. Errors in the estimation of magnitude imply either a smaller sample than could have been purchased (cost overestimated) with the resulting reduction in the statistical precision of our results, or a cost overrun (cost underestimated) with its attendant undesirable consequences. An error in the relative costs across cells implies a non-optimal allocation of sample points and hence the inefficient use of scarce experimental funds.

Following is a detailed description of the cost function for the assignment model.

Notation and Definitions

a_1 = Costs per initially enrolled unit that vary with respect to sample size, but <u>not</u> with respect to either attrition or response to treatment. These costs cover the variable cost components of screening, baseline, and enrollment.

a_2 = costs per initially enrolled unit that vary with sample size and assumed attrition, but not with respect to response to treatment. The primary component here is the cost of periodic interviews for the duration of the experiment.

a_3 = incremental cost per initially enrolled unit attributable to selecting a dispersed observation outside of Greater Winnipeg.

a_4 = costs per initially enrolled unit that vary with sample size and assumed labour supply response, before adjustment for attrition.

P_i = payments cost per year for a family in cell i <u>before</u> adjusting for attrition, but after adjustment for labour supply response.

R_i = retention rate for families in cell i for the duration of the experiment.

r = proportion of the experimental time period that the average attriting unit remains in the experiment.

C_i = total variable cost per initially enrolled family in cell i <u>after</u> adjustments for attrition and labour supply response.

B_i = breakeven level of income for i^{th} cell.

S_i = annual support level for cell i (when annual income is 0).

t_i = offset tax rate for cell i.

E_i = normal annual income for units in cell i after NIT.

E_i^* = normal annual income for units in cell i before NIT.

L - locational dummy variable = 0 if unit is in Winnipeg
= 1 if unit is outside Winnipeg

F = expenditure on research and follow-up per attrited unit to allow analysis of attrition bias.

h = variable components of accounting costs in payments system as a proportion of total payments made.

d = duration of experiment in years with regard to payments and surveys.

z_i = the average positive income tax rate for the i^{th} cell (based on 1973 Manitoba positive income tax rates and basic exemptions).

Cost Function of the Assignment Model

(1) General Specification of Cost Function

(i) $$C_i = a_1 + R_i(a_2 + a_4) + (1 - R_i)[r(a_2 + a_4) + F] + a_3L$$

where:

(ii) $$a_4 = (1 + h)d \cdot P_i$$

(iii) if $E_i \leq B_i$ then $P_i = \text{Max.}\{[(S_i - tE_i)^+ + z_iE_i], \$120\}$

if $E_i \geq B_i$ then $P_i = \text{Max.}\{[z_iE_i - t(E_i - B_i)], \$120\}$

$P_i = \$60$ for control families

(2) Specific Relationships and Values Employed

(a) Cost Implications of Attrition

In addition to creating problems for the interpretation of the data, the presence of attrition has cost implications which affect sample design. First, an attrited observation provides less than full information. Second, the presence of attrition lowers the budgetary cost of each initially assigned family but raises the cost of each usable sample point. Third, attrition assumptions which vary by household or treatment characteristics change the relative costs of observations in alternative cells and consequently affect the optimal assignment of sample observations.

Theory as well as empirical evidence would suggest that attrition is correlated with the payment levels associated with each cell. Accordingly, the process of relating retention rates in the design model to expected payment costs appears reasonable. The specific relationship of retention rate to expected payments employed was:

$$\text{(iv)} \quad R_i = .88 + .08 \text{ Min.}[P_i/1000, 1] \ .$$

Estimated Retention Rates

Payment	R_i
$ 0 (control)	.880
100	.888
300	.904
500	.920
700	.936
1000 and over	.960

(b) Cost Implications of Work Effort Response

The labour supply response resulting from the experimental stimulus can also affect the magnitude and pattern of costs. Specifically, the greater is the work disincentive for a given family, the lower will be their income from employment during the experiment and consequently the more costly will be the observations since payments to the unit are income-conditioned. For this reason some estimate of the work response must be incorporated into the cost function of the assignment model.

Several assumptions are possible. For example a household may take all payment benefits in the form of increased money income in which case there would be no work response (zero response case). Alternatively, a household may take all of its benefits in the form of leisure, thereby choosing its post-NIT income level so as to maintain its level of pre-experimental normal income (target income response). Define the work effort ratio, E_i/E_i^* , as the fraction of its pre-experimental work effort a family in the i^{th} cell displays after imposition of the treatment. The work effort ratio is simply unity in the case of zero work response and is assumed to be a function of S, t, and E^* in the case of the target income response.

After considering alternative specifications and given the information available, the Manitoba Basic Annual Income Experiment adopted a simple linear formulation relating post-NIT earnings to benefits received, the tax rate, and pre-experimental normal income. In effect, the estimated work disincentive is a function of the supplement, $S_i - t_iE_i^*$, and tax rate, t_i . The specific relationship employed as well as the associated estimated values of the work effort ratio, E_i/E_i^* , is as follows:

(v) $E_i = E_i^* - \alpha[S_i - t_iE_i^*] - \beta t_i$

For multiple earner families: $\alpha = .30$, $\beta = 1200$

For single earner families: $\alpha = .20$, $\beta = 800$

Subject to the constraint:

$$1.0 \geq \frac{E_i}{E_i^*} \geq (1 - t_i) \quad .$$

Estimated Values for E_i/E_i^*

Single Earner Units

$\alpha = .20$

$\beta = 800$

S_i	t_i	S_i/t_i	Normal Income Before NIT(E_i^*) $2000	$4000	$6000	$8000
4000	.3	13,333	.700[1]	.800	.887	.930
4000	.5	8,000	.500	.800	.900	.950
4000	.7	5,714	.460	.800	.913	.970
5000	.5	10,000	.500[1]	.750	.867	.925
5000	.7	7,142	.360	.750	.880	.945

Multiple Earner Units

$\alpha = .30$

$\beta = 1200$

S_i	t_i	S_i/t_i	Normal Income Before NIT(E_i^*) $2000	$4000	$6000	$8000
4000	.3	13,333	.700[1]	.700	.830	.895
4000	.5	8,000	.500[1]	.700	.850	.925
4000	.7	5,714	.300[1]	.700	.870	.955
5000	.5	10,000	.500[1]	.625	.800	.888
5000	.7	7,142	.300[1]	.625	.820	.918

[1] These values represent the value of constraint

$= 1 - t_i$

(c) Other Values Employed in Cost Function

$h = .03$

$F = \$200$ per attrited unit

$d = 3.2$ *

$r = .5$

$a_1 = \$200$ per initially enrolled unit

$a_2 = \$400$ per initially enrolled unit

$a_3 = \$440$ per initially enrolled unit outside of Winnipeg

* Assumes a phase-out period and a 7% per year rate of cost increase due to inflation.

Appendix A.2 Theoretical Sample Allocation for the Winnipeg Site

Socio-Economic		Plan (G/t)								
Family Type	Level of Normal Income	3800/.35	4600/.35	3800/.50	4600/.50	5400/.50	3800/.75	4600/.75	5400/.75	Controls
DHME	1300	.0	1.0	3.2	.0	.6	1.6	.0	.8	3.8
	4000	2.0	2.0	13.4	2.0	2.4	11.3	2.0	4.5	13.3
	6000	6.4	2.0	21.0	2.0	3.4	3.2	10.6	5.6	51.9
	8000	2.0	2.0	2.0	11.3	2.6	3.5	2.0	2.0	12.8
	11000	2.0	3.6	2.0	2.0	2.0	3.8	2.0	9.0	15.9
DHSE	1300	2.0	3.0	14.0	2.0	2.1	6.0	2.0	2.0	50.0
	4000	2.0	2.0	11.2	2.0	3.9	15.8	2.0	5.4	10.3
	6000	7.5	2.3	26.7	2.0	2.0	7.0	14.0	6.3	55.5
	8000	2.0	2.0	2.0	11.3	3.3	2.0	2.0	2.0	22.3
	11000	2.0	3.3	2.0	2.0	2.0	2.0	2.0	2.0	13.2
SH	400	2.0	2.0	6.2	2.0	2.0	2.0	2.0	2.0	17.5
	2000	2.0	2.0	4.6	2.0	2.0	7.7	2.0	2.4	9.8
	4500	2.6	2.0	15.6	2.0	2.0	3.2	6.4	2.8	31.4
	8000	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
SI	400	3.2	3.0	17.9	2.0	3.1	7.7	2.2	2.9	33.8
	1500	2.0	4.3	13.8	2.0	5.3	13.7	2.0	6.2	24.7
	2500	5.7	2.0	18.5	8.4	2.5	3.1	2.0	4.5	27.1
	4000	7.4	4.7	2.0	2.0	5.5	2.9	2.0	2.0	8.8
Total by Plan		55.1	45.5	178.2	59.3	48.8	98.7	59.3	64.5	404.0

Total Number of Observations = 1013.5

DHME = Double-Headed Multiple Earner
DHSE = Double-Headed Single Earner
SH = Single-Headed
SI = Single Individual

Appendix A.3 Theoretical Sample Allocation for the Rural Dispersed Sites

Socio-Economic Cell		Experimental Plan	
Family Type	Normal Income	Treatment	Control
Double-Headed Multiple Earner	1300	2.0	.0
	4000	6.0	2.0
	6000	24.1	14.1
	8000	6.3	12.5
	11000	8.9	13.0
Double-Headed Single Earner	1300	13.6	11.4
	4000	26.0	17.9
	6000	31.2	17.9
	8000	7.2	14.2
	11000	6.3	9.0
Single-Headed	400	3.6	2.8
	2000	12.9	5.1
	4500	8.5	2.5
	8000	2.0	2.0
Single Individual	400	13.2	4.8
	1500	2.0	2.0
	2500	2.0	2.0
	4000	3.4	3.6
Total by Plan		179.2	136.8

Total Number of Observations = 316.0

Appendix A.4 Theoretical Allocation of the Supplementary Sample

Socio-Economic Cell		Experimental Plan								
Family Type	Level of Normal Income	3800/.35	4600/.35	3800/.50	4600/.50	5400/.50	4600/.75	5400/.75	Controls	Total
Double-Headed	1300		1	6		1	1	4	8	21
	4000	2	2	20	1	2	8	6	20	61
	6000	7		42		6	13	9	44	121
	8000				79				14	93
Single-Headed	400	2	1	2	2	1	2	4	13	27
	2000		1	1		1	2		1	6
	4500			1			1	1	5	8
Single Individual	400			1						1
	1500							1	1	2
	2500	1			1					2
	4000	2								2
Total by Plan		14	5	73	83	11	27	25	106	344

Appendix A.5 Distribution of Enrollment Completions for the Winnipeg Site

Socio-Economic Cell		Plan* (G/t)								
Family Type	Level of Normal Income	3800/.35	4800/.35	3800/.50	4800/.50	5800/.50	3800/.75	4800/.75	5800/.75	Controls
Double-Headed	1300	3	4	9	6	6	5	2	4	31
	4000	5	4	13	7	6	15	5	11	18
	6000	9	6	23	8	7	8	10	4	61
	8000	9	15	6	41	16	10	9	11	59
	11000	12	15	1	10	10	11	4	14	47
Single-Headed	400	4	3	8	5	4	3	2	2	24
	2000	2	4	5	3	2	10	2	3	15
	4500	5	2	16	3	3	4	7	4	29
	8000	5	4	3	3	3	2	4	5	6
Single Individual	400	2	4	16	2	4	5	1	1	30
	1500	10	9	13	3	10	8	5	6	20
	2500	10	7	14	11	2	4	3	4	17
	4000	11	8	6	4	6	0	4	2	13
Total by Plan		87	85	133	106	79	85	58	71	370

Total Number of Observations = 1074

Furthermore, 5 additional families were enrolled as a result of residing with a selected household. Two of these were enrolled in plan (4800/.35), one in plan (4800/.50), and the remaining two were enrolled in plan (5800/.50).

* represents support levels at time of enrollment

Appendix A.6 Distribution of Enrollment Completions for the Rural Dispersed Sites

A. Non-Farm Households

Socio-Economic Cell		Experimental Plan	
Family Type	Normal Income	Treatment	Control
Double-Headed Multiple Earner	1300	8	0
	4000	6	2
	6000	5	7
	8000	9	17
Double-Headed Single Earner	1300	3	3
	4000	7	3
	6000	15	10
	8000	5	13
Single-Headed	400	13	6
	2000	11	2
	4500	6	3
	8000	5	6
Single Individual	400	7	3
	1500	3	1
	2500	0	0
	4000	0	2
Total by Plan		103	78

Total number of enrolled non-farm observations = 181 + one new treatment family created at enrollment as a result of a non-head marriage = 182

B. There were also 88 farm households enrolled as controls

Total number of enrolled households in the rural dispersed sites = 182 + 88 = 270

Appendix A.7 <u>Distribution of Enrollment Completions in the Supplementary Sample</u>

Socio-Economic Cell		Experimental Plan								
Family Type	Level of Normal Income	3800/.35	4800/.35	3800/.50	4800/.50	5800/.50	4800/.75	5800/.75	Controls	Total
Double-Headed	1300			5		1	1	2	6	15
	4000	2	2	14	1	1	7	6	19	52
	6000	6		36		5	10	8	40	105
	8000				68				13	81
Single-Headed	400	1	1	1	1	1	2	3	12	22
	2000			1		1	2		1	5
	4500			1			1	1	5	8
Single Individual	400									0
	1500							1	1	2
	2500	1								1
	4000	2								2
Total by Plan		12	3	58	70	9	23	21	97	293